



**HAL**  
open science

# INTERACTIONS GENE-ENVIRONNEMENT DANS L'ETIOLOGIE DES FENTES ORALES

Cécile Chevrier

► **To cite this version:**

Cécile Chevrier. INTERACTIONS GENE-ENVIRONNEMENT DANS L'ETIOLOGIE DES FENTES ORALES. Santé publique et épidémiologie. Université Paris Sud - Paris XI, 2005. Français. NNT: . tel-00669882

**HAL Id: tel-00669882**

**<https://theses.hal.science/tel-00669882>**

Submitted on 14 Feb 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**UNIVERSITE PARIS XI**  
**FACULTE DE MEDECINE PARIS-SUD**

Année 2005

N° attribué par la bibliothèque



**THESE**

pour obtenir le grade de  
**DOCTEUR DE L'UNIVERSITE PARIS XI**  
Spécialité : Epidémiologie

présentée et soutenue publiquement

par

Cécile CHEVRIER

le 4 octobre 2005

**INTERACTIONS GENE-ENVIRONNEMENT**

**DANS L'ETIOLOGIE DES FENTES ORALES**

Directeur de thèse : Mme Sylvaine CORDIER

**JURY**

M. Denis HEMON, Directeur de recherche (Inserm, Paris, France)

*Président*

Mme Laurence TIRET, Directeur de recherche (Inserm, Paris, France)

*Rapporteur*

M. Julian LITTLE, Professeur, *Canada Research Chair in Human genome Epidemiology*

(Université d'Ottawa, Ontario, Canada)

*Rapporteur*

Mme Janine GOUJARD, Directeur de recherche (Inserm, Paris, France)

M. Pierre DRUILHET, Enseignant-chercheur (Ensai, Rennes, France)

Mme Sylvaine CORDIER, Directeur de recherche (Inserm, Rennes, France)



## *Remerciements*

Je remercie Bernard Jégou, Directeur du Groupe d'Etude de la Reproduction chez l'Homme et les Mammifères (U625 Inserm), de m'avoir accueillie dans son laboratoire durant ces trois années de travail de thèse.

Je remercie sincèrement Sylvaine Cordier qui, outre sa qualité scientifique, a toujours su m'accorder de son temps, une écoute attentive, des réflexions justes et très constructives, et une confiance extrêmement valorisante.

Ce travail est le résultat de toute une équipe de personnels chercheurs et hospitaliers : Michel Bahuaud ; Elisabeth Robert-Gnansia, Janine Goujard, Christine Francannet, pour leur travail de coordination ; Agnès Nelva, Christine Herman, Sylvie Maisonnial, Dr Bonnet, and Dr Dye pour la collecte des données ; Claire Perret, Richard H Finnell, Huiping Zhu, Edward Lammer, David Iovannisci pour la détermination des polymorphismes ; Brigitte Dananché pour son expertise d'évaluation des expositions professionnelles ; Pr Vasquez, Pr Beziat, Dr James-Pangaud, Dr Paulus, Pr Raphaël et Pr Mondié pour le recrutement des enfants atteints de fente orale ; Pr Chatelain, Pr Dodat, Pr Malpuech, Pr Jouk et les maternités de Tenon et Rothschild (Pr Uzan) pour le recrutement des enfants témoins ; Pr Potier de Courcy et Pr Hercberg pour leur contribution et leur aide à l'élaboration du questionnaire alimentaire ; Dr Cambien, Pr Morel et Dr Creveaux pour leur aide à l'analyse de biologie moléculaire ; et enfin, toutes les familles qui ont choisi de participer à cette étude.

Parmi l'ensemble des participants à cette enquête épidémiologique, je souhaite remercier plus particulièrement Claire Perret pour sa disponibilité, sa patience et sa minutie dans son travail et tous nos échanges, Elisabeth Robert-Gnansia pour sa gentillesse, ses conseils et ses encouragements, Janine Goujard d'avoir accepté si naturellement son rôle d'évaluation de ce travail de thèse en tant que membre du jury, Richard H Finnell et Edward Lammer pour leur investissement spontané dans ce projet et leur générosité.

Je tiens à remercier Catherine Bonaïti-Pellié et Marguerite Guiguet de m'avoir soutenu et encouragé dès l'amorce de ce projet de thèse. Je remercie Clarice Weinberg et Richard Morris pour leur disponibilité et leur conseil technique.

Un grand merci à tous mes collègues, permanents et de passage, de bureau et d'équipe, qui sans le savoir ont amplement participé à l'élaboration de ce travail : Christine, Anne-Marie, Céline, Cristina, Luc, Johanne, Odile, Véronique, Florence, Laurence, Catherine et beaucoup d'autres.

Je remercie vivement Laurence Tiret et le Professeur Julian Little pour leur rigueur et leur excellence dans le travail d'évaluation et de relecture, en tant que rapporteurs de ma thèse. Un grand merci à Denis Hémon pour son enthousiasme à assurer le rôle de Président de jury pour ce travail de thèse.

Je remercie Pierre Druilhet qui depuis quelques années m'a toujours suivi et soutenu dans mes projets professionnels, d'avoir accepté de participer à l'évaluation de ce travail.

Enfin, je souhaite conclure ces remerciements en exprimant ma sincère et profonde reconnaissance pour trois de mes enseignants et instituteurs : Laurent Augier, professeur d'économie à l'université de La Rochelle, Madame Sardet, enseignante d'histoire-géographie au collège de Goise à Niort, Monsieur Benoît, instituteur à l'école Jean Jaurès à Niort.

## *Financements*

Ce travail de thèse a été soutenu par deux programmes de financement :

- Inserm (Institut National de la Santé Et de la Recherche Médicale) : Poste d'accueil pour Ingénieur de Grandes Ecoles (novembre 2002-octobre 2003 ; novembre 2003-octobre 2004)
- FRM (Fondation Recherche Médicale) : Aide financière « Fin de thèse » (novembre 2004-octobre 2005)

L'enquête épidémiologique a été financée par le programme « Interactions entre les déterminants de la Santé » de l'Inserm et par le programme « Recherche en Environnement-Santé » du Ministère de l'Environnement.

Les génotypages réalisés au Texas (R.H.Finnell) ont bénéficié du soutien financier des *National Institutes of Health* (grants DE12898 and HL66398), et de l'agence *US Environmental Protection Agency* (Grant No. 82829201).

## *Productions scientifiques issues du projet de thèse*

### **Publication**

Chevrier C, Perret C, Bahuau M, Nelva A, Herman C, Francannet C, Robert-Gnansia E, Cordier S. Interaction between *ADH1C* polymorphism and maternal alcohol intake in the risk of nonsyndromic oral clefts: an evaluation of the contribution of the child and maternal genotypes. *Birth Defects Res A Clin Mol Teratol.* 2005; 73(2):114-22.

### **Articles en révision**

Chevrier C, Perret C, Bahuau M, Zhu H, Nelva A, Herman C, Francannet C, Robert-Gnansia E, Finnell RH, Cordier S. Fetal and maternal *C677T* mutation of *MTHFR* gene, maternal folate intake and the risk of nonsyndromic oral clefts. (Soumis à *American Journal of Medical Genetics: Part A*)

Chevrier C, Dananché B, Bahuau M, Nelva A, Herman C, Francannet C, Robert-Gnansia E, Cordier S. Occupational exposure to organic solvent mixtures during pregnancy and the risk of nonsyndromic oral clefts. (Soumis à *Occupational and Environmental Medicine*)

### **Communications scientifiques**

Chevrier C, Dananché B, Bahuau M, Francannet C, Robert-Gnansia E, Perret C, Lammer E, Cordier S. Exposition maternelle aux solvants et risque de fentes orales. Congrès international organisé par l'Association des Epidémiologistes de Langue Française (ADELF), Environnement et Santé, 21-23 septembre 2005 Québec (Canada)

Chevrier C, Perret C, Bahuau M, Zhu H, Nelva A, Herman C, Francannet C, Robert-Gnansia E, Finnell RH, Cordier S. Rôles de l'apport alimentaire en folates de la mère et du polymorphisme *C677T* de la *MTHFR*, sur le Risque de fentes orales. Rencontre Rennaise sur la Reproduction, 7-8 February 2005 Rennes

Chevrier C, Bahuau M, Francannet C, Robert-Gnansia E., Perret C, Cordier S. Oral Clefts, Maternal Exposure To Solvents And *Cyp2E1* Polymorphism. The 16th Conference of the International Society for Environmental Epidemiology, 1-4 Août 2004 New York (USA). *Epidemiology* 2004; 15(4): S184-5.

Chevrier C, Cordier S. Mesure d'association génétique avec le risque de fentes orales. Congrès Biométrie et Epidémiologie. Société Française de Biométrie (SFB) et Association des Epidémiologistes de Langue Française (ADELF), 15-16 Septembre 2003 Lille

Chevrier C, Robert-Gnansia E, Bahuau M, Francannet C, Nelva A, Herman C, Guiguet M, Cordier S. Influence des polymorphismes du gène de l'*ADH3* sur le risque de fentes orales chez les enfants de mères exposées à l'alcool et à ses dérivés. Congrès de la Société Française de génétique Humaine (SFGH), 3-4 Avril 2003 Clermont Ferrand

# Sommaire

<b>Sommaire</b>	<b>5</b>
<b>INTRODUCTION GENERALE</b>	<b>9</b>
<b>Introduction</b>	<b>11</b>
A.	Définition des fentes orales et notions d'embryologie 11
B.	Fréquence des fentes orales 11
B. 1.	Mesure par la prévalence 14
B. 2.	Hétérogénéité des fentes orales 14
B. 3.	Variations temporelles de la prévalence 15
B. 4.	Variations géographiques de la prévalence 15
C.	Conséquence et problème de Santé Publique 16
<b>L'étiologie complexe des fentes orales</b>	<b>18</b>
A.	Epidémiologie environnementale des fentes orales 18
A. 1.	Sexe et facteurs socio-démographiques 18
A. 2.	Epilepsie et antiépileptiques 18
A. 3.	Vitamines 19
A. 4.	Consommation maternelle de tabac 22
A. 5.	Consommation maternelle d'alcool 23
A. 6.	Expositions aux solvants organiques et aux pesticides 25
A.6.i.	Expositions professionnelles de la mère aux solvants organiques 25
A.6.ii.	Expositions environnementales aux solvants organiques 26
A.6.iii.	Expositions aux pesticides 26
A. 7.	Autres facteurs de risque potentiels 26
B.	Influences génétiques 27
B. 1.	Ce que les études de jumeaux nous apprennent sur la contribution génétique ? 27
B. 2.	Les études d'association génétique 28
B.2.i.	Transforming Growth Factor alpha <i>TGF<math>\alpha</math></i> 28
B.2.ii.	Transforming Growth Factor beta3 <i>TGF<math>\beta</math>3</i> 30
B.2.iii.	Homeobox <i>MSX1</i> 30
B.2.iv.	Autres gènes impliqués dans le développement embryonnaire 31
B.2.v.	MéthylèneTétraHydroFolate Réductase <i>MTHFR</i> 32
B.2.vi.	Autres gènes de métabolisme 33
B.2.vii.	Marqueurs anonymes 34
C.	Interactions gène-environnement 34
C. 1.	Maryland 34
C. 2.	Californie 35
C. 3.	Iowa 36
C. 4.	Danemark 37
C. 5.	Pays-Bas 37
C. 6.	Norvège 38
C. 7.	Conclusion 39
D.	Bilan 43

<b>Etat de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement</b>	<b>44</b>
A.	Randomisation mendélienne 45
A. 1.	Les lois de Mendel 45
A. 2.	Le concept de la randomisation mendélienne 46
A. 3.	Contribution de la randomisation mendélienne à l'épidémiologie 46
A. 4.	Limites de la randomisation mendélienne 47
B.	Mesure d'une association entre un gène et le risque d'une maladie 47
B. 1.	Approche cas-témoins 47
B.1.i.	Principe de la mesure d'une association 47
B.1.ii.	Modélisation de l'association : logistique ou log-linéaire. 49
B.1.iii.	Test de l'association et Intervalle de confiance 49
B.1.iv.	Biais de stratification génétique de population 50
iv.a.	Définition 50
iv.b.	Quel biais ? 50
iv.c.	Détection et mesure d'association en présence d'une stratification génétique de population 51
iv.d.	Quantification du biais et discussion 52
B. 2.	Approche cas-parents 53
B.2.i.	Association mesurée par un déséquilibre de transmission allélique 53
i.a.	Premières approches 53
i.b.	Les tests de déséquilibre de transmission allélique 54
i.c.	Extensions du TDT 56
i.d.	Hypothèses et Limites du TDT 57
B.2.ii.	Approche génotypique 57
ii.a.	Approche des pseudo-sibs 57
ii.b.	Méthodologie log-linéaire (Weinberg et al. 1998) 60
ii.c.	Hypothèses, principe de fonctionnement et limites 61
ii.d.	Extensions lorsque les génotypes des parents sont manquants 63
C.	Evaluation d'une interaction entre un gène et un facteur environnemental 64
C. 1.	Approche cas-témoins 64
C.1.i.	Modélisation logistique 64
C.1.ii.	Hypothèses, Limite, Puissance, Faisabilité 66
C.1.iii.	L'alternative Cas-seuls 67
iii.a.	Principe 67
iii.b.	Hypothèses, Limite, Puissance 68
C.1.iv.	L'alternative mixte : modélisation log-linéaire 69
C.1.v.	Autres alternatives 70
C. 2.	Approche cas-parents 70
C.2.i.	Méthodologie log-linéaire 71
C.2.ii.	Hypothèses, Limites et Puissance 72
D.	Tableau-bilan de l'évaluation d'une interaction gène-environnement 73
E.	Mesure des effets d'empreintes parentales 74
<b>Objectif du projet</b>	<b>75</b>
<b>MATERIEL &amp; METHODES</b>	<b>77</b>
<b>Population</b>	<b>79</b>
A.	Enquête épidémiologique 79
A. 1.	Recrutement des sujets 79
A.1.i.	Les cas 79

A.1.ii.	Les témoins	79
A. 2.	Collecte d'informations	80
A.2.i.	Questionnaire à la mère	80
A.2.ii.	Evaluation des expositions professionnelles	81
B.	Informations génétiques	82
B. 1.	Prélèvements biologiques	82
B. 2.	Génotypage	82
B. 3.	Effectifs des informations génétiques disponibles	84
B. 4.	Nomenclature	85
<b>Stratégie d'analyse</b>		<b>87</b>
<b>Description de la population</b>		<b>88</b>
A.	Description des populations de cas non syndromiques et de témoins	88
B.	Discussion des biais potentiels de l'enquête épidémiologique	90
B. 1.	Biais de sélection	90
B.1.i.	Le recrutement hospitalier	90
B.1.ii.	La non-réponse	91
B. 2.	Biais de classement	92
B. 3.	Biais de confusion	92
<b>RESULTATS</b>		<b>95</b>
<b>Rôle d'un polymorphisme du gène <i>ADH1C</i> en interaction avec la consommation maternelle d'alcool sur le risque de fentes orales non syndromiques</b>		<b>97</b>
A.	Matériel et Méthodes	98
A. 1.	Définition des expositions	98
A. 2.	Analyses statistiques	99
A.2.i.	Comparaison cas-témoins	99
A.2.ii.	Analyse des triades	99
B.	Résultats	100
B. 1.	Comparaison cas-témoins	100
B. 2.	Analyse des triades	101
B.2.i.	Rôle des génotypes <i>ADH1C</i> de l'enfant et de la mère	101
B.2.ii.	Empreinte parentale	101
B.2.iii.	Interaction entre la consommation maternelle d'alcool et les génotypes <i>ADH1C</i> de l'enfant et de la mère	102
B. 3.	Tableaux et figures	103
C.	Discussion	107
<b>Les génotypes <i>MTHFR C677T</i> de l'enfant et de la mère, l'apport maternel en folates et le risque de fentes orales non syndromiques</b>		<b>112</b>
A.	Matériel et Méthodes	113
A. 1.	Définition des apports maternels en folates	113
A. 2.	Analyse statistique	113
A.2.i.	Comparaison cas-témoins	113
A.2.ii.	Analyse des triades	114
B.	Résultats	114
B. 1.	Génotypes <i>MTHFR C677T</i>	115
B. 2.	Interaction Génotype <i>MTHFR C677T</i> et statut en folates de la mère	115

B. 3.	Tableaux et figures	116
C.	Discussion	121
<b>L'exposition professionnelle de la mère aux solvants organiques en interaction avec des gènes de métabolisme/détoxication, sur le risque de fentes orales non syndromiques</b>		<b>125</b>
A.	Matériel et Méthodes	127
A. 1.	Définition de l'exposition	127
A. 2.	Analyse statistique	127
B.	Résultats	128
B. 1.	Expositions maternelles aux solvants	128
B. 2.	Interaction gène-environnement	129
B.2.i.	Interaction statistique	129
B.2.ii.	Rôles propres des gènes de métabolisme	130
B. 3.	Tableaux	131
C.	Discussion	137
<b>CONCLUSION GENERALE</b>		<b>143</b>
<b>Résumé des résultats</b>		<b>145</b>
<b>Discussion du biais potentiel de stratification de population</b>		<b>147</b>
<b>Puissance statistique des tests d'interactions gène-environnement</b>		<b>149</b>
<b>Discussion du choix <i>a priori</i> des polymorphismes</b>		<b>154</b>
<b>BIBLIOGRAPHIE &amp; AUTRES REFERENCES</b>		<b>157</b>
<b>ANNEXES</b>		<b>175</b>
<b>Index des tableaux</b>		<b>177</b>
<b>Glossaire</b>		<b>179</b>
<b>Annexe 1 : Tests associés à l'estimation du maximum de vraisemblance</b>		<b>181</b>
<b>Annexe 2 : Causes d'hospitalisation des témoins</b>		<b>182</b>
<b>Annexe 3 : Complément des résultats du gène <i>ADH1C</i> sur le risque de fentes orales, en interaction avec l'exposition maternelle à l'alcool</b>		<b>183</b>
<b>Annexe 4 : Complément des résultats sur le gène <i>MTHFR</i> et l'apport alimentaire en folates</b>		<b>186</b>
<b>Annexe 5 : Complément des résultats sur les gènes <i>CYP2E1</i>, <i>GSTM1</i> et <i>GSTT1</i> et le risque de fentes orales</b>		<b>192</b>
<b>Annexe 6 : Effectifs du tableau 42</b>		<b>199</b>
<b>Annexe 7 : Puissance statistique <i>a posteriori</i> des tests d'association pour chaque facteur environnemental et chaque gène, par le logiciel QUANTO</b>		<b>200</b>

# **INTRODUCTION GENERALE**



# 1

## Introduction

Il est convenu de distinguer parmi les fentes orales deux groupes anatomocliniques correspondant à des processus de développement embryonnaire différents : les fentes labiopalatines (regroupant les fentes labiales isolées ou associées à une fente palatine), notées FL/P, et les fentes palatines isolées, FP.

### A. Définition des fentes orales et notions d'embryologie

Le développement de la face débute dès la 4<sup>ème</sup> et 5<sup>ème</sup> semaine de gestation. L'apparition d'une fente orale résulte d'une anomalie de fusion impliquant les bourgeons faciaux et/ou les processus palatins. A la fin de la 5<sup>ème</sup> semaine, les bourgeons faciaux se sont formés : il s'agit du bourgeon frontal, du bourgeon maxillaire, des bourgeons nasaux internes et externes (Figure 1). Au cours des deux semaines suivantes, le développement de la face progresse considérablement. Le processus maxillaire supérieur se développe et rétrécit la fente buccale d'origine (Figure 2). Les bourgeons nasaux internes se rejoignent et forment en partie intérieure le palais primaire. Les parties intérieures des processus maxillaires supérieurs sont nommées processus palatins. A ce stade du développement, une fente labiale peut se constituer en présence de perturbations dans la fusion du bourgeon nasal interne avec le processus maxillaire supérieur. Une fente labiale peut être uni- ou bi-latérale et peut être associée à une fente du palais primaire (communément appelée fente labiopalatine).

Une autre zone de fusion entre le bourrelet nasal externe et le processus maxillaire supérieur forme le sillon lacrymo-nasal, qui se prolongera jusqu'au coin interne de l'œil à la 10<sup>ème</sup> semaine pour former l'aile du nez. A la 7<sup>ème</sup> semaine de développement, les processus palatins qui se situent bilatéralement sous le bord de la langue (Figure 3), se redressent (Figure 4) et fusionnent au-dessus de la langue (Figure 5). Cette fusion forme le palais secondaire qui divise l'espace entre le conduit nasal et la cavité buccale définitifs. Une anomalie lors du redressement ou lors de la fusion des processus palatins engendre une fente palatine (partielle ou totale).

### B. Fréquence des fentes orales

Un rapport sur les anomalies craniofaciales issu d'un congrès organisé par l'Organisation Mondiale de la Santé (*Global registry and database on craniofacial anomalies, Report of a World Health Organization Registry Meeting on Craniofacial Anomalies, Brazil 2001*), estime, en moyenne, qu'un enfant sur 700 naissances vivantes dans le monde en 2001 serait atteint d'une FL/P ou d'une FP.

Figure 1: Vue frontale de la face. (A) Embryon de 5 semaines. (B) Embryon de 6 semaines.<sup>1</sup>

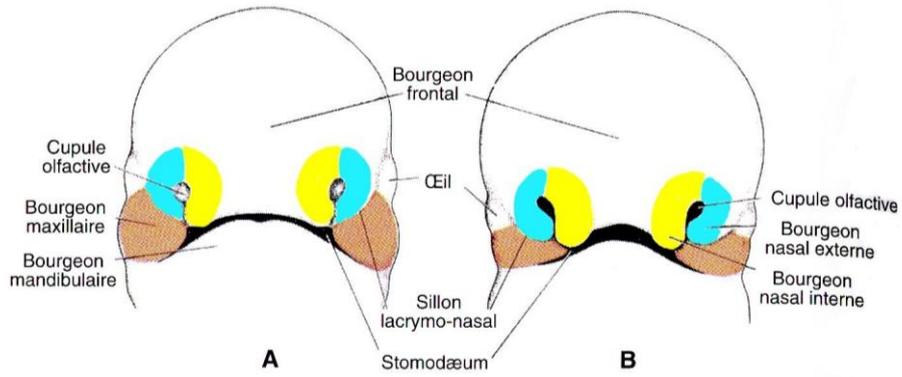


Figure 2 : Vue frontale de la face. (A) Embryon de 7 semaines. (B) Embryon de 10 semaines.<sup>1</sup>

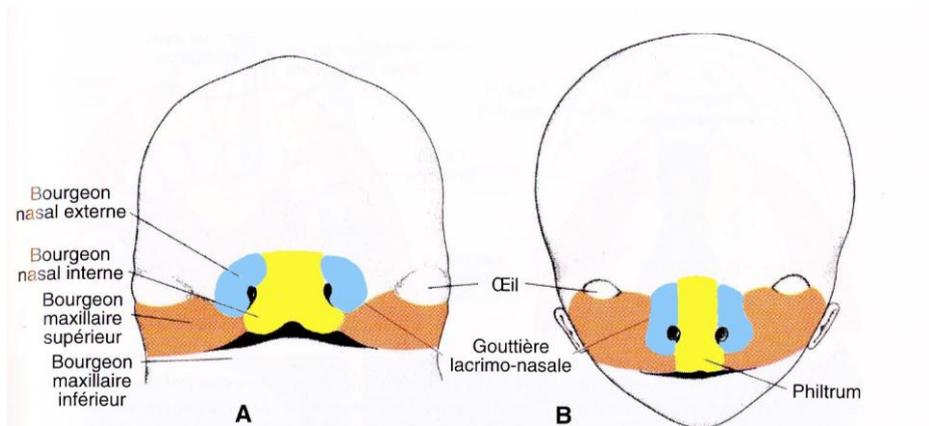


Figure 3: Embryon de 6 semaines et demie. (A) Coupe frontale de la région céphalique. (B) Vue centrale des processus palatins<sup>1</sup>

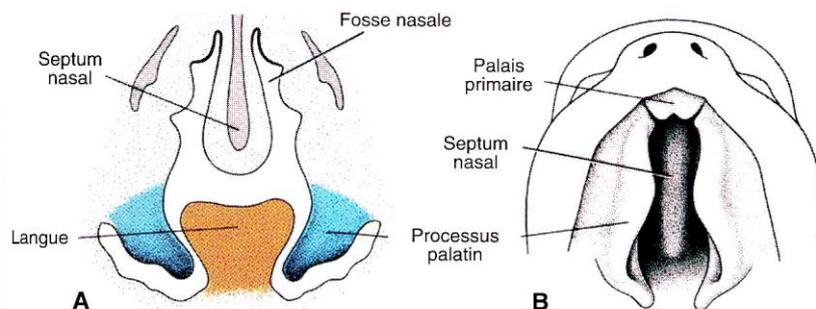


Figure 4: Embryon de 7 semaines. (A) Coupe frontale de la région céphalique. (B) Vue centrale des processus palatins<sup>1</sup>

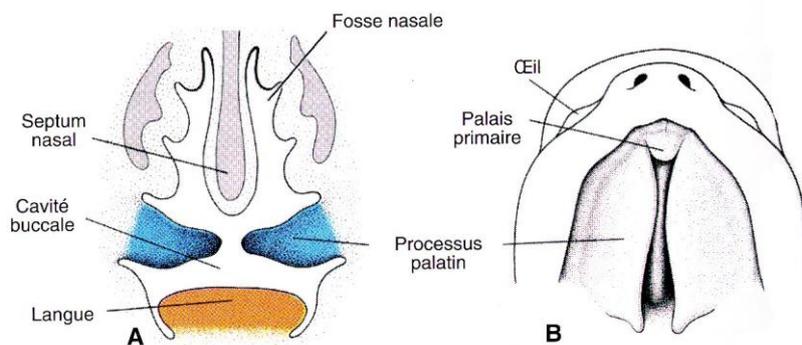
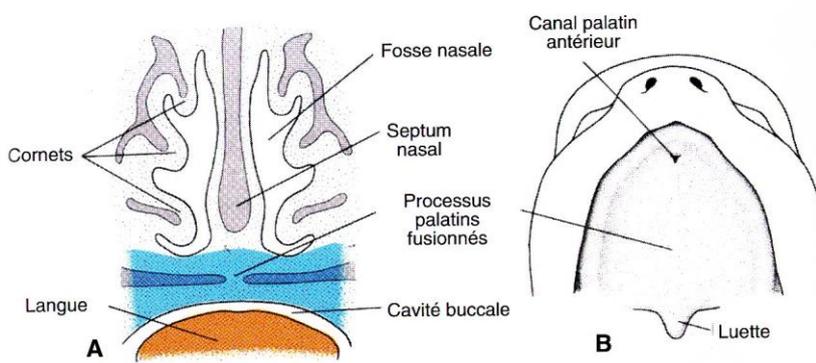


Figure 5: Fusion des processus palatins. (A) Coupe frontale de la région céphalique. (B) Vue centrale des processus palatins<sup>1</sup>



<sup>1</sup> Sources : Sulik KK. Department of Cell Biology and Anatomy. University of North Carolina. ([http://www.med.unc.edu/embryo\\_images/unit-welcome/welcome\\_htms/contents.htm](http://www.med.unc.edu/embryo_images/unit-welcome/welcome_htms/contents.htm))

Langman J, Sadler TW. Embryologie médicale. 6<sup>ème</sup> édition française. Editions Pradel. Paris, 1996.

## B. 1. Mesure par la prévalence

L'étude de la population à risque de fentes orales devrait inclure toutes les conceptions vivantes à la fin du processus normal de fusion. Or il n'est pas possible d'évaluer l'ensemble des conceptions qui se produisent et survivent jusqu'à cette date précoce du développement embryonnaire, ni d'évaluer si une fusion pathologique s'est produite ou non à cette date. L'incidence des fentes orales ne peut donc pas être obtenue. En revanche, il est classique de mesurer la proportion de fentes orales parmi les naissances (d'enfants vivants ou mort-nés) et les interruptions de grossesse se produisant après le premier trimestre, soit la prévalence des fentes orales (Leck 1972).

Dans un objectif de recherche étiologique, la mesure de prévalence est cependant limitée et son usage peut entraîner quelques résultats paradoxaux. Lorsqu'une exposition modérée conduit à un risque accru de fentes orales et qu'une exposition intense du même facteur engendre des fausses couches spontanées avant le premier trimestre, cette situation pourrait amener à conclure qu'une réduction du niveau d'exposition est associée à une augmentation de la prévalence des fentes orales.

## B. 2. Hétérogénéité des fentes orales

Les fentes orales sont un groupe de malformations congénitales très hétérogènes selon des éléments cliniques mais aussi étiologiques. Dans les études épidémiologiques à visée étiologique, il est important de distinguer différents types de fentes orales afin de déterminer les facteurs de risque spécifiques.

Les fentes orales peuvent être dissociées selon leur forme unilatérale (la plus commune, environ 80% de l'ensemble des fentes orales), ou bilatérale (forme la plus sévère).

La ressource OMIM (*the Online Mendelian Inheritance in Man*) estime la présence de **syndrome**<sup>2</sup> génétique connu pour environ 15% des fentes orales. Cette proportion est variable selon la définition et la classification d'une fente orale non syndromique : d'autres ressources rapportent entre 2 et 25% de cas syndromiques pour les FL/P, et entre 10 et 50% pour les FP (Saal 2002, Mossey et Little 2002). Parmi les 85% restantes (dites non syndromiques), OMIM estime que la moitié est associée à une autre anomalie ; l'autre moitié est dite « forme isolée ». Au total, selon la source OMIM, nous pouvons lister 174 **désordres mendéliens** associés à une FL/P et 312 à une FP (Saal 2002).

De façon générale, une FP est plus souvent syndromique ou associée à une autre malformation qu'une FL/P. Le syndrome de Stickler est identifié comme la forme syndromique la plus fréquente des FP (Saal 2002).

---

<sup>2</sup> Les mots du texte en **gras** sont définis dans le glossaire situé en annexe.

### B. 3. Variations temporelles de la prévalence

L'existence d'une tendance temporelle à l'augmentation de la prévalence des fentes orales est très controversée. Les tendances apparentes peuvent être dues à des meilleurs diagnostics, des processus d'enregistrement plus performants, une mortalité néonatale décroissante, un accroissement du nombre de tératogènes environnementaux... Les meilleures données sur lesquelles nous pouvons baser notre jugement d'une tendance croissante ou décroissante de la prévalence des fentes orales dans le temps, sont les registres qui requièrent des critères et des infrastructures relativement constants dans le temps sur une population géographique définie.

Le registre du Danemark, qui intègre un des systèmes d'enregistrement des enfants atteints les plus exhaustifs (depuis 1936) et des procédures rigoureuses d'évaluation, rapporte tout d'abord une augmentation légère de la prévalence des fentes orales jusqu'en 1961, puis des prévalences constantes sur la période de 1962 à 1987, avec environ 1.4 FL/P et 0.8 FP pour 1000 naissances (Christensen 1999a). Une mise à jour récente du registre a permis d'inclure toutes les naissances vivantes du Danemark de 1988 à 2001. Les nouvelles données du registre ne rapportent aucune tendance à la diminution ou à l'augmentation sur cette période de la prévalence des fentes orales non syndromiques (malgré une baisse de la consommation maternelle de tabac pendant la grossesse sur cette période et des recommandations de suppléments vitaminiques en acide folique depuis 1997 au Danemark, Bille et al. 2005b).

A partir de registres européens de malformations congénitales, le réseau EUROCAT (*European Registry Of Congenital Anomalies and Twins*) n'observe pas de tendance globale en Europe en faveur d'un accroissement ou d'une diminution des prévalences de FL/P ou de FP sur la période de 1980 à 2003 (Données de prévalence disponibles sur <http://www.eurocat.ulster.ac.uk/>).

### B. 4. Variations géographiques de la prévalence

En France, trois systèmes d'enregistrement de malformations congénitales sont disponibles :

- 1- le registre Centre-Est qui a débuté en 1973 et s'étend aujourd'hui aux régions Rhône-Alpes et Auvergne (environ 90 000 naissances annuelles).
- 2- un programme à Paris initié en 1975 couvrant les maternités parisiennes (38 500 naissances annuelles)
- 3- un registre sur la région autour de Strasbourg (limitée au département du Bas-Rhin), débutant en 1979 (13 500 naissances annuelles).

Ces trois registres sont membres du système d'enregistrement d'EUROCAT. Les prévalences de FL/P et de FP mesurées dans ces registres sont présentées dans le Tableau 1. Aucune tendance temporelle n'est observée sur la période de 1987 à 2001. En revanche, des disparités régionales en France sont constatées, particulièrement importantes pour la prévalence des FL/P (de 6.3 à 10.7 FL/P pour 10 000 naissances entre 1997 et 2001).

**Tableau 1: Prévalence des fentes orales en France**

Prévalence pour 10 000 naissances (vivantes ou mort-nées)	FL/P	FP
Registre Centre-Est		
1987-1991	6.5	4.6
1992-1996	7.7	6.3
1997-2001	6.3	5.5
Paris		
1987-1991	6.6	4.2
1992-1996	6.7	4.4
1997-2001	6.6	5.1
Strasbourg		
1987-1991	9.5	8.9
1992-1996	12.8	9.2
1997-2001	10.7	5.1

Sources :

Rapport annuel de l'ICBDMS (*International Clearinghouse for Birth Defects Monitoring Systems*) 2004, <http://www.icbd.org/document/AR2004/index.htm>

En Europe, un gradient nord-sud est observé pour la prévalence des FL/P, avec les plus fortes prévalences en Norvège, Danemark, Suède et Irlande (>10 pour 10 000 naissances), des prévalences moyennes en Europe Centrale et les plus faibles prévalences en Europe du Sud (<7 pour 10 000). Une variation de la prévalence des FP s'observe entre les pays européens mais aussi à l'intérieur de certains pays. Parmi les fentes orales non syndromiques, un ratio FL/P : FP de 2 : 1 est généralement admis dans la population européenne.

Les quelques études réalisées sur le continent africain estiment des prévalences faibles de FL/P et de FP. La population afro-américaine se caractérise par des taux de prévalence de FL/P et de FP plus faibles que la population blanche aux Etats-Unis. Les études en Asie rapportent des prévalences élevées de FL/P, similaires aux taux de FL/P mesurés dans la population du continent Nord-Américain d'origine chinoise ou japonaise.

Aujourd'hui, l'ensemble des recherches, notamment avec l'appui des études de populations migrantes, s'accordent à dire que les variations globales des taux de prévalences de FL/P et de FP observées dans différentes régions du monde sont liées aux caractéristiques des populations, reflétant l'influence de caractères génétiques mais aussi de modes de vie et expositions environnementales (au sens large), voire d'interaction entre ces facteurs.

### **C. Conséquence et problème de Santé Publique**

L'impact d'une fente orale chez l'enfant et sa famille varie considérablement et peut engendrer divers problèmes médicaux, psychologiques, chirurgicaux, dentaires... De nombreux spécialistes interviennent dans les processus de prévention, de réparation, de suivi et de soin d'une fente orale chez un individu (pédiatre, chirurgien, orthophoniste, orthodontiste, psychologues...). Le parcours de réparation est long, de la naissance à l'âge

adulte. Le succès de réparation dépend des coopérations développées entre les divers acteurs intervenant dans les étapes de ce parcours.

Cette approche multidisciplinaire semble essentielle pour permettre au sujet d'atteindre une expression orale et une apparence esthétique idéales. Elle dépend cependant des systèmes de santé de chaque pays et les coûts (économiques) associés varient considérablement selon les pays et la définition de la réussite du traitement. Le CDC (*Centers for Disease Control and Prevention*) estime aux Etats-Unis en 1992 un coût annuel global de 697 millions de dollars qui refléterait l'ensemble des moyens nécessaires pour que chaque enfant atteint d'une FL/P ou d'une FP ait la possibilité de poursuivre une vie « productive ».

# 2

## L'étiologie complexe des fentes orales

Les fentes orales (syndromiques ou non) constituent un groupe de malformations congénitales hétérogènes et connues pour être déterminées à la fois par des facteurs génétiques et environnementaux. Leur étiologie multifactorielle est étudiée depuis de nombreuses décennies et la littérature scientifique sur le sujet est dense.

### A. Epidémiologie environnementale des fentes orales

#### A. 1. Sexe et facteurs socio-démographiques

Dans la population européenne, il est bien établi que les enfants atteints de FL/P sont plus souvent des garçons, et les enfants atteints de FP sont plus souvent des filles (EUROCAT 2000).

Il n'y a pas d'évidence concordante selon laquelle l'âge de la mère et/ou du père jouent un rôle dans la survenue d'une fente orale : bien qu'une étude récente, basée sur les données de 1.5 million de naissances vivantes au Danemark entre 1973 et 1996, suggère l'influence de l'âge des parents sur le risque de fentes orales, et en particulier de l'âge paternel sur le risque de FP (Bille et al.2005a), cette relation n'a pas toujours été rapportée (Vieira et al. 2002).

Un faible niveau socio-économique a été souvent associé au risque de fente orale, l'étude la plus récente se base sur les données de plus de 600 000 naissances vivantes en Ecosse entre 1989 et 1998 (Clark et al. 2003).

Les enfants atteints de fentes orales semblent avoir une durée de gestation plus courte et un poids de naissance plus faible que les enfants normaux (Cohen 2002, Wyszynski et al. 2003).

#### A. 2. Epilepsie et antiépileptiques

Il est bien établi que les mères épileptiques ont un risque accru d'avoir un enfant atteint de malformations congénitales majeures, et plus particulièrement de fentes orales (Hayes 2002). Cependant, il est difficile de discerner si cette association s'explique par l'épilepsie maternelle *per se*, les traitements associés anti-épileptiques ou anti-convulsifs, ou un lien génétique entre l'épilepsie et la survenue d'une fente orale.

L'hypothèse d'une prédisposition génétique semble écartée, notamment par une revue de neuf études s'intéressant à l'existence d'une possible agrégation familiale de l'épilepsie et d'une fente orale (Hecht et Annegers 1990).

Certains auteurs rapportent que l'utilisation maternelle d'antiépileptiques pendant la grossesse affecte le risque de survenue de fentes orales, et suggèrent un mécanisme de dégradation des folates par certains antiépileptiques (Dravet et al. 1992, Abrishamchian et al. 1994, Wyszynski et Beaty 1996a, Arpino et al. 2000, Matalon et al. 2002). Cette association serait accentuée en présence de polythérapie de l'épilepsie (Abrishamchian et al. 1994). Les principaux traitements antiépileptiques incriminés sont le phénobarbital, le méthylphénobarbital, la phénytoïne, le primidone, et la carbamazépine. Hernández-Díaz et al. (2000) observent des risques de fentes orales associés aux prises d'anti-épileptiques, augmentés, significatifs et similaires, quelle que soit la prise d'acide folique par la mère pendant la grossesse, suggérant un effet tératogène des antiépileptiques, en particulier, la phénytoïne et le phénobarbital, indépendants des folates.

D'autres traitements, tels que les benzodiazépines (BDZ), peuvent être prescrits aux mères épileptiques pour décroître l'anxiété (généralement sous forme de tranquillisants). Les études s'intéressant à l'association entre l'utilisation de BDZ et le risque de fentes orales présentent des résultats variables, contradictoires et peu convaincants.

### A. 3. Vitamines

Parmi les vitamines, l'intérêt a beaucoup porté sur l'acide folique en raison de son rôle prédominant dans la prolifération cellulaire et donc potentiellement dans le développement embryonnaire. Il existe à ce jour des preuves suffisantes pour confirmer le rôle spécifique des suppléments en acide folique dans la prévention du risque de récurrence et d'occurrence des anomalies du tube neural. Un rôle semblable dans la prévention des fentes orales est toujours discuté.

Pourtant, plusieurs études cas-témoins ont évalué l'association entre la prise de multivitamines contenant de l'acide folique pendant la grossesse et le risque de fentes orales. Les résultats sont variables d'une étude à l'autre. Ils sont présentés dans le Tableau 2 par population et par ordre chronologique. La variabilité des résultats reflète l'hétérogénéité des études : par exemple, l'évaluation de l'utilisation maternelle des multivitamines varie (interview, questionnaire ou enregistrement médical), la période pertinente de prise de vitamines (1<sup>er</sup> trimestre de grossesse avec ou non une période préconceptionnelle), le mode de prise (régulière par jour ou irrégulière), la dose d'acide folique (inconnue jusqu'à 3mg/tablette et 1-3 tablettes/jour (Czeizel et al. 1999)), le groupe de référence de non utilisatrice de vitamines (exclusion ou non des femmes qui commencent la prise de vitamine après le 2<sup>ème</sup> mois de grossesse, exclusion ou non des prises irrégulières...), la diversité ethnique des populations.

Malgré les coûts logistiques et financiers des essais d'intervention randomisés, ces études ont la possibilité de fournir des preuves convaincantes sur le rôle des vitamines, et en particulier de l'acide folique dans la prévention de fentes orales. Czeizel et al. publient en 1999 la mise à jour des résultats de l'essai randomisé de prévention d'anomalies congénitales mis en place sur la population hongroise. A partir de 26 centres du Programme hongrois de planning familial, des multivitamines (MV, contenant des vitamines A, B, C, D, E, acide folique (15mg), minéraux) ou des suppléments-témoins (contenant minéraux, lactose et vitamine C), sont prescrits aux mères consentantes à usage régulier un mois avant la conception jusqu'à 3 mois après. Les résultats rapportent un total de 4 fentes orales et aucune anomalie du tube

neural parmi les enfants des 2471 femmes supplémentées en MV, et 5 fentes orales et 6 anomalies du tube neural parmi les 2391 enfants du groupe témoin. Ils indiquent finalement que l'effet de l'acide folique dans la réduction du risque de fentes orales, s'il existe, serait plus faible que pour les anomalies du tube neural. Les auteurs notent cependant que ce résultat manque de puissance statistique. Pour la même population, à partir d'une comparaison appariée de deux cohortes de femmes supplémentées et non supplémentées, ces auteurs suggèrent aucun effet protecteur de la prise d'acide folique sur le risque de fentes orales (Czeizel et al. 2004).

L'hypothèse dominante actuellement permettant d'expliquer le rôle protecteur probable des multivitamines contenant de l'acide folique dans la survenue d'une fente orale est qu'un faible statut en folates (traduit par un faible apport en acide folique ou par une altération dans les mécanismes de transport ou de métabolisme des folates) engendrerait un statut d'hyperhomocystéinémie maternel, suspecté tératogène chez certains mammifères. Or, le métabolisme de l'homocystéine est complexe et nécessite l'intervention de nombreux autres nutriments comme les vitamines B6 et B12.

Une autre approche plus récente pour l'étude du rôle des vitamines dans l'étiologie des fentes orales est l'utilisation de biomarqueur du statut vitaminique, permettant ainsi de cibler spécifiquement sur l'étude du rôle d'un nutriment particulier. Wong et al. (1999, Pays-Bas) rapportent tout d'abord un statut en folates (mesuré par deux indicateurs : *serum folate*, *RBC folate - Red Blood Cell*) plus élevé chez les mères d'enfants atteints de fentes orales que chez les mères de témoins, puis une concentration dans le sang en vitamine B6 (par le niveau de pyridoxal 5'-phosphate) significativement plus faible chez les mères d'enfants atteints de fentes orales que chez les mères de témoins. A l'aide des mêmes marqueurs, Van Rooij et al. (2003a, Pays-Bas) observent un statut en folate de la mère équivalent pour les deux groupes d'enfants-cas et d'enfants-témoins, et des concentrations de vitamines B6 et B12 dans le sang plus faibles chez les mères d'enfant atteint de fente orale que chez les mères-témoins. Munger et al. (2004) observent, sur deux populations des Philippines, des résultats contradictoires quant à l'association entre le statut en folates des mères et le risque de FL/P, puis des résultats concordants suggérant un risque augmenté de FL/P associé à une déficience en vitamine B6 des mères. Ce risque est plus fort parmi les mères ayant un faible niveau en folates.

Enfin, des études d'interaction nutriment-**génotype** peuvent aider à clarifier le rôle des folates et autres nutriments dans la prévention des fentes orales. Ces études sont présentées dans la section suivante C.

La relation entre vitamine A et fentes orales est encore très peu étudiée dans la littérature épidémiologique. Quelques études épidémiologiques ont examiné la vitamine A, via la supplémentation, la consommation de foie ou de légumes riches en bêta-carotène, en relation avec le risque de fentes orales (Werler et al. 1990a, Natsume et al. 1995, Shaw et al. 1996a, Mitchell et al. 2003). Les résultats sont contradictoires et ont montré à la fois un risque augmenté et un effet protecteur de l'apport en vitamine A.

**Tableau 2 : Etudes cas-témoins évaluant l'association entre la prise de multivitamines par la mère pendant la grossesse et le risque de fentes orales**

	Population	Mode de recueil	Période et mode d'utilisation des multivitamines (MV)	OR [IC95%] (utilisation versus non utilisation des MV)
<b>Europe</b>				
Saxen et al. 1975*	Finlande 1967-1971 232 FL/P, 232 témoinsT1 232 FP, 232 témoinsT2	Questionnaire auto-administré	1 <sup>er</sup> trimestre de grossesse	% d'utilisation MV : FL/P : 63.8% ; T1: 54.4% FP : 46.6% ; T2 : 47.8%
Hill et al. 1988*	Royaume-Uni 1983-1984 676 paires Cas-Témoins	Enregistrement médical	3 mois précédant la conception	% d'utilisation MV C: 1.9%; T: 1.3%
Czeizel et al. 1999	Hongrie 1980-1996 1246 paires FL/P-témoins 537 paires FP-témoins	Interview Dose d'acide folique par enregistrement médical	FL/P : mois 1et 2 de grossesse FP : mois 1 à 3 de grossesse	FL/P : 0.72 [0.55-0.92] FP : 0.86 [0.66-1.13]
Mitchell et al. 2003	Danemark 1991-1994 302 cas, 567 témoins	Questionnaire (acide folique :100µg vitamine A :2640 UI)	Utilisation régulière au 1 <sup>er</sup> trimestre de grossesse	OR inversés: FL/P : 1.68 [1.12-2.52] FP: 0.92 [0.51-1.92]
van Rooij et al. 2004	Pays-Bas 1998-2000 174 FL/P et 203 témoins	Questionnaire (questions ouvertes) + Questionnaire ALIMentaire	Utilisation régulière entre 4 semaines avant et 8 semaines après conception	MV : 0.53 [0.33-0.85] MV (+ ALIM>242µg/jour): 0.26 [0.09-0.72]
<b>Etats-Unis</b>				
Khoury et al. 1989*	Atlanta 1968-1980 345 cas, 2809 témoins	Interview au téléphone	3 mois avant jusqu' à 3 mois après conception	FL/P : 0.74 [0.56-0.97] FP : 0.93 [0.62-1.40]
Itikala et al. 2001*	Atlanta 1968-1980 309 cas, 3029 témoins	Interview au téléphone	Utilisation régulière à partir du 1 <sup>er</sup> mois de grossesse au plus tard	FL/P : 0.52 [0.34-0.80] FP : 0.81 [0.44-1.52]
Shaw et al. 1995	Californie 1987-1989 731 cas, 734 témoins	Interview au téléphone + Consommation de CEREALes fortifiées en acide folique au petit déjeuner (oui/non)	MV : 1 mois avant jusqu'à 2 mois après conception  CEREAL : tous les matins	MV : FL/P : 0.50 [0.36-0.68] FP : 0.73 [0.46-1.20] CEREAL (+ pas de MV): FL/P: 0.41 [0.17-0.98] FP : 0.26 [0.03-1.5]
Hayes et al. 1996	Boston, Philadelphie, Toronto 1988-1991 303 cas, 1167 témoins	Interview avec infirmière	Utilisation régulière 1 mois avant à 16 semaines après conception	FL/P: 1.3 [0.2-2.1] FP : 0.9 [0.5-1.6]
Werler et al. 1999*	Boston, Philadelphie, Toronto 1993-1996 160 cas, 521 témoins	Interview avec infirmière	Utilisation régulière à partir de 1 mois avant conception jusqu'à... ...3 mois après (FL/P) ...4 mois après (FP)	FL/P : 0.7 [0.4-1.1] FP : 0.4 [0.2-0.9]
Beaty et al. 2001	Maryland 1998-2000 135 cas, 152 témoins	Interviews à l'hôpital et au téléphone	Utilisation régulière 1 mois avant conception et 1 <sup>er</sup> trimestre de grossesse	FL/P : 0.6 [0.3-1.1] FP : 0.7 [0.3-1.6]
<b>Amérique du Sud</b>				
Loffredo et al. 2001	Sao Paulo (Brésil) 1991-1992 450 paires cas-témoins	Interview à l'hôpital ou maternité	Les 4 premiers mois de grossesse	FL/P : 0.58 [0.43-0.79] FP : 0.60 [0.37-0.98]

OR : odds-ratio

IC95% : intervalle de confiance à 95%

\*Sources : Munger 2002 et Botto et al. 2004

#### A. 4. Consommation maternelle de tabac

L'association entre la consommation maternelle de tabac pendant la grossesse et le risque de fentes orales a été examinée par de nombreuses études épidémiologiques cas-témoins ou de cohorte. Plusieurs revues de cette littérature rapportent et discutent des résultats conflictuels entre ces études (Werler 1997, Little 1999, Hayes 2002, Lorente 2001, Little et al. 2004a).

Certaines études rapportent un accroissement modéré du risque de fentes orales non syndromiques tandis que d'autres ne confirment pas cette association (OR : odds-ratio ; IC95% : intervalle de confiance à 95%) :

Malloy et al. 1989 <sup>3</sup> :	OR=0.8 ; IC95%=0.7,1.1,
Kelsey et al. 1978 <sup>4</sup> :	OR=1.10 ; IC95%=1.0,1.3
van Rooij et al. 2001:	OR=1.1 ; IC95%=0.6,2.2
Wyszynski et Wu 2002 :	OR=1.33 ; IC95%=1.2,1.5,
Chung et al. 2000:	OR=1.34 ; IC95%=1.2,1.6
Evans et al. 1979 <sup>2</sup> :	OR=1.35 ; IC95%=0.9,2.0
van Rooij et al. 2002:	OR=2.0 ; IC95%=0.8,5.0.

La majorité des études considèrent les FL/P et les FP séparément :

- Un excès modéré de risque FL/P (Ericson et al. 1979<sup>5</sup>, Khoury et al. 1987<sup>4</sup>, 1989<sup>4</sup>, van den Eeden et al. 1990<sup>4</sup>, Shaw et al. 1996b, Kallen et al. 1997<sup>4</sup>, Beaty et al. 1997, Christensen et al. 1999b, Romitti et al. 1999, Lorente et al. 2000a, Mitchell et al. 2001 [similaire à Christensen et al. 1999b], Meyer et al. 2004 [extension de Kallen et al. 1997], Little et al. 2004a)
- ainsi qu'un accroissement modéré du risque FP sont suggérés (Ericson et al. 1979, Khoury et al. 1987, 1989, van den Eeden et al. 1990, Shaw et al. 1996b, Kallen et al. 1997, Beaty et al. 1997, Romitti et al. 1999, Meyer et al. 2004, Little et al. 2004a),
- tandis que d'autres ne confirment pas ces associations (Shiono et al. 1986<sup>4</sup>, Werler et al. 1990b<sup>4</sup>, Hwang et al. 1995 [extension des études de Khoury et al. 1987,1989], Lieff et al. 1999 [extension de Werler et al. 1990b], Beaty et al. 2001 [extension de Beaty et al. 1997]).

Certaines de ces études examinent les relations dose-effet (Khoury et al. 1987,1989, Shaw et al. 1996b, Christensen et al. 1999b, Romitti et al. 1999, Lieff et al. 1999, Chung et al. 2000, Lorente et al. 2000a, Mitchell et al. 2001, van Rooij et al. 2002, Wyszynski et Wu 2002, Meyer et al. 2004, Little et al. 2004a) et un accroissement linéaire du risque avec la dose est suggérée par quelques-unes d'entre elles (Khoury et al. 1987, Shaw et al. 1996b, Romitti et al. 1999 (FP), Chung et al. 2000, Lorente et al. 2000a (FL/P), Wyszynski et Wu 2002, Little et al. 2004a).

Chung et al. (2000) est actuellement la plus grande étude publiée sur cette association et se base sur près de 4 millions de naissances enregistrées aux Etats-Unis depuis 1996 (couverture à 99% à partir des certificats de naissance). En revanche, elle ne peut pas discriminer les FL/P des FP.

---

<sup>3</sup> Source : Lorente 2001

<sup>4</sup> Source : Chung et al. 2000

<sup>5</sup> Source : Hayes 2002

Deux méta-analyses des études publiées jusqu'en 1996 (Wyszynski et al. 1997a) ou jusqu'en 2002 (Little et al. 2004b) ont estimé les risques relatifs globaux de FL/P et de FP en relation avec la consommation maternelle de tabac pendant la grossesse. Les résultats sont concordants et concluent à une association positive, modérée et statistiquement significative pour les deux groupes des FL/P et des FP (le biais potentiel de publication est testé dans l'étude de Little et al. (2004b), qui ne détecte pas de signal). Ils sont présentés dans le Tableau 3. Little et al. (2004b) suggèrent un accroissement modeste du risque de FL/P avec les niveaux de consommation maternelle de tabac.

**Tableau 3: Résultats des deux méta-analyses s'intéressant à l'association entre la consommation maternelle de tabac pendant la grossesse et le risque de fentes orales non syndromiques**

	FL/P	FP
	RR [IC95%]	RR [IC95%]
Wyszynski et al. 1997a (11 études)	1.29 [1.18-1.42]	1.32 [1.10-1.62]
Little et al. 2004b (24 études)	1.34 [1.25-1.44]	1.22 [1.10-1.35]

RR : risque relatif  
IC95% : intervalle de confiance à 95%

Depuis une dizaine d'années, quelques études épidémiologiques ont choisi d'examiner cette association en interaction avec différents **gènes** du développement embryonnaire, ou impliqués dans le métabolisme de substances toxiques issues de la consommation de tabac. Ces travaux sont présentés dans la section suivante C.

### A. 5. Consommation maternelle d'alcool

Une consommation excessive d'alcool pendant la grossesse provoque des altérations dans le développement embryonnaire incluant des dysmorphies craniofaciales, l'ensemble étant décrit sous le nom de « syndrome d'alcoolisation fœtal ». Pourtant, la littérature épidémiologique s'intéressant à la relation entre le risque de fentes orales et la consommation maternelle d'alcool est moins dense que la littérature ciblant sur la relation avec le tabac. L'existence d'un biais de publication peut être envisagé (Little et al. 2004b).

Certaines études rapportent des associations positives (Werler et al. 1991, Munger et al. 1996, Shaw et Lammer 1999a, Romitti et al. 1999, Lorente et al. 2000a) mais les accroissements de risque observés et les doses d'alcool ingérées par la mère sont très variables d'une étude à l'autre, rendant les résultats difficilement comparables. Ces études observent cependant de façon concordante un risque augmenté pour le groupe des FL/P avec la consommation maternelle d'alcool, exceptée l'étude de Lorente et al. (2000a) qui rapporte un risque augmenté et statistiquement significatif uniquement pour le groupe des FP.

D'autres études n'observent pas de relation entre le risque de fentes orales et la consommation maternelle d'alcool, quelle que soit la mesure utilisée : pas de consommation d'alcool pendant

la grossesse versus toute consommation, ou versus différentes doses ingérées (Mills et Graubard 1987, Beaty et al. 2001, Mitchell et al. 2001, Meyer et al. 2003).

Le tableau suivant présente les risques estimés de FL/P et de FP pour les plus fortes doses ingérées d'alcool observées dans chaque étude.

**Tableau 4: Risques de FL/P et de FP associés aux plus fortes doses d'alcool ingérées par la mère pendant la grossesse, des études épidémiologiques mesurant l'association entre alcool et risque de fentes orales**

	FL/P OR [CI95%]	FP OR [CI95%]
<b>Mills et Graubard 1987</b>		
Toute consommation	1.3 [0.8-2.2]	
<b>Werler et al. 1991</b> <b>Boston, Philadelphie, Toronto, régions d'Iowa, 1983-1987</b>		
≥ 5 verres / jour	3.0 [1.1-8.5]	0.9 [0.1-7.2]
<b>Munger et al. 1996</b> <b>Iowa, 1987-1991</b>		
≥10 verres / mois	4.0 [1.1-15.1]	1.8 [0.3-12.1]
<b>Romitti et al. 1999</b> <b>Iowa, 1987-1994</b>		
≥4 verres / mois	2.8 [1.2-6.6]	1.7 [0.5-6.4]
<b>Shaw et Lammer 1999a</b> <b>Californie, 1987-1989</b>		
≥ 5 verres / occasion et ≥ 1 occasion / semaine	3.4 [1.1-9.7] (FL/P isolées)	1.0 [0.2-8.5] (FP isolées)
<b>Lorente et al. 2000a</b> <b>France, Italie, Royaume-Uni, Pays-Bas, 1989-1992</b>		
≥ 70g/ semaine (≈ 7 verres de vin/semaine)	1.12 [0.4-2.8] (FL/P isolées)	1.88 [0.6-5.9] (FP isolées)
<b>Beaty et al. 2001</b> <b>Maryland, 1992-1998</b>		
Toute consommation	0.4 [0.2-0.8]	0.5 [0.2-1.2]
<b>Mitchell et al. 2001</b> <b>Danemark, 1991-1994</b>		
> 6 verres /semaine	0.55 [0.2-1.5]	0.68 [0.2-3.1]
<b>Meyer et al. 2003</b> <b>Boston, Philadelphie, Toronto, régions d'Iowa, 1983-1997</b>		
≥3 verres / jour	1.0 [0.6-1.7]	0.9 [0.4-2.1]

OR : odds-ratio

IC95% : intervalle de confiance à 95%

## A. 6. Expositions aux solvants organiques et aux pesticides

### A.6.i. Expositions professionnelles de la mère aux solvants organiques

L'association entre exposition professionnelle de la mère aux solvants et risque de fente orale a été évoquée dans une première étude finlandaise de Holmberg et al. (1982) puis confirmée par des résultats statistiquement significatifs dans d'autres études en France (Cordier et al. 1992; Laumon et al. 1996) et en Europe, en particulier avec les éthers de glycol (Cordier et al. 1997, 2001).

Deux études récentes évaluant cette association ne rapportent pas de résultat statistiquement significatif (étude européenne multicentrique de Lorente et al. 2000b, étude sur une population californienne de Shaw et al. 2003a). L'étude européenne observe cependant des excès de risque FL/P avec l'exposition aux aldéhydes aliphatiques (OR=2.1 ; IC95%=0.8,5.9) et aux éthers de glycol (OR=1.7 ; IC95%=0.9,3.3), un risque élevé de FP avec l'exposition au trichloroéthylène (OR=6.7 ; IC95%=0.9,49.7). L'étude californienne observe des excès de risque de FL/P avec l'exposition aux colorants aromatiques (OR=2.7 ; IC95% :0.9,7.7) et des risques élevés de FP avec l'exposition aux hydrocarbures aliphatiques (OR=2.2 ; IC95%=0.9,5.7), aux cétones (OR=2.3 ; IC95%=0.7,9.1), aux amines aromatiques (OR=1.5 ; IC95%=0.5,5.6) et aux sulfures et disulfures (OR=1.8 ; IC95%=0.6,6.8).

D'autres études ont examiné le risque de fentes orales en fonction du secteur professionnel et des activités professionnelles de la mère. S'appuyant respectivement sur des populations italiennes (1980-1991) et espagnoles (1993-1994), Bianchi et al. (1997) et Garcia et Fletcher (1998) suggèrent un risque augmenté de fentes orales pour les mères travaillant dans l'industrie du cuir pendant la grossesse, et plus particulièrement pour les FP (Bianchi et al. 1997). Ces auteurs évoquent le rôle possible d'une exposition à des solvants tels que les hydrocarbures aliphatiques et/ou le toluène. Les activités de coiffure et de tâches ménagères ont été associées à un risque élevé de fentes palatines dans l'étude de Lorente et al. (2000b), incriminant possiblement une exposition aux éthers de glycol. Le travail en laboratoire de la mère pendant la grossesse a été examiné en relation avec le risque de fentes orales par quelques études, suspectant des expositions intenses au benzène, toluène, xylène et hydrocarbures aromatiques, mais les résultats sont inconstants. Une étude suédoise observe un risque augmenté et significatif d'anomalies craniofaciales et conotruncales avec l'exposition au benzène dans des laboratoires de recherche biomédicales (Wennborg et al. 2005).

Très peu d'études ont porté sur les expositions professionnelles des pères aux solvants en relation avec le risque de fente orale (Olshan et al. 1991, Kristensen et al. 1993).

Finalement, l'ensemble de ces résultats d'études épidémiologiques (européennes pour la majorité) ne permet pas de dresser des conclusions définitives quant au risque de survenue de fente orale lié à l'exposition des parents aux solvants organiques. La multiplicité des expositions professionnelles aux solvants organiques, à d'autres substances chimiques et à des agents biologiques, ainsi que la possible influence des conditions physiques de travail et des expositions domestiques aux solvants organiques rendent difficile l'attribution des effets observés à un agent spécifique. De plus, la diversité des stratégies d'évaluation des expositions professionnelles produit des résultats difficilement comparables.

### **A.6.ii. Expositions environnementales aux solvants organiques**

Ciblant plus spécifiquement sur l'exposition aux solvants chlorés, différentes études ont recherché les effets des dérivés des produits chlorés dans l'eau de consommation sur la survenue d'une fente orale. Des effets toxiques des dérivés, comme les trihalométhanes, ont été trouvés (Bove et al. 1995, revue dans Bove et al. 2002) mais de façon inconstante (Kristensen et al. 1997, Dodds et King 2001, Shaw et al. 2003b). L'étude de Bove (1995) a montré un lien entre l'ingestion d'eau contaminée par les solvants organiques (tétrachloroéthylène, dichloroéthane et benzène) et la survenue de fentes orales.

### **A.6.iii. Expositions aux pesticides**

Le rôle du travail agricole, et en particulier de l'exposition aux pesticides des parents pendant la grossesse dans l'étiologie des fentes orales a été suggéré (Gordon et Shy 1981, Nurminen et al. 1995, Schnitzer et al. 1995 (père agriculteur), Garcia et al. 1999, Hanke et Jurewicz 2004) mais de façon inconstante (Thomas et al. 1992, Shaw et al. 1999b). L'étude de Shaw et al. (1999b) évalue, de plus, le risque lié à l'utilisation domestique de pesticides et insecticides et le risque d'un lieu de résidence proche de terrains agricoles. Les auteurs rapportent un seul risque augmenté et significatif de FP avec l'utilisation d'insecticides.

## **A. 7. Autres facteurs de risque potentiels**

Les corticostéroïdes, utilisés pendant la grossesse pour le traitement de l'asthme par exemple, ont été associés à un risque accru de FL/P par quelques études et de façon constante. Les résultats pour le groupe de FP ne sont, en revanche, pas concordants (Hayes 2002).

De nombreuses autres études se sont intéressées à d'autres facteurs de risque probables des fentes orales (Hayes 2002) tels que le diabète de la mère, l'hyper-/l'hypo-thyroïdie de la mère, la présence de fièvre de la mère pendant la grossesse, l'usage maternel de drogue, la consommation de café, les conditions physiques de travail de la mère, l'exposition aux radiations ionisantes, à des substances chimiques telles que les biocides et les phtalates, les composés de plomb, l'exposition aux nitrates, la proximité d'usine de traitement des déchets, de champs électromagnétiques, l'exposition à une pollution de l'air ambiant et le stress maternel. Ces études sont isolées et pour la plupart souffrent de faiblesses méthodologiques. Peu d'informations sont alors disponibles pour pouvoir dresser des conclusions sur le rôle de ces facteurs de risque dans la survenue ou la prévention d'une fente orale.

## B. Influences génétiques

La présence d'antécédents familiaux de fentes orales est un des facteurs de risque les mieux identifiés des fentes orales non syndromiques (Wyszynski et al. 1996b, Mitchell 2002). Cette histoire familiale positive n'est cependant pas suffisante (ni nécessaire) pour prédire la survenue de l'anomalie. Les modes d'hérédité génétique d'une FL/P et d'une FP sont complexes.

Plusieurs approches peuvent être utilisées pour identifier un gène de susceptibilité à la survenue d'une fente orale. Les analyses de ségrégation sont des outils statistiques utilisés sur des données familiales pour déterminer le mode d'hérédité d'une maladie (par exemple, un modèle dépendant d'un seul gène majeur versus un modèle à hérédité multifactorielle). Les analyses de liaison déterminent si les **allèles** de deux loci sont hérités ensemble ou non, et estiment le taux de **recombinaison** (ou mesure d'une distance génétique) entre deux loci. Enfin, les études d'association comparent différents facteurs de risque entre un groupe de malades et un groupe de référence. De façon générale, les études de liaison sont plus appropriées aux maladies ayant un faible nombre de gènes déterminants. Les études d'association semblent, en revanche, mieux adaptées aux maladies possiblement déterminées par un grand nombre de gènes à effet modeste sur le risque (Marazita et Neiswanger 2002). Deux types d'analyses sont utilisés dans ces études d'association, à partir du schéma d'étude cas-témoins traditionnel ou d'un schéma d'étude basé sur des données familiales.

### B. 1. Ce que les études de jumeaux nous apprennent sur la contribution génétique ?

Les études de jumeaux observent la présence d'une fente orale pour chaque individu et calculent des taux de concordance parmi les paires de jumeaux monozygotes (MZ : génétiquement identiques) et les paires de jumeaux dizygotes (DZ : qui partagent la moitié de leur matériel génétique). Une différence entre les taux de concordance observés chez les MZ et chez les DZ peut être alors attribuée à leur degré variable de similarité génétique, et permet donc de conclure à l'importance de la contribution d'une influence génétique dans la survenue d'une fente orale. De plus, un taux de concordance parmi les jumeaux MZ inférieur à 100% indique les influences d'une altération génétique se produisant après le clivage de l'embryon et/ou d'autres facteurs non hérités (ou facteurs environnementaux dans le sens large) sur le risque de maladie.

Wyszynski et al. (1996b) rapportent des taux de concordance pour les paires de jumeaux MZ et DZ, respectivement de 36% (MZ) et 4.7% (DZ) pour les FL/P, et de 22%(MZ) et 4.6% (DZ) pour les FP. Malgré la variabilité des estimations rapportées par les études de jumeaux atteints de fentes orales, tous les auteurs s'accordent à conclure à l'existence d'influences à la fois génétique et environnementale dans l'étiologie des fentes orales. En particulier pour le groupe des FL/P, la part de la contribution génétique serait la plus importante (Mitchell 2002).

## B. 2. Les études d'association génétique

La littérature des études d'association génétique dans l'étiologie des fentes orales est dense. La revue de littérature présentée dans cette section se limite aux études de populations caucasiennes (Europe, Amérique du Nord, Australie). Par défaut, l'association étudiée considère le gène de l'enfant.

### B.2.i. *Transforming Growth Factor alpha TGF $\alpha$*

La première étude rapportant une association génétique positive avec le risque de fentes orales est une étude cas-témoins d'association entre les FL/P et le **polymorphisme** *TaqI* du gène *TGF $\alpha$* , situé dans la région chromosomique 2p13 (OMIM\*190170, Ardingier et al. 1989).

Cette association a été très étudiée par la suite et reproduite dans plusieurs études. En revanche, les études les plus récentes ne confirment pas cette association, quel que soit le choix du schéma d'étude (Tableau 5).

Mitchell (1997) évalue par une méta-analyse l'association du *TGF $\alpha$*  avec le risque de FL/P à partir de 7 études cas-témoins sur des populations caucasiennes publiées avant 1997, et rapporte un résultat significatif. Ce résultat est cependant peu convaincant puisque l'auteur observe une forte hétérogénéité entre les groupes de cas, non liée aux critères ethniques des populations, tandis que les groupes de témoins paraissent plus homogènes : l'auteur note que les études incluses dans la méta-analyse présentent des proportions variables de cas ayant une histoire familiale de fentes orales ou ne distinguent pas les cas sévères (FL/P bilatérales) des autres, ce qui pourrait possiblement avoir une importance dans l'estimation de l'association (Mitchell 1997, Lidral et al. 1998).

Depuis 1997, aucune des études présentées dans le Tableau 5 s'intéressant à cette association ne rapporte de résultat significatif.

Les résultats de l'association du *TGF $\alpha$*  avec le risque de FP sont contradictoires.

Finalement, le rôle du *TGF $\alpha$*  dans l'étiologie des FL/P et des FP n'est pas établi. L'association du *TGF $\alpha$*  avec les risques de FL/P et de FP a été étudiée en interaction avec des expositions environnementales, ces études sont commentées dans la section suivante c.

Tableau 5: Etudes d'association du gène *TGF $\alpha$*  avec le risque de fentes orales (populations caucasiennes)

Référence	Population	Schéma d'étude	Marqueur du TGF $\alpha$	Résultat de l'association	
				FL/P	FP
Ardinger et al. 1989	Iowa	Cas-témoins	TaqI	+	nr
			BamHI	+	nr
			RsaI	-	nr
Holder et al. 1992	Royaume-Uni	Cas-témoins	TaqI	+	nr
			BamHI	-	nr
			RsaI	-	nr
Chenevix-Trench et al. 1992	Australie	Cas-témoins	TaqI	+	nr
			BamHI	+	nr
			Hinfl	-	nr
Sassani et al. 1993	Philadelphie	Cas-témoins	TaqI	+	nr
Stoll et al. 1993*	France	Cas-témoins	TaqI	-	-
			BamHI	-	nr
Feng et al. 1994†	Philadelphie	TDT	TaqI	+	nr
Hwang et al. 1995	Maryland	Cas-témoins	TaqI	-	+
Shaw et al. 1996b	Californie	Cas-témoins	TaqI	-	« + »
Maestri et al. 1997	Maryland	TDT	D2S443	-	-
Beaty et al. 1997	Maryland	Cas-témoins	TaqI	-	-
Wyszynski et al. 1997b†	Etats-Unis	TDT	D2S443	-	nr
Lidral et al. 1998	Iowa	Cas-témoins	TaqI	-	-
			D4S433	-	-
		TDT	TaqI	-	-
			D4S433	-	-
Scapoli et al. 1998	Italie	TDT	TaqI	-	nr
Christensen et al. 1999b	Danemark	Cas-témoins	TaqI	-	-
Romitti et al. 1999	Iowa	Cas-témoins	TaqI	-	-
Beaty et al. 2001	Maryland	Cas-témoins	TaqI	-	-
			D2S443	-	-
Beaty et al. 2002	Maryland	TDT	D2S443	-	-
Jugessur et al. 2003a	Norvège	Triades	TaqI	-	+

TDT : Test de Déséquilibre de Transmission ; Triades : schéma d'étude cas-parents avec analyse log-linéaire

+ : association statistiquement significative ; « + » : association non significative, mais suggérée car forte ; - : association non significative ; nr : non renseigné

\*source Mitchell et al. 1997

† à partir de l'abstract

### B.2.ii. Transforming Growth Factor beta3 *TGFβ3*

L'expression du gène du *TGFβ3* (région chromosomique 14q24, OMIM\*190230) a été étudiée de façon abondante dans les expérimentations animales qui rapportent un rôle crucial du gène dans le développement du palais secondaire chez la souris (Chong et al. 2002).

Les résultats épidémiologiques de l'association du *TGFβ3* avec le risque de FP sont contradictoires (chez l'homme, Tableau 6).

L'association du gène *TGFβ3* avec les FL/P, rapportée de façon inconstante par les études chez l'homme, est inattendue, compte tenu de l'absence d'observation de FL/P chez des souris *knock-out* (Juriloff 2002). Certains auteurs expliquent ce résultat par la présence probable d'une interaction gène-gène entre le *TGFβ3* et l'homéobox *MSX1* (Lidral et al. 1998, Beaty et al. 2002). D'autres suggèrent une expression du gène altérée par la présence de facteurs exogènes (consommation maternelle d'alcool ou de tabac) (Maestri et al. 1997).

Tableau 6: Etudes d'association du gène *TGFβ3* avec le risque de fentes orales (populations caucasiennes)

Référence	Population	Schéma d'étude	Marqueur du <i>TGFβ3</i>	Résultat de l'association	
				FL/P	FP
Maestri et al. 1997	Maryland	TDT	D14S61	+	+
Lidral et al. 1998	Iowa	Cas-témoins	CA	-	-
			X5.1	-	-
			5'UTR.1	-	-
		TDT	CA	-	-
			X5.1	« + »	-
			5'UTR.1	+	-
Romitti et al. 1999	Iowa	Cas-témoins	CA	+	-
			X5.1	-	-
			5'UTR.1	-	-
Beaty et al. 2001	Maryland	Cas-témoins	D14S61	-	-
Mitchell et al. 2001	Danemark	Cas-témoins	CA	-	+
			X5.1	-	-
Beaty et al. 2002	Maryland	TDT	D14D61	-	+
Jugessur et al. 2003a	Norvège	Triades	CA	« + »	-

TDT : Test de Déséquilibre de Transmission ; Triades : schéma d'étude cas-parents avec analyse log-linéaire

+ : association statistiquement significative ; « + » : association non significative, mais suggérée car forte ; - : association non significative

### B.2.iii. Homeobox *MSX1*

La majorité des études épidémiologiques du gène *MSX1* (région chromosomique 4p16, OMIM\*142983) rapporte une association avec le risque de FP (Tableau 7). Deux études suggèrent aussi un lien entre deux **marqueurs** du *MSX1* et le risque de FL/P. Fallin et al. (2003) s'appuient sur la même source de données que Beaty et al. (2002) et complètent les résultats des marqueurs isolés (présentés dans le Tableau 7) par une analyse d'**haplotypes** de trois marqueurs successifs de la région de *MSX1* parmi les sept marqueurs étudiés. Pour le groupe des FP, les auteurs montrent une association significative (ou un excès de transmission par le TDT adapté à l'étude des haplotypes) de l'ensemble des haplotypes contenant le

marqueur CA. Un seul haplotype (snp1, snp2, snp5) a été associé de façon statistiquement significative à la survenue d'une FL/P.

**Tableau 7: Etudes d'association du gène *MSX1* avec le risque de fentes orales (populations caucasiennes)**

Référence	Population	Schéma d'étude	Marqueur du <i>MSX1</i>	Résultat de l'association	
				FL/P	FP
Lidral et al. 1998	Iowa	Cas-témoins	CA	-	+
			X1.1	-	-
			X1.3	+	+
			X2.1	-	-
		TDT	CA	-	-
			X1.1	-	-
			X1.3	-	-
			X2.1	-	-
Romitti et al. 1999	Iowa	Cas-témoins	CA	-	-
			X1.1	-	-
			X1.3	-	-
			X2.1	-	-
			X2.4	-	-
Beaty et al. 2001	Maryland	Cas-témoins	CA	-	+
			Mitchell et al. 2001	Danemark	Cas-témoins
			N8		
Beaty et al. 2002	Maryland	TDT	CA	-	+
			Fallin et al. 2003	Maryland	TDT
			G219A (snp1)		
			G232A (snp2)	-	-
			G590T (snp5)	-	« + »
			T1929G (snp6)	-	-
			C2204A (snp7)	-	+
			G2284A (snp8)	-	« + »
Jugessur et al. 2003a	Norvège	Triades	CA	-	« + »
			X1.3	-	-

TDT : Test de Déséquilibre de Transmission ; Triades : schéma d'étude cas-parents avec analyse log-linéaire  
 + : association statistiquement significative ; « + » : association non significative, mais suggérée car forte ; - : association non significative

L'ensemble de ces résultats sont concordants avec les observations récentes rapportées par les expérimentations animales : des mutations du gène *MSX1* perturberaient le développement du palais secondaire et le développement des dents, provoquant chez les souris *knock-out* des fentes palatines et autres anomalies craniofaciales (Lidral et al. 1998, Chong et al. 2002).

#### **B.2.iv. Autres gènes impliqués dans le développement embryonnaire**

De nombreuses autres études d'association avec le risque de fentes orales se sont intéressées à des gènes suspectés d'être impliqués dans le développement embryonnaire :

- les proto-oncogènes *BCL3* (Maestri et al. 1997, Lidral et al. 1998) ou *BCL2* (Chenevix-Trench et al. 1992),
- les homéobox *HOX2F*, *HOX2G*, *EN2* (Chenevix-Trench et al. 1992),

- les *Epidermal Growth Factors EGF* (Ardinger et al. 1989)
- et leurs récepteurs *EGFR*, communs avec les *TGF* (Ardinger et al. 1989).

Une étude récente réalisée sur dix populations montre un rôle important du gène *IRF6* (*Interferon Regulatory Factor 6*) dans la survenue d'une FL/P à partir du site *V274I* polymorphique (pour les populations incluses dans l'étude), et d'autres sites polymorphiques du gène ou proche du gène (Zuccherro et al. 2004).

Récemment, certains auteurs ont examiné le rôle simultané de multiples loci dans la survenue d'une fente orale par des interactions gène-gène. Il s'agit des travaux de Beaty et al. (2002) et de Jugessur et al. (2003b, 2003c) qui suggèrent l'existence d'interactions entre *TGFβ3* et *MSX1* et entre *TGFβ3* et *BCL3* dans l'étiologie de fentes orales non syndromiques, et des interactions entre *TGFα* et *MSX1*, et *TGFα* et *MTHFR* (gène impliqué dans le métabolisme des folates, cf partie suivante) plus spécifiquement dans l'étiologie des fentes palatines.

#### **B.2.v. MéthylèneTétraHydroFolate Réductase MTHFR**

La prise de vitamines contenant de l'acide folique a été associée au risque de fentes orales par certaines études, mais de façon inconstante. Pour clarifier cette relation, certains auteurs ont choisi d'examiner le rôle des gènes impliqués dans le transport ou le métabolisme des folates dans l'étiologie des fentes orales.

Le gène *MTHFR* (région chromosomique 1p36.3, OMIM\*607093) joue un rôle majeur dans le métabolisme des folates. Son polymorphisme *C677T*, commun dans la population caucasienne, est associé à une activité enzymatique réduite et à une thermolabilité accrue. La forme mutée du polymorphisme *A1298C* est aussi caractérisée par une activité réduite, mais dans une moindre mesure. Quelques études associent l'allèle variant *677T* au risque de FP ou de FL/P, mais les résultats sont contradictoires et les conclusions peu convaincantes (Tableau 8). Certaines de ces études différencient les rôles du gène porté par le fœtus de celui de la mère, en supposant que ce dernier puisse modifier l'environnement intra-utérin de l'enfant.

Shaw et al. (2003c) ne montrent pas d'association entre le gène du transport des folates *RFC1* (*Reduced Folate Carrier 1*) et le risque de FL/P ou de FP.

**Tableau 8: Etudes d'association du gène *MTHFR* avec le risque de fentes orales (populations caucasiennes)**

Référence	Population	Schéma d'étude	Marqueur du MTHFR	Résultat de l'association	
				FL/P	FP
Shaw et al. 1998a, 1999c	Californie	Cas-témoins	C677T	-	-
Mills et al. 1999	Irlande	Cas-témoins	C677T	-	+
Martinelli et al. 2001	Italie	Cas-témoins	C677T enfant	-	nr
			C677T mère	+	nr
Beaty et al. 2002	Maryland	TDT	C677T	-	nr
		TDT	C677T	-	-
			A1298C	-	-
Prescott et al. 2002	Royaume-Uni	Test d'HWE	C677T mère	« + »	nr
Jugessur et al. 2003b	Norvège	Triades	C677T enfant	-	« + »
			A1298C enfant	-	-
			C677T mère	« + »	-
			A1298C mère	-	-
van Rooij et al. 2003b	Pays-Bas	Cas-témoins	C677T enfant	-	nr
			A1298C enfant	-	nr
			C677T mère	-	nr
			A1298C mère	-	nr
		TDT	C677T	-	nr
			A1298C	-	nr
Pezzetti et al. 2004	Italie	TDT	C677T	-	nr
			A1298C	-	nr
		Cas-témoins	C677T enfant	-	nr
			C677T mère	+	nr

TDT : Test de Déséquilibre de Transmission ; Triades : schéma d'étude cas-parents avec analyse log-linéaire ; HWE : Equilibre d'Hardy-Weinberg  
 + : association statistiquement significative ; « + » : association non significative, mais suggérée car forte ; - : association non significative  
 nr : non renseigné

### **B.2.vi. Autres gènes de métabolisme**

Le gène *RARA*, Récepteur Alpha de l'Acide Rétinoïque, a été associé positivement au risque de FL/P par deux études (Chenevix-Trench et al. 1992, Maestri et al. 1997). Cette association n'est pas retrouvée par Mitchell et al. (2003).

Les gènes N-acétyl-transférases 1 et 2 (*NAT1*, *NAT2*), impliqués dans la biotransformation de divers composés toxiques exogènes ou endogènes, ont été examinés par quelques études en association avec le risque de fentes orales (van Rooij et al. 2002, Lammer et al. 2004a, 2004b). Une association positive du *NAT1* sur les FL/P est rapportée (Lammer et al. 2004a, 2004b).

Des études plus isolées se sont intéressées à la relation entre la survenue d'une fente orale et des gènes impliqués dans le métabolisme de xénobiotiques toxiques tels que les Glutathion-S-Transférases (*GST*) et les cytochromes P450 (*CYP*) (van Rooij et al. 2001, Lammer et al. 2005).

### **B.2.vii. Marqueurs anonymes**

A partir de résultats suggérés par des analyses de liaison, quelques études d'association ont rapporté des liens entre la survenue de fentes orales et des marqueurs anonymes situés dans les régions chromosomiques 2p13, 2q37, 3p21.2, 4q25-4q31, 10p13, 11p12-14, 12q13, 16p13.3, 16p24 (Mitchell et al. 1995, Blanton et al. 2004a, 2004b).

## **C. Interactions gène-environnement**

La littérature sur les études d'interaction gène-environnement dans l'étiologie des fentes orales est très récente et ne concerne actuellement que quelques populations, qui ont été bien étudiées. Ces études sont présentées dans le Tableau 10 placé à la fin de cette section, et elles sont décrites par et pour chaque population.

### **C. 1. Maryland**

La première étude d'interaction gène-environnement dans la survenue de fente orale s'appuie sur des données d'un registre de malformations congénitales du Maryland entre 1984 et 1992 par Hwang et al. (1995).

La population des enfants atteints de fentes orales a été complétée par un recrutement de 1992 à 1996 pour les études de Beaty et al. (1997) et Maestri et al. (1997), puis jusqu'en 1998 pour Beaty et al. (2002) à partir de centres de traitement dans le Maryland et une clinique spécialisée de Washington dans les anomalies craniofaciales, recouvrant ainsi près de 72% des enfants atteints de fentes orales à Maryland (Beaty et al. 2001).

Les mères des enfants ont été interviewées en deux étapes (1<sup>ère</sup> visite, téléphone) et des échantillons d'ADN ont été collectés chez les enfants et leurs parents par prélèvements veineux, par tâches de sang ou par frottis buccal.

Les études sur la population du Maryland ont examiné la consommation maternelle de tabac en interaction avec divers gènes candidats de l'enfant : *TGF $\alpha$* , *TGF $\beta$ 3*, *BCL3*, *RARA*, *MSX1* et *F13A*.

Hwang et al. (1995) observent, par une approche cas-témoins, un risque augmenté et significatif de FP chez les enfants porteurs de l'allèle variant *C2* du *TGF $\alpha$*  et de mères fumeuses, suggérant ainsi la présence d'une interaction (pas de test d'interaction rapporté). Beaty et al. (1997) ne confirment pas le résultat. L'approche cas-parents de Maestri et al. (1997) permet de conclure à un excès de transmission de l'allèle '4' du marqueur *D2S443* (proche du gène candidat du *TGF $\alpha$* ) parmi les mères fumeuses. Beaty et al. (2002) manquent de données familiales pour confirmer ce résultat.

Un rôle de la consommation maternelle de tabac sur le risque de fentes orales est suggéré en interaction avec deux autres gènes candidats. Un excès significatif de transmission de l'allèle '6' du marqueur *DI4S61* (proche du **locus** du *TGF $\beta$ 3*) parmi les enfants atteints de fentes orales et de mères fumeuses est observé (Maestri et al. 1997). Les enfants-cas de mères fumeuses ont plus de chance d'être homozygotes par l'allèle '4' du marqueur *CA* de *MSX1* versus les enfants de mères non fumeuses (Beaty et al. 2002). Ce résultat est concordant avec

les interactions significatives, rapportées par Fallin et al. (2003), entre l'ensemble des haplotypes de 3 marqueurs de *MSXI*, dont le marqueur *CA*, et la consommation maternelle de tabac, sur le risque de fentes orales. Les auteurs observent, de plus, que le signal des interactions est plus fort parmi le groupe des FP.

## C. 2. Californie

L'ensemble des études gène-environnement sur le risque de fentes orales en Californie repose sur des données issues du Programme de surveillance des malformations congénitales en Californie (*California Birth Defects Monitoring Program*) de 1987 à 1989.

Les mères ont été interviewées par téléphone, et les échantillons de sang des enfants ont été collectés par les buvards recueillant des tâches de sang à la naissance (pour le test de Guthrie).

Trois études sur la population californienne se sont intéressées à l'interaction entre la consommation maternelle de tabac et un gène-candidat dans la survenue d'une fente orale.

Shaw et al. (1996b) rapportent que les mères fumant plus de 20 cigarettes par jour ont un risque augmenté d'avoir un enfant atteint de FP ou de FL/P, et encore davantage augmenté si l'enfant est porteur de l'allèle variant *C2* du *TGF $\alpha$* , suggérant la présence d'une interaction (tests d'hétérogénéité entre les strates définies par les allèles :  $p=0.15$  pour les FL/P isolées et  $p=0.09$  pour les FP isolées).

La consommation maternelle de tabac a été étudiée récemment par Lammer et al. (2004a, 2005) en interaction avec des gènes impliqués dans la biotransformation de composés toxiques dérivés du tabac (N-acétyl-transférases, *NAT*, et Glutathion-s-transférases, *GST*). Pour chacun des deux polymorphismes du *NATI* (*1088AA* et *1095AA*), les auteurs ne rapportent pas d'interaction statistiquement significative. Ils observent, parmi les enfants de mères fumeuses, un excès de risque de FL/P en particulier associé à l'homozygotie pour chaque variant (OR d'environ 4, pour l'estimation réalisée versus les enfants porteurs de l'homozygotie sauvage et de mères non fumeuses).

Lammer et al. (2005) ne réalisent pas de test statistique pour l'interaction entre la consommation de tabac et les gènes *GSTM1* et *GSTT1*. Ils notent cependant un risque de FL/P associé à la présence conjointe de la délétion à l'état homozygote du gène *GSTT1* de l'enfant et de la consommation maternelle de tabac, deux fois plus élevé qu'il ne serait sous l'hypothèse d'une interaction multiplicative.

La majorité des études californiennes a évalué le risque de fentes orales associé à l'utilisation de multivitamines des mères pendant la grossesse en interaction avec un facteur génétique, tel que *TGF $\alpha$* , *MTHFR*, *RFC1* et *NATI*.

Shaw et al. (1998b) rapportent des excès de risque de FL/P isolées (OR=3.0 ; IC95% :1.4,6.6) ou de FP isolées (OR=2.6 ; IC95% :1.0,7.7) parmi les enfants porteurs de l'allèle rare du *TGF $\alpha$*  et de mères non supplémentées en multivitamines, comparés aux enfants homozygotes pour l'allèle commun du gène et de mères supplémentées, suspectant la présence d'une interaction gène-nutriment (tests non significatifs d'hétérogénéité entre les strates de mères utilisatrices ou non de suppléments en multivitamines). Les auteurs notent cependant qu'il n'existe pas de liens biologiques connus entre le gène du *TGF $\alpha$*  et un composant de l'apport

en multivitamines. L'hypothèse *a priori* de l'étude de cette interaction s'appuie sur des résultats précédents des mêmes auteurs (1996a) qui suggèrent l'existence d'une interaction, dans l'étiologie des fentes orales, entre le gène du *TGF $\alpha$*  et la consommation maternelle de tabac, susceptible d'engendrer une décroissance de la concentration de folate dans le sérum.

La prise de multivitamines de la mère pendant la grossesse a ensuite été testée en interaction avec le gène *MTHFR*, impliqué dans le métabolisme des folates, et son polymorphisme *C677T* (Shaw et al. 1998a ou Wyszynski et Diehl 2000, Shaw et al. 1999c). Les études ne rapportent aucune interaction significative dans l'étiologie des fentes orales, mais les auteurs suspectent un effet accru de l'homozygotie *TT* de l'enfant comparé à l'homozygotie *CC* parmi les mères non supplémentées dans la survenue d'une FL/P (OR=1.4 ; IC95% :0.6,3.6).

Aucune interaction n'est suggérée avec le gène *RFC1* impliqué dans le transport des folates (Shaw et al. 2003c).

Enfin, l'influence des polymorphismes du *NAT1* dans la survenue d'une fente orale a été examinée en relation avec la prise de multivitamines de la mère (Lammer et al. 2004b), puisque la forme la plus rare du gène conduit à un **catabolisme** plus important des folates. Aucune interaction significative n'est rapportée ; les auteurs suggèrent toutefois un risque augmenté de FL/P chez les enfants homozygotes pour le variant du polymorphisme *NAT1(1095)* comparé à l'homozygotie sauvage, parmi les mères non supplémentées (OR=2.9 ; IC95% :1.0,9.5).

Enfin, une étude a examiné le risque de fentes orales associées aux expositions chimiques professionnelles de la mère à 74 agents chimiques, en interaction avec deux gènes impliqués dans le métabolisme de phase II de xénobiotiques toxiques, *GST* et *NAT* (Shaw et al. 2003a). Les estimations des interactions impliquent des effectifs très faibles, conduisant à une interprétation prudente et seulement suggestive des résultats.

### C. 3. Iowa

L'étude de Romitti et al. (1999) s'appuie sur le Registre de malformations congénitales de l'Iowa (*Iowa Birth Defects Registry*) et inclut l'ensemble des naissances vivantes, mort-nées, et des avortements porteurs d'une FL/P ou une FP, et les naissances vivantes sans anomalies congénitales, entre 1987 et 1994.

Pour les naissances de 1987 à 1991, les mères ont répondu à un questionnaire puis ont été interviewées par téléphone, les échantillons d'ADN ont été collectés par prélèvement veineux ou tache de sang. Pour les naissances de 1992 à 1994, les mères ont répondu à un questionnaire sans interview et les échantillons d'ADN ont été collectés grâce à des prélèvements de cellules buccales.

Romitti et al. s'intéressent aux consommations maternelles de tabac et d'alcool en interaction avec trois gènes de l'embryogénèse, *TGF $\alpha$* , *TGF $\beta$ 3* et *MSX1*, dont plusieurs marqueurs pour *TGF $\beta$ 3* (*CA*, *X5.1*, *5'UTR.1*) et *MSX1* (*CA*, *X1.1*, *X1.3*, *X2.1*, *X2.4*). L'étude n'apporte pas d'évidence de présence d'interaction entre le *TGF $\alpha$*  et la consommation de tabac ou d'alcool dans la survenue d'une fente orale. En revanche, elle rapporte les présences significatives d'une interaction entre le *TGF $\beta$ 3* et la consommation de tabac sur le risque de FL/P, d'une interaction entre le *MSX1* et la consommation de tabac sur le risque de FP, et d'une interaction

entre le *MSX1* et la consommation d'alcool sur le risque de FL/P. Ces effets ne sont cependant pas observés pour l'ensemble des marqueurs d'un même gène.

#### C. 4. Danemark

Les études d'interactions gène-environnement dans l'étiologie des fentes orales au Danemark se basent sur les données d'une étude cas-témoins réalisée à l'échelle nationale entre 1991 et 1994.

L'ensemble des enfants nés vivants durant cette période, atteints d'une fente orale et sans autre malformation majeure, constitue les enfants-cas de l'étude. Les témoins sont les deux naissances vivantes, sans malformation majeure, précédant la naissance d'un enfant-cas dans le même hôpital. L'ensemble des inclusions dans l'étude se limite aux enfants dont les parents parlent couramment la langue danoise, reflétant ainsi une population génétiquement homogène. Les mères ont été interviewées pour la majorité à l'hôpital par une infirmière formée ou plus rarement à domicile. Les échantillons d'ADN des enfants ont été obtenus à l'aide de tâches de sang collectées sur des buvards à la naissance.

Les études sur cette population ne rapportent aucune évidence d'interaction gène-environnement dans l'étiologie des FL/P et des FP. La consommation maternelle de tabac pendant la grossesse a été étudiée en interaction avec le polymorphisme commun *TaqI* du *TGF $\alpha$*  de l'enfant (Christensen et al. 1999b), puis en interaction avec les deux marqueurs *CA* et *X5.1* du gène *TGF $\beta$ 3* et les deux marqueurs *CA* et *N8* du gène *MSX1* (Mitchell et al. 2001).

Mitchell et al. (2003) examinent l'effet de la vitamine A sur la survenue ou la prévention d'une fente orale en prenant en compte l'utilisation de suppléments vitaminiques et la consommation de foie, et en étudiant le gène du Récepteur Alpha de l'Acide Rétinoïque (*RARA*). Les auteurs n'observent pas d'hétérogénéité de l'effet de l'exposition à la vitamine A entre les enfants de génotype *RARA* différent, ne suggérant pas de présence d'interaction *RARA*-exposition à la vitamine A.

L'ensemble de ces résultats est reproduit par l'étude récente de Etheredge et al. (2005), dans laquelle les analyses statistiques exploitent l'hypothèse d'indépendance entre les facteurs génétiques et environnementaux, et permettent ainsi d'offrir possiblement une meilleure puissance statistique des tests d'interaction gène-environnement.

#### C. 5. Pays-Bas

Les études d'interaction gène-environnement réalisées aux Pays-Bas s'appuient sur les mêmes sources de données collectées entre 1998 et 2000 : les enfants atteints de fentes orales ont été recrutés à partir des neuf équipes spécialisées dans le diagnostic et traitement des fentes aux Pays-Bas ; les enfants témoins ont été inclus par l'intermédiaire des familles des cas (familles, connaissances, voisins) ou dans des garderies ou centres de soins pédiatriques proches de Nijmegen.

Les mères ont répondu à un questionnaire à domicile, puis interviewées par téléphone dans le cas de réponses manquantes trop nombreuses. Des échantillons d'ADN ont été collectés par prélèvements de cellules buccales chez les enfants et leurs parents. Des prélèvements d'urine

des mères ont été recueillis pour la détermination du statut faible ou rapide acétylateurs par *NAT2* (van Rooij et al. 2002).

Les études néerlandaises présentent leurs résultats à travers les risques estimés par une analyse stratifiée par le facteur environnemental dichotomique et ne rapportent pas de test d'interaction.

Van Rooij et al. (2001) suggèrent ainsi l'existence d'une interaction entre *GST* et tabac dans la survenue d'une fente orale, en montrant un effet conjoint suffisamment élevé de la délétion du gène *GSTT1* de l'enfant et/ou de la mère, à l'état homozygote, et du statut fumeur de la mère (pour être supérieur au produit des effets propres du polymorphisme ou du statut fumeur). Aucune interaction n'est suggérée entre la consommation maternelle de tabac et le gène du *CYP1A1* porté par l'enfant ou la mère (van Rooij et al. 2001), et le gène *NAT2* (van Rooij et al. 2002).

Van Rooij et al. (2003b) concluent à l'existence possible d'une interaction entre l'apport en acide folique de la mère pendant la grossesse (via les multivitamines ou l'alimentation) et le gène *MTHFR* de la mère, en montrant des risques élevés de FL/P parmi les mères ayant un faible apport en folates et un statut homozygote pour le variant des deux polymorphismes *C677T* et *A1298C* du gène *MTHFR*.

## C. 6. Norvège

Les études norvégiennes exploitent des données de 1996 à 1998 par un schéma d'étude cas-parents. Ces données sont issues d'une étude cas-témoins plus large, mise en place en Norvège de 1996 à 2001, actuellement en cours d'analyse (<http://dir.niehs.nih.gov/direb/studies/ncl/home.htm>).

Les familles des enfants-cas ont été recrutées à partir des deux centres de traitement de fentes orales en Norvège. Les mères ont répondu à un auto-questionnaire sur la période des trois mois suivant la naissance. Un total de 74% des familles éligibles a fourni un prélèvement sanguin de l'enfant (collecté au moment de la chirurgie initiale) et des parents (à l'hôpital).

Les deux études norvégiennes ne montrent pas d'interaction entre la consommation maternelle de tabac ou d'alcool et les gènes *TGF $\alpha$*  ou *MTHFR* sur le risque de FL/P et de FP (Jugessur et al. 2003b, 2003c). Les auteurs rapportent un risque relatif de FP associé à l'allèle *T* du polymorphisme *MTHFR C677T* (comparé au génotype *CC*) significativement supérieur parmi les mères supplémentées en multivitamines, comparé aux mères non supplémentées. Ce résultat est inattendu puisqu'il est généralement admis que l'allèle *T* du gène *MTHFR* est impliqué dans une activité réduite du métabolisme des folates, induisant un état similaire à un faible apport en folates. Cependant les auteurs notent que l'approche cas-parents proposé ne permet pas de distinguer si le risque est réduit par l'acide folique parmi les enfants sans l'allèle *T*, ou si l'acide folique produit un accroissement du risque parmi les enfants possédant l'allèle *T* (Jugessur et al. 2003b). Les auteurs observent un risque relatif de FP associé à l'homozygotie pour le variant du *TGF $\alpha$*  (comparé au génotype homozygote sauvage) supérieur et non significatif parmi les mères non supplémentées en multivitamines, comparé aux mères supplémentées. Enfin, aucune indication d'interaction entre l'apport en multivitamines et les gènes *TGF $\alpha$*  ou *MTHFR* (pour les deux polymorphismes *C677T* et *A1298C*) n'est rapportée sur le risque de FL/P.

## C. 7. Conclusion

La consommation maternelle de tabac est fortement suspectée de jouer un rôle dans la survenue d'une fente orale. De nombreux auteurs ont ainsi choisi d'explorer comment l'activité d'un gène impliqué dans le développement de la face est susceptible d'être modulée par ce facteur environnemental. Les premières études examinant le gène du *TGF $\alpha$*  en relation avec le risque de fentes orales suggèrent que la consommation maternelle de tabac modifie l'effet du gène. Cependant les études les plus récentes ne reproduisent pas ce résultat. Une méta-analyse des cinq études cas-témoins rapporte que l'existence d'une interaction entre le génotype de l'enfant du polymorphisme *TaqI* du gène *TGF $\alpha$*  et la consommation maternelle de tabac ne peut être suggérée que pour le groupe des FP (Zeiger et al. 2005). L'ensemble des études s'intéressant à l'interaction entre la consommation maternelle de tabac et un gène candidat dans la survenue d'une fente orale est répertorié dans le Tableau 9.

**Tableau 9: Résumé des résultats des études s'intéressant à la consommation maternelle de tabac en interaction avec un gène candidat**

Gène candidat	Auteur	Schéma d'étude	Suggestion d'une interaction* ?
<b>Gènes du développement</b>			
<i>TGF<math>\alpha</math></i>	Hwang et al. 1995	Cas-témoins	+ (FP)
	Shaw et al. 1996b	Cas-témoins	+
	Beaty et al. 1997	Cas-témoins	-
	Maestri et al. 1997	TDT	+
	Romitti et al. 1999	Cas-témoins	-
	Christensen et al. 1999b	Cas-témoins	-
	Beaty et al. 2002	TDT	-
<i>TGF<math>\beta</math>3</i>	Jugessur et al. 2003c	Triades	-
	Maestri et al. 1997	TDT	+
	Romitti et al. 1999	Cas-témoins	+ (FL/P)
	Mitchell et al. 2001	Cas-témoins	-
<i>MSX1</i>	Beaty et al. 2002	TDT	-
	Romitti et al. 1999	Cas-témoins	+ (FP)
	Mitchell et al. 2001	Cas-témoins	-
<i>GST</i>	Beaty et al. 2002	TDT	+
	Romitti et al. 1999	Cas-témoins	-
<b>Gènes de métabolisme de xénobiotiques toxiques</b>			
<i>GST</i>	van Rooij et al. 2001	Cas-témoins	+
	Lammer et al. 2005	Cas-témoins	+ (FL/P)
<i>NAT</i>	van Rooij et al. 2002	Cas-témoins	-
	Lammer et al. 2004a	Cas-témoins	+ (FL/P)

TDT : Test de Déséquilibre de Transmission ; Triades : schéma d'étude cas-parents avec analyse « pseudo-sibs » ou « cas-parents »

+ : oui ; - : non

\* Interaction sous l'échelle multiplicative

Peu d'études rapportent l'existence d'une interaction entre la consommation maternelle d'alcool pendant la grossesse et un gène candidat dans l'étiologie des fentes orales.

Enfin, la majorité des études présentées dans le Tableau 10 examinant l'association entre la prise de multivitamines (contenant de l'acide folique) par les mères et le risque de fentes orales suggèrent l'existence d'une susceptibilité génétique individuelle (de l'enfant ou de la mère) par le gène du *TGF $\alpha$*  (Shaw et al. 1998b, Jugessur et al. 2003c) ou par le gène *MTHFR* (Shaw et al. 1998a, van Rooij et al. 2003b, Jugessur et al. 2003b).

Tableau 10 : Etudes épidémiologiques d'interaction gène-environnement dans l'étiologie des fentes orales

Référence	Facteur environnemental	Gène candidat	Sujets disponibles pour l'interaction	Résultats : Test d'interaction* et Force de l'effet conjoint**
<b>MARYLAND</b>				
(1984-1992) Hwang et al. 1995	Tabac maternel	<i>TGF<math>\alpha</math></i> de l'enfant ( <i>TaqI</i> )	113 FL/P 69 FP 281 témoins	Pas de test d'interaction FP : OR++ pour allèle variant et tabac = 7.0[1.8-28]
(1992-1996) Beaty et al. 1997	Tabac maternel	<i>TGF<math>\alpha</math></i> de l'enfant ( <i>TaqI</i> )	75 FL/P 46 FP 86 témoins	<i>p</i> interaction ns (25% de puissance pour détecter la même interaction que Hwang et al. 1995)
(1992-1996) Maestri et al. 1997	Tabac maternel	<i>TGF<math>\alpha</math></i> <i>TGF<math>\beta</math>3</i> <i>BCL3</i> (4 marqueurs) <i>RARA</i> <i>F13A</i>	160 trios cas-parents	<i>p</i> interaction (ou hétérogénéité de transmission) = 0.06 pour <i>TGF<math>\alpha</math></i> et Tabac = 0.04 pour <i>TGF<math>\beta</math>3</i> et Tabac
(1992-1998) Beaty et al. 2002	Tabac maternel	<i>TGF<math>\alpha</math></i> <i>TGF<math>\beta</math>3</i> <i>MSX1</i> <i>BCL3</i>	-130 trios cas-parents -198 trios cas-parents -166 trios cas-parents -207 trios cas-parents	<i>p</i> interaction (ou hétérogénéité de transmission) = ns = ns = 0.03 = ns
(non précisés) Fallin et al. 2003	Tabac maternel	<i>MSX1</i> (5 haplotypes)	(analyse cas-seuls) 141 FL/P 63 FP	<i>p</i> interaction Haplotype et Tabac < 0.05 pour l'ensemble haplotypes étudiés de <i>MSX1</i>
<b>IOWA (1987-1994)</b>				
Romitti et al. 1999	Tabac maternel Alcool maternel	<i>TGF<math>\alpha</math></i> <i>TGF<math>\beta</math>3</i> (3 marqueurs) <i>MSX1</i> (4 marqueurs)	Selon marqueurs : de 90 à 118 FL/P de 39 à 51 FP de 166 à 338 témoins	<i>p</i> interaction <0.05 pour FL/P : <i>TGF<math>\beta</math>3</i> et Tabac ; <i>MSX1</i> et alcool pour FP : <i>TGF<math>\beta</math>3</i> et Tabac ; <i>MSX1</i> et Tabac

[Suite Tableau 10 : Etudes épidémiologiques d'interaction gène-environnement dans l'étiologie des fentes orales]

Référence	Facteur environnemental	Gène candidat	Sujets disponibles pour l'interaction	Résultats : Test d'interaction* et Force de l'effet conjoint**
<b>CALIFORNIE (1987-1989)</b>				
Shaw et al. 1996b	Tabac maternel (mères fumant + de 20 cigarettes/jour)	<i>TGF<math>\alpha</math></i> de l'enfant ( <i>TaqI</i> )	191 FL/P isolées 87 FP isolées 379 témoins	$p$ interaction = ns FL/P : OR++ pour allèle variant et tabac (+20cig./j.) = 6.1 [1.1-37] FP : OR++ pour allèle variant et tabac (+20cig./j.) = 9.0 [1.4-62]
Shaw et al. 1998b	Multivitamines	<i>TGF<math>\alpha</math></i> de l'enfant ( <i>TaqI</i> )	255 FL/P isolées 106 FP isolées 557 témoins	$p$ interaction = ns FL/P : OR++ pour allèle variant et non supplément. = 3.0 [1.4-6.6] FP : OR++ pour allèle variant et non supplément. = 2.6 [1.0-7.7]
Shaw et al. 1998a Wyszynski et Diehl 2000	Multivitamines contenant acide folique	<i>MTHFR</i> de l'enfant ( <i>C677T</i> )	310 FL/P 383 témoins	$p$ interaction = ns OR++ pour <i>TT</i> et non supplément.
Shaw et al. 1999c	Multivitamines contenant acide folique	<i>MTHFR</i> de l'enfant ( <i>C677T</i> )	117 FP 383 témoins	Pas de test d'interaction Aucun effet conjoint suggéré
Shaw et al. 2003c	Multivitamines contenant acide folique	<i>RFC1</i> de l'enfant ( <i>A80G</i> )	300 FL/P 120 FP 362 témoins	Pas de test d'interaction Aucun effet conjoint suggéré
Shaw et al. 2003a	74 groupes chimiques d'exposition professionnelle de la mère	<i>GSTM1</i> <i>GSTT1</i> <i>NAT1</i> <i>NAT2</i>	Non précisés	Pas de test d'interaction Quelques effets conjoints élevés et significatifs, en particulier pour le groupe des FP
Lammer et al. 2004b	Multivitamines contenant acide folique	<i>NAT1</i> (1088) <i>NAT1</i> (1095)	- 298 FL/P ; 123 FP ; 298 témoins - 279 FL/P ; 112 FP ; 281 témoins	$p$ interaction = ns pour FL/P : OR++ pour <i>NAT1</i> homozygote (1095) et non supplément
Lammer et al. 2004a	Tabac maternel	<i>NAT1</i> (1088) <i>NAT1</i> (1095) <i>NAT2</i> (2 phénotypes : s/r)	- 303 FL/P ; 282 témoins - 283 FL/P ; 299 témoins - 297 FL/P ; 123 FP ; 288 témoins	$p$ interaction = ns pour FL/P : OR++ pour <i>NAT1</i> (1088) homozygote et Tabac OR++ pour <i>NAT1</i> (1095) homozygote et Tabac
Lammer et al. 2005	Tabac maternel	<i>GSTM1</i> <i>GSTT1</i>	297 FL/P 125 FP 294 témoins	Pas de test d'interaction pour FL/P : OR++ pour délétion <i>GSTT1</i> (enfant) et Tabac

**[Suite Tableau 10: Etudes épidémiologiques d'interaction gène-environnement dans l'étiologie des fentes orales]**

Référence	Facteur environnemental	Gène candidat	Sujets disponibles pour l'interaction	Résultats : Test d'interaction* et Force de l'effet conjoint**
<b>DANEMARK (1991-1994)</b>				
Christensen et al. 1999b	Tabac maternel	<i>TGF<math>\alpha</math></i> de l'enfant ( <i>TaqI</i> )	188 FL/P 64 FP 444 témoins	$p$ interaction = ns
Mitchell et al. 2001	Tabac maternel Alcool maternel	<i>TGF<math>\beta</math>3</i> (2 marqueurs) <i>MSX1</i> (2 marqueurs)	Non précisés	$p$ interaction = ns
Mitchell et al. 2003	Multivitamines contenant vitamine A Consommation de foie	<i>RARA</i>	109 FL/P 272 témoins	$p$ interaction additive = ns
<b>PAYS-BAS (1998-2000)</b>				
van Rooij et al. 2001	Tabac maternel	<i>CYP1A1</i> <i>GSTT1</i> Chez l'enfant et la mère	Selon marqueurs (enfants et mères) : de 93 à 100 cas de 84 à 89 témoins	Pas de test d'interaction OR ++ pour délétion <i>GSTT1</i> (enfant et/ou mère) et Tabac
van Rooij et al. 2002	Tabac maternel Médicaments	<i>NAT2</i>	45 cas 72 témoins	Pas de test d'interaction Aucun effet conjoint suggéré
van Rooij et al. 2003b	Multivitamines contenant acide folique Folate alimentaire	<i>MTHFR</i> ( <i>C677T</i> ) ( <i>A1298C</i> ) Chez l'enfant et la mère	Selon marqueurs (enfants et mères) : de 86 à 137 FL/P de 110 à 164 témoins	Pas de test d'interaction OR ++ et significatifs : pour <i>677TT</i> maternel et non supplément. et/ou folate alim. faible pour <i>1298CC</i> maternel et non supplément. et folate alim. faible
<b>NORVEGE (1996-1998)</b>				
Jugessur et al. 2003b	Multivitamines contenant acide folique	<i>MTHFR</i> ( <i>C677T</i> ) ( <i>A1298C</i> ) Chez l'enfant et la mère	173 trios cas-parents FL/P 88 trios FP	Pour FP : $p$ interaction = 0.03 avec RR( <i>T</i> foetal et supplément.) > RR( <i>T</i> foetal et non supplément.)
Jugessur et al. 2003c	Tabac maternel Alcool maternel Multivitamines contenant acide folique	<i>TGF<math>\alpha</math></i> ( <i>TaqI</i> )	173 trios FL/P 88 trios FP	$p$ interaction = ns pour FP: RR ++ pour enfant <i>TGF<math>\alpha</math></i> -homozygote et non supplément.

ns : non statistiquement significatif

\* : Test d'interaction sous l'échelle multiplicative par défaut

\*\* OR ++ : OR de l'effet conjoint positif et supérieur au produit des OR des effets propres du gène et du facteur environnemental

RR ++ : RR positif et supérieur au RR de l'effet génétique sur l'autre strate du facteur environnemental

## D. Bilan

Les études recherchant des facteurs environnementaux ou des facteurs génétiques dans l'étiologie des fentes orales sont nombreuses et une grande diversité de population est impliquée dans ces études. En revanche, encore peu de populations disposent de données suffisantes pour l'étude d'interaction entre ces facteurs. De façon générale, les études évaluant une interaction gène-environnement dans l'étiologie des fentes orales manquent de concordance dans les résultats et de puissance statistique pour pouvoir dresser des conclusions définitives. Les associations sont modestes, parfois inconstantes et les résultats observés sont seulement suggestifs.

La majorité des études cas-témoins d'interaction gène-environnement suit un protocole d'étude lourd, et les populations concernées sont déjà bien étudiées. Ces études rapportent généralement un triplet de résultats : l'association entre le risque de fentes orales et le facteur environnemental, le gène et l'interaction gène-environnement.

De façon équivalente aux mesures d'association avec un facteur environnemental, l'estimation et l'interprétation d'une interaction gène-environnement sont sensibles et fragilisées par des erreurs de classement de l'exposition, la présence d'expositions multiples, ou une période d'exposition mal définie.

Dans certaines études et de façon identique aux études d'association génétique, plusieurs marqueurs d'un même gène sont utilisés pour l'étude d'une interaction gène-environnement, mais tous ne rapportent pas le même résultat. Certains auteurs expliquent cette observation par des degrés de **déséquilibre de liaison** différents entre les marqueurs et le vrai locus de la maladie, et/ou entre les populations. Cependant, les connaissances actuelles sur ces variabilités sont encore très limitées.

Deux schémas d'étude, cas-témoins et triades (ou cas-parents), sont utilisés selon la disponibilité des sujets et de leurs informations.

Certains auteurs ne rapportent pas de test statistique d'interaction, et suggèrent la présence d'une interaction en observant une hétérogénéité de l'association du facteur génétique avec le risque de fentes orales entre chaque strate du facteur environnemental (ou de façon symétrique, une hétérogénéité de l'association du facteur environnemental avec le risque de fentes orales entre chaque strate du facteur génétique). Lorsqu'un test d'interaction est rapporté, les auteurs s'appuient sur l'hypothèse de l'échelle multiplicative.

Lorsqu'un résultat d'une interaction gène-environnement est suggéré dans l'étiologie des fentes orales, les hypothèses de plausibilité biologique sont établies à partir de résultats issus d'expérimentation animale ou de connaissances biochimiques et moléculaires chez l'homme adulte. Peu de connaissances sont actuellement disponibles pour discuter de l'activité et l'expression embryonnaires précoces des gènes étudiés chez l'humain.

Enfin, quelques études, parmi les plus récentes, suspectent un rôle des gènes de la mère sur l'environnement *in utero* du fœtus, en interaction ou non avec des expositions maternelles ; il s'agit de gènes impliqués dans le transport ou le métabolisme de facteurs environnementaux.

# 3

## Etat de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement

Dans les études épidémiologiques traditionnelles, une interaction gène-environnement est définie comme « un effet d'un facteur environnemental sur le risque de maladie différent selon le génotype de l'individu » ou, de façon symétrique, « un effet du génotype sur le risque de maladie différent selon l'exposition des sujets au facteur environnemental ».

La présence ou l'absence d'une interaction statistique s'évalue selon une échelle de mesure multiplicative ou additive. En termes de risque de maladie, il y a présence d'interaction multiplicative si l'effet conjoint de deux facteurs A et B est différent du produit des effets propres de A et B. Il y a présence d'interaction additive si l'effet conjoint de A et B est différent de la somme des deux effets propres moins un. Ainsi, en présence d'une interaction gène-environnement, les effets propres d'un caractère génétique et d'une exposition environnementale ne permettent pas de prédire l'effet observé de la présence simultanée des deux facteurs.

Le choix de l'échelle de mesure a été intensivement débattu dans la littérature épidémiologique (revue rapide par Ottman 1996) ; certains auteurs recommandent l'utilisation de l'échelle multiplicative lorsque les facteurs agissent à des étapes distinctes, d'autres conseillent l'échelle multiplicative pour les recherches à visée étiologique. En l'absence de consensus, il nous semble pertinent de présenter les estimations sous la forme stratifiée afin d'observer l'hétérogénéité des résultats, puis de préciser l'échelle utilisée lors de l'interprétation d'un résultat de test d'une interaction, comme Greenland et Rothman (1998) le recommandent.

Cette définition de l'interaction gène-environnement est strictement méthodologique. L'étude statistique d'une interaction doit être accompagnée d'une interprétation biologique pour contribuer à la compréhension des mécanismes. Cependant, un modèle d'interaction statistique sur le risque d'une maladie peut être obtenu par différents modèles de mécanismes biologiques. Ottman (1990) décrit quatre modèles de relation entre un gène, une exposition environnementale et le risque d'une maladie, biologiquement plausibles et impliquant possiblement la détection statistique d'une interaction. (1) L'effet de synergie : les deux facteurs génétique et environnemental augmentent le risque de maladie chacun séparément, et le risque est plus fort si les deux sont simultanément présents. (2) Le caractère génétique étudié exacerbe l'effet du facteur de risque environnemental alors que le génotype n'a pas d'effet parmi les individus non exposés. (3) L'exposition environnementale exacerbe l'effet du gène de susceptibilité à la maladie, et n'a pas d'effet sans le génotype à risque. (4) L'interaction dite « pure » : seule la présence conjointe des deux facteurs affecte le risque. De nombreuses autres possibilités de modèles d'interaction gène-environnement existent et se

complexifient en présence d'effets protecteurs des facteurs génétiques et environnementaux (Ottman 1996).

Nous ciblerons cette section sur les méthodes d'analyse statistique utilisées récemment dans la discipline de l'épidémiologie intégrant les informations génétiques. Une variété de schémas d'étude et de modèles statistiques est à ce jour développée pour évaluer une association entre un gène et une maladie ou une interaction entre un gène et un facteur environnemental sur le risque d'une maladie. Cette section se limitera spécifiquement aux schémas d'étude cas-témoins et cas-parents, qui se distinguent par leur groupe de témoins, issus respectivement, d'un recueil externe si les témoins sont recrutés dans la population-source de l'étude, ou d'un recueil interne lorsque le groupe de témoins est construit à partir des informations des parents des enfants-cas. Nous passerons en revue les principales méthodes proposées dans la littérature, et tenterons de décrire leur principe de fonctionnement, leurs hypothèses, leurs limites, leur performance et faisabilité et leurs évolutions. Nous présenterons plus en détails les quelques méthodes statistiques qui nous semblent les plus adaptées au contexte du projet d'étude de malformations congénitales, comme les fentes orales.

Avant toute évaluation statistique d'une interaction gène-environnement, certains auteurs épidémiologistes privilégient l'exploitation d'un concept dénommé « la randomisation mendélienne » (Clayton et McKeigue 2001). Nous décrivons en première partie de cette section, les principes et les limites de ce concept.

## **A. Randomisation mendélienne**

### **A. 1. Les lois de Mendel**

L'ensemble des observations empiriques de G. Mendel, établies en 1865, a constitué le fondement de la théorie chromosomique de l'hérédité. Deux lois fondamentales y sont décrites : la première loi, appelée de ségrégation, décrit l'hérédité d'un seul trait. Mendel postule l'existence d'un facteur qui se transmet des parents à l'enfant à travers les gamètes avec différents traits possibles. Ces facteurs sont aujourd'hui appelés gènes. Il fait l'hypothèse que ces facteurs sont présents par paires pour chaque individu et que chaque facteur possède une chance égale d'être transmis.

La seconde loi, la loi d'indépendance, décrit le comportement conjoint de loci qui contrôlent deux traits différents. Les allèles à un locus « se distribuent » de façon indépendante des allèles de l'autre locus. Cette proposition est maintenant connue pour être vraie strictement pour les loci de chromosomes différents ou suffisamment éloignés pour ne pas être en déséquilibre de liaison. Parce que le génome est de taille immense, la seconde loi de Mendel est usuellement convenable et forme ainsi une base d'hypothèses fondamentales pour le développement de toute méthodologie faisant intervenir des notions génétiques.

## A. 2. Le concept de la randomisation mendélienne

Le concept de « randomisation mendélienne » exploite la seconde loi de Mendel et garantit une distribution aléatoire des allèles lors de la formation des gamètes. Elle résulte ainsi en des associations aléatoires entre des loci quelconques dans une population (lorsqu'ils sont assez éloignés pour ne pas être en déséquilibre de liaison, en particulier si ceux-ci sont sur le même chromosome).

L'idée de la randomisation mendélienne se visualise plus clairement dans un schéma d'héritage d'un enfant de ses deux parents : dans le cas d'un locus biallélique (allèle  $a$ , allèle muté  $A$ ), un parent hétérozygote a autant de chance de transmettre l'allèle  $a$  ou l'allèle  $A$  à son enfant. Ainsi un couple de parents hétérozygotes  $aA$  peut transmettre, selon les lois mendéliennes, trois possibilités de génotypes  $aa$ ,  $aA$  et  $AA$ , avec respectivement les probabilités suivantes de  $\frac{1}{4}$ ,  $\frac{1}{2}$  et  $\frac{1}{4}$ .

## A. 3. Contribution de la randomisation mendélienne à l'épidémiologie

Cette attribution aléatoire des allèles contribue à surmonter certaines limites des techniques classiques de l'épidémiologie. Le génotype d'un sujet ne peut être lié *a priori* à aucun facteur de confusion environnemental, tel que le niveau socio-économique, la consommation de tabac... L'erreur potentielle dans la mesure du génotype est liée au génotypage lui-même : une technique de génotypage faiblement reproductible, un statut cas-témoins connu du personnel sont des exemples de sources d'erreur potentielle.

Dans le cadre d'une étude rigoureusement organisée et coordonnée, l'association entre un génotype et une maladie ne peut être sujette à des biais de confusion. La randomisation mendélienne offre ainsi les mêmes avantages que les essais d'intervention randomisés, qui sont construits de telle sorte que les groupes de sujets aient une distribution similaire des divers facteurs de confusion, qu'ils soient mesurés ou non.

Certains auteurs font du concept de la randomisation mendélienne une « renaissance » ou un « tournant » de l'épidémiologie. Dans ces débats et réflexions, ils développent une idée prometteuse qui place l'épidémiologiste à un carrefour de plusieurs disciplines.

Pour un gène polymorphique dont la fonction connue altère le métabolisme ou la régulation d'un facteur environnemental, l'association mesurée avec la maladie permet d'identifier une voie métabolique causale, contribuant ainsi directement à la compréhension du rôle du facteur environnemental dans l'étiologie de la maladie. Ce phénomène est parfois décrit sous le terme d'« interaction biologique » gène-environnement. La capacité de la randomisation mendélienne à contourner les erreurs de mesure et les biais « classiques » des études épidémiologiques qui s'intéressent aux facteurs environnementaux, peut être finalement exploitée pour étudier l'effet suspecté d'un facteur environnemental sur une maladie. Ce concept fournit ainsi un moyen prometteur d'examiner ou de découvrir les effets d'une exposition sur le risque d'une maladie.

#### A. 4. Limites de la randomisation mendélienne

Alors que les comportements ou les expositions environnementales ne peuvent déterminer le génotype d'un sujet, il est possible à l'inverse que le génotype influence un comportement (engendrant par exemple une aversion pour l'alcool), ou un paramètre physiologique (comme le niveau de cholestérol dans le sang). Le variant du gène d'intérêt peut également être associé, ou en déséquilibre de liaison, avec le variant d'un autre polymorphisme qui détermine un comportement à risque de maladie. Aussi, la contribution de la randomisation mendélienne à l'épidémiologie se trouve limitée dans le cas de gène **pléiotropique** (c'est-à-dire un polymorphisme qui est impliqué dans différentes activités métaboliques et qui peut aboutir à plusieurs **phénotypes** spécifiques).

Il est alors impératif d'identifier les loci en déséquilibre de liaison au locus d'intérêt, puis de distinguer les divers processus enzymatiques activés par le gène étudié. C'est pourquoi, le développement des recherches biologiques, moléculaires et biochimiques, sur les rôles et la densité des marqueurs des gènes régulant le transport, le métabolisme ou la bio-disponibilité de substances exogènes ou endogènes issues d'exposition environnementale, doit être encouragé.

Finalement, même si une association génétique ne démontre pas une association causale avec la maladie, si elle est biologiquement plausible, c'est-à-dire si le polymorphisme d'intérêt est suffisamment étudié et identifié, elle offre une première réponse dans la recherche d'une cible environnementale (exogène ou endogène).

### B. Mesure d'une association entre un gène et le risque d'une maladie

De façon générale, une association entre un génotype à un locus marqueur et une maladie peut se produire de différentes façons (1) le locus est lié à la maladie et les différents allèles peuvent porter des risques différents, (2) le locus n'est pas lié lui-même à la maladie mais est suffisamment proche du « vrai » locus, c'est-à-dire qu'il se trouve en déséquilibre de liaison avec le locus « coupable » (3) l'association s'explique par un biais de l'étude ou par un artefact statistique.

#### B. 1. Approche cas-témoins

##### B.1.i. Principe de la mesure d'une association

Deux approches sont disponibles, selon si l'on s'intéresse à l'estimation des effets liés aux génotypes (dans un cas bi-allélique :  $aa$ ,  $aA$ ,  $AA$ ) ou à estimer l'effet de l'allèle d'intérêt ( $A$ ). La première approche consiste à comparer les distributions des génotypes chez les cas et les témoins, parmi  $n$  sujets. La seconde compare la distribution allélique parmi les  $2n$  allèles des cas et des témoins.

Nous présentons ci-dessous les principes de mesure d'une association entre un gène et le risque de maladie, par comparaison cas-témoins, pour la situation la plus flexible qui s'intéresse aux effets de chaque génotype.

**Tableau 11: Distribution dans la population des génotypes selon le statut de la maladie**

En terme de probabilité			En terme d'effectif			
Génotype	Maladie		Génotype	Maladie		Total
	Oui	Non		Oui	Non	
<i>aa</i>	$\pi_{aa}$	$1 - \pi_{aa}$	<i>aa</i>	$n_{aa^{(+)}}$	$n_{aa^{(-)}}$	$n_{aa}$
<i>aA</i>	$\pi_{aA}$	$1 - \pi_{aA}$	<i>Aa</i>	$n_{aA^{(+)}}$	$n_{aA^{(-)}}$	$n_{aA}$
<i>AA</i>	$\pi_{AA}$	$1 - \pi_{AA}$	<i>AA</i>	$n_{AA^{(+)}}$	$n_{AA^{(-)}}$	$n_{AA}$

L'association est définie par la différence entre les **pénétrances**  $\pi_{AA}$ ,  $\pi_{aA}$ ,  $\pi_{aa}$  ( $\pi_i$  = la proportion de malades parmi les porteurs du génotype *i*) ou, de manière équivalente, entre les distributions des génotypes chez les cas et chez les témoins.

Une mesure commune est le risque relatif (RR) qui mesure l'écart des pénétrances par rapport à une pénétrance choisie comme référence. Si l'allèle *a* est l'allèle le plus courant, la force de l'association est mesurée par :

$$RR(AA \text{ vs } aa) = \pi_{AA} / \pi_{aa}$$

et

$$RR(aA \text{ vs } aa) = \pi_{aA} / \pi_{aa}$$

Une autre mesure plus couramment utilisée en épidémiologie est le rapport des *odds* (odds-ratio, noté OR), où l'*odds* de la maladie pour les porteurs d'un génotype X mesure le rapport entre la probabilité que la maladie soit présente parmi les porteurs du génotype et la probabilité que la maladie soit absente parmi les porteurs du génotype.

$$OR(AA \text{ vs } aa) = (\pi_{AA}/1-\pi_{AA}) / (\pi_{aa}/1-\pi_{aa})$$

$$\text{et } OR(aA \text{ vs } aa) = (\pi_{aA}/1-\pi_{aA}) / (\pi_{aa}/1-\pi_{aa})$$

Le RR et l'OR sont des indices définis au niveau de l'ensemble de la population. Dans le cas d'une maladie rare, la valeur du risque relatif peut être approchée par la valeur de l'odds-ratio. Cela découle du fait que les valeurs  $\pi$  sont petites, et ainsi que  $(1-\pi)$  est proche de 1.

Dans le cadre d'une enquête cas-témoins où le nombre de témoins est fixé en fonction du nombre de cas, la fréquence de la maladie pour chaque génotype ne peut pas être estimée. Le RR ne peut donc pas être calculé dans ce type d'enquête. En revanche, nous pouvons montrer que dans le calcul de l'odds-ratio, le génotype et la maladie ont un rôle symétrique, nous pouvons ainsi écrire l'odds-ratio comme le rapport des *odds* des génotypes. Cette écriture requiert l'estimation des fréquences des génotypes chez les malades et chez les non malades, informations disponibles dans une enquête cas-témoins. Finalement les calculs des odds-ratios se simplifient et s'expriment par les effectifs de chaque cellule.

$$OR(AA \text{ vs } aa) = (n_{AA^{(+)}}/n_{AA^{(-)}}) / (n_{aa^{(+)}}/n_{aa^{(-)}})$$

$$\text{et } OR(aA \text{ vs } aa) = (n_{aA^{(+)}}/n_{aA^{(-)}}) / (n_{aa^{(+)}}/n_{aa^{(-)}})$$

**B.1.ii. Modélisation de l'association : logistique ou log-linéaire.**

La modélisation des données cas-témoins par une régression logistique utilise le statut cas ou témoin comme variable aléatoire. Nous nous intéressons à la probabilité d'un sujet d'être un cas ( $\pi$ ), exprimée par la fonction inverse du Logit d'une combinaison linéaire incluant les variables indicatrices (notées I) de la présence/absence de chacun des trois génotypes.

$$\begin{aligned}\text{Logit}(\pi) &= \log(\pi / 1 - \pi) \\ &= {}^tX \beta \\ &= \beta_0 + \beta_{aa} I_{(\text{sujet porteur de } aa)} + \beta_{aA} I_{(\text{sujet porteur de } aA)} + \beta_{AA} I_{(\text{sujet porteur de } AA)}\end{aligned}$$

Par transformation, nous retrouvons l'expression des OR en fonction des coefficients sous la forme :

$$\begin{aligned}\text{OR}(AA \text{ vs } aa) &= \exp(\beta_{AA} - \beta_{aa}) \\ \text{et OR}(aA \text{ vs } aa) &= \exp(\beta_{aA} - \beta_{aa})\end{aligned}$$

En pratique, pour prévenir la situation d'*aliasing* (dans laquelle la matrice X n'est pas de plein rang), nous avons besoin de poser une contrainte sur les paramètres. Nous choisissons par exemple le génotype *aa* comme référence en fixant  $\beta_{aa} = 0$ .<sup>6</sup>

La modélisation log-linéaire se base sur une table de contingence (ici, 2x3) dans laquelle les fréquences sont supposées être distribuées selon une loi de Poisson, et le logarithme de leurs valeurs attendues obéit à un modèle linéaire. Le modèle (sous sa forme saturée) fait intervenir le statut cas ou témoin, le génotype et leurs interactions. Notons G1 et G2 les variables indicatrices associées respectivement aux génotypes *aA* et *AA*, M la présence (M=1, sujet cas) ou l'absence (M=0, sujet témoin) de la maladie. On note  $u_{mg}$  le nombre de sujets avec M=m et le génotype g.

$$\ln u_{mg} = \mu_0 + \beta_{01} G1 + \beta_{02} G2 + \mu_1 M + \beta_{11} MG1 + \beta_{12} MG2$$

Les coefficients  $\beta_{11}$  et  $\beta_{12}$  du modèle log-linéaire et les coefficients  $\beta_{aA}$  et  $\beta_{AA}$  (sachant la contrainte  $\beta_{aa} = 0$ ) du modèle logistique fournissent la même description du risque de maladie, et s'interprètent exactement de la même façon.

**B.1.iii. Test de l'association et Intervalle de confiance**

Plusieurs tests d'association entre le génotype et la maladie sont disponibles : le test de Wald, le test du rapport de vraisemblance et le test du score (Annexe 1). Ces trois tests sont asymptotiquement équivalents, le test de Wald étant le moins robuste sur de petits échantillons. Plus facile à mettre en œuvre, nous utilisons le test de rapport de vraisemblance qui mesure deux fois l'écart des log-vraisemblances des deux modèles à comparer (avec ou sans les indicatrices associées aux génotypes).

$$-2 \{ \log V(\text{modèle contraint sans G1 ni G2}) - \log V(\text{modèle complet}) \} \sim \chi^2_{ddl=2}$$

<sup>6</sup> Pour la suite du rapport, les écritures des modèles intègrent la contrainte.

Cette statistique suit asymptotiquement une loi du Khi-deux à  $k$  degrés de liberté, où  $k$  est égal à la différence de paramètres entre le modèle contraint et le modèle complet (ici  $k=2$ ).

Les modélisations logistiques et log-linéaires peuvent permettre de tester différentes hypothèses sur les effets des génotypes : modèle « multiplicatif » ( $\beta_{aA} = \beta_{AA}$  ou  $\beta_{01} = \beta_{02}$ , avec une variable codant le nombre de copies d'allèle  $A$  portées par le sujet), allèle  $A$  **dominant** ( $\beta_{aA} = \beta_{AA}$  ou  $\beta_{01} = \beta_{02}$ , avec une indicatrice codant la présence/absence de l'allèle  $A$  dans le génotype) ou **récessif** ( $\beta_{aa} = \beta_{aA}$  ou  $\beta_{01} = 0$ , avec une indicatrice codant la présence/absence de l'allèle  $a$  dans le génotype). Dans ces situations, un test de rapport de vraisemblance peut être employé et suit une distribution du Khi-deux à 1 degré de liberté.

Plusieurs méthodes d'estimation d'intervalle de confiance à 95% pour chaque coefficient peuvent être réalisées. Nous choisirons les intervalles de confiance usuels de Wald (Annexe 1), basés sur l'hypothèse de normalité des coefficients estimés, aisément calculables par l'encadrement du coefficient estimé plus ou moins 1.96 (fractile de la loi normale à 97.5%) multiplié par l'écart-type estimé du coefficient. Les intervalles de confiance des risques se calculent par les exponentielles des bornes inférieures et supérieures de l'intervalle de Wald.

#### **B.1.iv. Biais de stratification génétique de population**

La mesure d'association entre un gène et le risque de maladie dans une étude cas-témoins n'est pas exempte des biais méthodologiques traditionnels observés pour la mesure de l'association avec un facteur environnemental. Aussi nous portons une attention particulière au biais de stratification génétique de population.

##### iv.a. DEFINITION

Une stratification génétique de la population est présente dans un échantillon cas-témoins si le groupe des témoins est issu d'une population génétiquement différente de celle des cas. Elle peut être le résultat d'échantillonnage réalisé sans considération des groupes ethniques des sujets ou le résultat de groupes de cas et de témoins qui possèdent une répartition différente des groupes ethniques.

Il peut sembler aisé d'identifier et de contrôler cette forme de stratification par un ajustement des analyses ou par un appariement des cas et des témoins, sur leurs origines ethniques (Wacholder 2000), cependant une distinction plus ou moins fine des groupes ethniques peut ne pas suffire à détecter la présence d'une stratification de population. Aussi, dans le cas de population « génétiquement » jeune, il peut s'avérer nécessaire de connaître l'origine ethnique des ancêtres.

##### iv.b. QUEL BIAIS ?

La conclusion raisonnable par un généticien d'une association statistique entre le génotype défini par le locus du marqueur et le phénotype est qu'elle implique usuellement un lien physique entre le locus du marqueur et le locus de la maladie. En présence de stratification de population, il est possible d'observer des associations statistiques entre le phénotype et le marqueur qui pourtant n'a pas de lien physique avec le locus « causal ». En effet, le phénomène de stratification permet aux fréquences de l'allèle marqueur de varier parmi

différentes sous-populations. Si la maladie est plus prévalente dans l'une de ses sous-populations, elle sera statistiquement associée avec les allèles les plus fréquents dans cette sous-population.

Une solution alternative qui a souvent été proposée pour contourner ce problème de stratification est l'utilisation d'un schéma d'étude avec un groupe « interne » de témoins, comme les allèles non transmis des parents à l'enfant atteint. Ce groupe de témoins fictifs est issu de la même population génétique que les allèles des sujets cas, et offre ainsi une robustesse face au biais de stratification de population. Il s'agit de l'ensemble des études basées sur des données familiales, incluant l'enfant-cas, les parents ou les frères et sœurs, voire les grands-parents.

C'est seulement récemment que certains auteurs ont développé des méthodologies nouvelles et adaptées au schéma d'étude cas-témoins, permettant de détecter l'existence d'un biais de stratification ou/et de mesurer une association génétique en présence de ce biais.

#### iv.c. DETECTION ET MESURE D'ASSOCIATION EN PRESENCE D'UNE STRATIFICATION GENETIQUE DE POPULATION

L'approche basique pour détecter la présence d'une stratification génétique de population serait de vérifier l'hypothèse d'**équilibre d'Hardy-Weinberg** (HWE) pour divers loci suffisamment éloignés pour qu'ils ne soient pas en déséquilibre de liaison. S'ils dévient de l'équilibre HWE, nous pourrions admettre qu'il puisse exister une stratification génétique de la population. La difficulté de cette approche est qu'il s'agit ici de tests multiples et qu'elle nécessite une correction du degré de signification (*p-value*). De plus, il est maintenant bien établi que la puissance du test de HWE pour détecter une stratification de population est faible (Deng et al. 2001).

De façon générale, ces développements sont possibles grâce aux nouvelles techniques de la génomique qui offre aujourd'hui un accès rapide aux informations du génome humain. Ils s'appliquent à des jeux de données pour lesquels il est possible de réaliser des génotypages supplémentaires pour un panel de marqueurs non liés au locus de la maladie appelés « témoins génomiques » (traduit littéralement de *genomic control* (GC)).

Pritchard and Rosenberg (1999) proposent un test de la présence de stratification, adapté à des jeux de données cas-témoins, pour lesquels une association a été observée pour plusieurs loci. Ils considèrent que si une stratification est présente dans l'échantillon de population, alors les marqueurs additionnels non liés au locus de la maladie devraient montrer à leur tour une association avec le phénotype étudié. Les simulations des auteurs montrent que le test proposé obtient un risque d'erreur de première espèce de 6% à partir de 30 marqueurs bialléliques. L'approche développée par Devlin and Roeder (1999) est similaire.

La stratification de la population génère un phénomène de sur-dispersion (inflation de la variance) des statistiques de l'évaluation de l'association entre le locus d'intérêt et le risque de maladie. Certains auteurs proposent d'utiliser ces « témoins génomiques » pour estimer un facteur de correction et corriger la statistique de test (Bacanu et al. 2000, Devlin et al. 2001, Reich and Goldstein 2001).

Une autre approche, appelée *structured association* (SA), a pour principe d'utiliser les GC dans le but d'évaluer les informations manquantes, telles que le nombre de sous-populations qui constitueraient la stratification. Elle mélange des méthodologies statistiques d'approximation bayésienne ou/et classiques basées sur la vraisemblance des modèles. L'hypothèse de l'approche est que la population est hétérogène mais composée de sous-populations non observées homogènes. Elle estime ainsi la structure de la population (par un modèle latent) puis teste l'association. Il s'agit, entre autres, des travaux de Pritchard et al. (2000, 2001), de Satten et al. (2001) et de Hoggart et al. (2003).

Enfin, les GC peuvent permettre de constituer des groupes homogènes de sujets à l'aide d'un algorithme de classification, et les classes ainsi identifiées sont ensuite intégrées dans une analyse de régression comme ajustement (Schork et al. 2001).

L'ensemble des méthodes est intéressant et prometteur. Cependant encore peu d'applications sont actuellement publiées dans la littérature. Ces développements nécessitent davantage de recherche pour mieux apprécier leurs limites et leurs points forts.

#### iv.d. QUANTIFICATION DU BIAIS ET DISCUSSION

Il n'existe aujourd'hui que peu de publications d'études cas-témoins rapportant une preuve solide de l'existence de biais de stratification de population.

Une étude récente a proposé de quantifier le biais potentiel causé par une stratification de population (Wacholder et al. 2000) sur une étude cas-témoins s'intéressant à l'association entre un polymorphisme commun et le risque de cancer (vessie pour les hommes, sein pour les femmes) pour une population non hispanique des Etats-Unis d'ancêtres issus de huit pays européens. Les simulations suggèrent que le biais causé par la stratification de population est minimal et diminue avec le nombre de sous-population pris en compte. Aussi, lorsqu'une maladie est causée par différents gènes de différents loci non proches, la possibilité d'un biais important de stratification est sévèrement réduite (Wacholder et al. 2002a).

Cardon et Palmer (2003) suggèrent que, même en présence d'un biais de stratification de population dans une étude cas-témoins, la force du biais est faible, sauf si la structure est définie par une situation de différences extrêmes du risque de maladie et de la fréquence allélique entre les sous-populations.

L'ensemble de ces résultats est reproduit plus récemment par les simulations de l'étude de Wang et al. (2004), réalisées sur une population afro-américaine.

Nombreux auteurs développent l'idée que le biais de stratification de population dans les études cas-témoins correctement construits n'est pas une menace sérieuse pour la validité des conclusions, et modèrent ainsi l'enthousiasme porté sur les schémas d'étude basés sur les données familiales.

Morton et Collins (1998) considèrent que l'effet de stratification de population est négligeable à l'intérieur des groupes ethniques même s'ils sont vagues tels que les « Caucasiens, Noirs, Asiatiques et Hispaniques ».

L'idée majeure développée par Wacholder et al. (2002a) est que l'appartenance ethnique à des pratiques culturelles et linguistiques, lorsqu'elles sont auto-déclarées, est un meilleur indicateur que les marqueurs génétiques pour identifier une structure de la population : les auteurs concluent ainsi qu'une étude épidémiologique cas-témoins qui prend en compte ces appartenances ethniques ne peut être que faiblement concernée par un biais de stratification de population.

Rosenberg et al. (2002) renforce cette hypothèse par la construction d'une classification basée sur 377 loci, à partir de 1056 sujets issus de 52 populations définies selon la géographie, la culture et la langue (caractéristiques décrites par les sujets eux-mêmes). L'étude identifie 6 groupes principaux génétiquement homogènes dont 5 correspondent aux zones géographiques majeures définies par les délimitations des continents. Des sous-populations plus finement décrites à l'intérieur des zones géographiques majeures peuvent être distinguées dans la classification des loci, excepté pour le continent Européen pour lequel aucune structure génétique ne peut être distinguée (précisément l'étude isole trois sous-groupes génétiquement homogènes selon les 377 loci qui correspondent à la population basque, à la population de la Sardaigne, et la population restante).

Enfin, l'existence et l'ampleur du biais de stratification génétique de population dans les études cas-témoins de mesure d'association génétique a été et est encore largement débattu. Les développements méthodologiques (génomique ou statistique) permettant de contrôler ce biais doivent être améliorés et leur faisabilité évaluée, puisque les discussions des études cas-témoins traditionnelles feront toujours l'objet de cette critique.

## B. 2. Approche cas-parents

### B.2.i. Association mesurée par un déséquilibre de transmission allélique

#### i.a. PREMIERES APPROCHES

Répondant à la difficulté de construire un groupe approprié de témoins, Falk et Rubinstein (1987) proposent d'utiliser les allèles des parents non transmis à l'enfant-cas comme échantillon de témoins.

A partir des génotypes des parents et de l'enfant-cas, le principe de la méthode proposée est de reconstituer le « génotype non-transmis » et de comparer les fréquences de la présence de l'allèle d'intérêt entre les génotypes des cas (génotype constitué par les allèles transmis) et les génotypes des pseudo-témoins (constitué par les allèles non transmis).

**Tableau 12: Analyse non appariée des génotypes transmis et non transmis, pour la construction du test HRR de Falk et Rubinstein (1987 - Cas d'un gène candidat à 2 allèles a et A, où A est l'allèle d'intérêt)**

	Cas (enfant-cas)	Témoin (génotype non transmis reconstruit)	
Porteur de A	d	g	
Non porteur de A	n-d	n-g	
	n	n	2n génotypes

L'association entre l'allèle d'intérêt et le statut cas est estimée de façon équivalente aux études cas-témoins traditionnelles par le calcul usuel du produit croisé des effectifs du tableau 2x2. Cet odds-ratio a été appelé initialement *Haplotype Relative Risk* (HRR) puis *Genotype-based HRR* (GHRR).

$$\text{GHRR (Statistique de test)} = 2n (d-g)^2 / (d+g)(2n-d-g)$$

D'autres auteurs se sont inspirés de cette méthode et proposent des statistiques de test plus puissantes, qui diffèrent entre elles dans les détails. Ces travaux s'appuient sur les hypothèses d'indépendance de transmission entre les allèles pour chaque parent et d'indépendance de transmission des allèles entre les parents. Ils comptabilisent les 4n allèles transmis et non transmis aux n enfants-cas, et construisent un tableau 2x2 croisant la présence ou l'absence de l'allèle d'intérêt avec le statut cas (allèle transmis) ou témoin (allèle non transmis parmi les allèles parentaux). Il s'agit des travaux de Terwilliger et Ott (1992, *Haplotype-based HRR* ou HHRR) ou de Thomson (1995, *Affected Family-BAsed Controls* ou AFBAC).

$$\text{HHRR (Statistique de test)} = 4n (w-y)^2 / (w+y)(4n-w-y)$$

**Tableau 13: Analyse non appariée des allèles transmis et non transmis, pour la construction des tests étendus du HRR**

	Cas (allèle transmis)	Témoin (allèle non transmis)	
A	w	y	
a	2n-w	2n-y	
	2n	2n	4n allèles

Les statistiques GHRR et HHRR sont des statistiques usuelles de table de contingence 2x2 de Pearson, qui suivent approximativement une loi du Khi-deux à 1 degré de liberté.

Ces premières approches utilisant un groupe interne de témoins, restent cependant sensibles au biais de stratification génétique de population, puisqu'une structure ou un mélange de populations crée une dépendance entre les allèles transmis et non transmis si le locus marqueur n'est pas le locus de la maladie (Ewens et Spielman, 1995).

#### i.b. LES TESTS DE DESEQUILIBRE DE TRANSMISSION ALLELIQUE

L'idée du test du déséquilibre de transmission (*Transmission Disequilibrium Test*, TDT) développé par Spielman et al. (1993) est inspirée des tests précédents. Il offre une version appariée de l'analyse des allèles transmis et non transmis présentée dans le Tableau 13. Il s'agit d'évaluer la déviation possible aux lois mendéliennes de la transmission des allèles à partir des parents hétérozygotes à leur enfant-cas. Chaque paire d'allèles d'un parent hétérozygote est considérée comme une paire appariée constituée d'un allèle transmis et d'un allèle non transmis. Les deux allèles parentaux ont une chance égale d'être transmis s'il n'y a pas d'association (ou de déséquilibre de liaison entre le locus marqueur et le locus de la maladie), et inversement. Le test est ainsi insensible au biais de structure de population.

**Tableau 14: Analyse appariée des allèles transmis et non transmis, pour la construction du TDT**

		Allèles Non Transmis		
		A	a	
Allèle Transmis	A	a	b	w = a+b
	a	c	d	2n-w = c+d
		y = a+c	2n-y = b+d	2n

Le TDT est une statistique de test de McNemar pour séries appariées, qui suit approximativement une loi du Khi-deux à 1 degré de liberté pour un nombre important de paires discordantes (b+c) (nombre total de parents hétérozygotes, d'après le Tableau 14).

$$\text{TDT (Statistique de test)} = (b-c)^2 / (b+c)$$

L'odds-ratio (de transmission de l'allèle d'intérêt A comparé à l'allèle a) est mesuré de façon équivalente aux études appariées traditionnelles par le rapport b/c, tandis que les effectifs a et d issus des parents homozygotes n'apportent aucune information de transmission. Si b>c ou l'allèle A est préférentiellement transmis dans le groupe d'enfants-cas, nous rapportons que l'allèle A est à risque pour la maladie étudiée. Inversement, si l'allèle a est transmis préférentiellement à A, l'allèle A est dit protecteur.

L'idée initiale du développement du TDT est de localiser les gènes de la maladie par l'existence d'un déséquilibre de liaison entre le locus marqueur et le locus de la maladie. Ainsi, les auteurs présentent le TDT comme un test de liaison en présence d'une association (ou de déséquilibre de liaison) ou de façon équivalente comme un test d'association en présence de liaison. Ceci peut se comprendre de la façon suivante.

Soient deux loci, le locus de la maladie défini par deux allèles M/s et le locus marqueur A/a. L'allèle à risque de maladie M peut être associé à l'un ou l'autre des allèles du locus-marqueur et ainsi former deux haplotypes: M-A/s-a ou M-a/s-A. Un enfant est dit non-recombinant par rapport à l'un de ses parents hétérozygotes s'il conserve la même forme haplotypique, et recombinant s'il a hérité d'un gamète recombiné. Il n'existe pas de liaison entre les deux loci si les gamètes recombinés ou non recombinés ont autant de chance de se produire. Il existe une liaison s'il a été transmis chez l'enfant plus fréquemment un type spécifique d'haplotype. En l'absence d'association, nous accordons autant de chances pour les parents hétérozygotes de porter l'une ou l'autre forme haplotypique des deux loci. Sans la visibilité de l'information haplotypique chez l'enfant (en pratique, le locus de la maladie est inconnu), nous sommes dans l'impossibilité de classer l'enfant « recombinant » ou « non-recombinant », c'est-à-dire de savoir si, par exemple chez les enfants, l'allèle M est plus souvent associé à l'allèle A du locus-marqueur, et donc de détecter une liaison. Les allèles-marqueurs des parents sont alors transmis de façon aléatoire aux enfants quelle que soit la proportion de gamètes recombinés, et le TDT ne détecte aucun signal. Inversement, considérons la situation sans liaison et en présence d'association dans la population : les fréquences des formes haplotypiques chez les parents hétérozygotes sont disproportionnées, chaque forme haplotypique a autant de chance de transmettre un gamète recombiné et un gamète non recombiné, ce qui implique ainsi une transmission aléatoire des allèles-marqueurs des parents à l'enfant, et finalement aucun signal du TDT. Le test détecte un signal lorsque

l'apparente transmission des allèles des parents hétérozygotes dévie des proportions mendéliennes (1:1).

#### i.c. EXTENSIONS DU TDT

Le test de déséquilibre de transmission a été proposé initialement pour un cas simple d'un gène possédant deux allèles. De nombreux auteurs ont étendu ce test aux situations de locus marqueur avec plus de deux allèles. Il s'agit des travaux de Bickeboller et Clerget-Darpoux (1995), Sham et Curtis (1995), Schaid (1996), Spielman et Ewans (1996), Lazzeroni et Lange (1998), qui calculent tous une statistique du Khi-deux pour un test global de déséquilibre de transmission allélique.

Des méthodes basées sur des modélisations logistiques ont été proposées pour tester un déséquilibre de transmission allélique. Les travaux de Sham et Curtis (1995) considèrent pour chaque parent hétérozygote la probabilité de transmettre l'allèle k sachant que l'allèle j n'est pas transmis, notée  $p_{kj}$ . Le modèle logistique s'écrit sous la forme :

$$p_{kj} = \exp(b_k - b_j) / [1 + \exp(b_k - b_j)]$$

où  $(b_k - b_j)$  représente le log de l'odds-ratio de transmission de l'allèle k sachant le génotype parental 'kj'. Si  $(b_k - b_j) > 0$  alors l'allèle k est transmis de façon préférentielle à l'allèle j, correspondant au TDT. Pour K allèles, nous fixons arbitrairement  $b_K = 0$ .

D'autres auteurs proposent d'utiliser la forme conditionnelle de la régression logistique en modélisant la probabilité de transmission allélique expliquée par la présence ou l'absence de l'allèle d'intérêt, conditionnellement à chaque parent hétérozygote (Maestri et al. 1997). La modélisation logistique possède l'avantage d'évaluer et de tester simplement une hétérogénéité de la transmission allélique, par l'introduction supplémentaire dans le modèle d'un terme d'interaction (Maestri et al. 1997, Waldman 1999, Eaves et Sullivan 2001). Par exemple, Maestri et al. présentent une application de cet avantage en testant les interactions entre des marqueurs proches de loci de gènes *Transforming Growth Factor* et l'ethnicité, l'histoire familiale, la consommation maternelle de tabac ou le type de phénotypes, chez les enfants atteints de fentes orales. Le modèle utilisé s'écrit :

$$\text{logit}(p) = \alpha_0 + \alpha_s + \beta_k A_k + \gamma_k A_{ks} E$$

avec  $k=1, \dots, (K-1)$  et  $\beta_K = 0, \gamma_K = 0$ ,

où

p mesure la probabilité de transmission (le « cas » est défini par l'allèle transmis et le « témoin » par l'allèle non transmis),

$\alpha_s$  est le risque de base de chaque parent hétérozygote s (strate) de la régression logistique conditionnelle ( $\alpha_0$  est la constante globale du modèle),

$A_k$  code 1 pour la transmission de l'allèle k et 0 pour les autres, et E est la variable indicatrice du facteur environnemental (ou facteur d'hétérogénéité à laquelle l'étude s'intéresse).

L'exponentielle de  $\beta_k$  correspond à l'odds-ratio de transmission de l'allèle k chez les non exposés au facteur E ; l'exponentielle de  $(\beta_k + \gamma_k)$  est l'odds-ratio de transmission de l'allèle k chez les exposés au facteur E (ou  $\gamma_k$  est le coefficient d'interaction qui peut être testé par un test du rapport de vraisemblance).

D'autres sources potentielles d'hétérogénéité de transmission allélique peuvent être évaluées : effets d'origine parentale de l'allèle d'intérêt (Eaves et Sullivan 2001, section suivante E, l'hétérogénéité ethnique...

Enfin, plus récemment quelques auteurs ont proposé des extensions au TDT évaluant l'association entre un trait quantitatif et un déséquilibre de transmission (Allison 1997, Waldman 1999, Rabinowitz 1997). Des approches alternatives utilisent les génotypes des frères et sœurs non atteints des enfants-cas (Boehnke et Langefeld 1998, Schaid et Rowland 1998 qui offrent une très bonne revue de cette littérature, Spielman et Ewens 1998 qui présentent le S-TDT).

#### i.d. HYPOTHESES ET LIMITES DU TDT

La principale limite du TDT et de ses extensions est l'unité allélique de la transmission étudiée. Elle impose ainsi l'hypothèse d'effets multiplicatifs des allèles : si  $r$  est l'odds-ratio pour une copie de l'allèle, alors l'odds-ratio pour l'homozygotie pour l'allèle d'intérêt est calculé par  $r^2$ . Cette approche est puissante lorsque cette hypothèse est vérifiée. En revanche, elle peut aboutir à des biais sévères dans les autres situations (Schaid 1999a).

Le test utilise comme groupe d'allèles de référence, les allèles non transmis des parents. Il suppose implicitement que ce groupe joue un rôle négligeable dans la survenue de la maladie chez l'enfant.

Basé sur des taux de transmission, le calcul du TDT admet que les transmissions des parents hétérozygotes sont des événements indépendants, même si en général elles ne le sont pas (Umbach et Weinberg 2000). Par exemple, pour une famille de parents hétérozygotes, la probabilité que la mère transmette l'allèle d'intérêt à son enfant quand le père ne le transmet pas est différente de la probabilité que la mère transmette l'allèle d'intérêt à son enfant quand le père le transmet aussi (qui dépend du risque de porter deux copies), si l'allèle d'intérêt est lié à la survenue de la maladie.

Lorsque nous nous intéressons à évaluer une hétérogénéité de transmission, l'effet principal du facteur étudié, spécifique à l'enfant, ne peut pas être estimé par ce test ou ses extensions, puisque les « cas » et « témoins » sont identiques à chaque paire d'allèles parentaux et sont ainsi parfaitement appariés (excepté sur l'allèle d'intérêt).

### **B.2.ii. Approche génotypique**

#### ii.a. APPROCHE DES PSEUDO-SIBS

Certaines limites de l'étude de la transmission allélique peuvent être dépassées en considérant les génotypes des deux parents simultanément.

#### **Les pseudo-sibs**

Notons  $g_c$ ,  $g_m$  et  $g_p$  les génotypes respectifs de l'enfant-cas, de sa mère et de son père. Selon les règles de Bayes, la probabilité du génotype de l'enfant conditionnellement au statut cas et aux génotypes des parents, peut s'exprimer de la façon suivante :

$$P(g_c / g_m, g_p, \text{enfant cas}) = \frac{P(\text{enfant cas} / g_c, g_m, g_p)P(g_c / g_m, g_p)P(g_m, g_p)}{\sum_{g^* \in G} P(\text{enfant cas} / g^*, g_m, g_p)P(g^* / g_m, g_p)P(g_m, g_p)}$$

où  $G$  est l'ensemble des quatre génotypes possibles créées par les combinaisons potentielles des deux allèles de chaque parent : l'un est celui de l'enfant-cas, les trois autres sont fictifs, ce sont les *pseudo-sibs*.  $g^*$  est l'un de ces quatre génotypes.

En supposant que la probabilité d'être un enfant-cas sachant son génotype, ne dépend pas des génotypes des parents, l'équation se simplifie et s'écrit sous la forme :

$$P(g_c / g_m, g_p, \text{enfant cas}) = \frac{r(g_c)}{\sum_{g^* \in G} r(g^*)}$$

où  $r$  est le risque d'être un enfant-cas pour le génotype  $g_c$ .

En considérant la forme générale log-linéaire du risque pour le génotype  $g$ ,  $\log(r(g)) = X'\beta$ , la contribution d'un enfant-cas et de ses parents à la vraisemblance de l'échantillon constitué uniquement d'enfants-cas conditionnellement aux génotypes des parents, s'écrit :

$$L_i = \frac{\exp(X_i' \beta)}{\sum_{g^* \in G_i} \exp(X^* \beta)}$$

où :  $i$  est l'indice de la famille  $i$

$X_i$  est le vecteur codant le génotype observé du cas  $i$

$X^*$  est le vecteur codant le génotype  $g^*$

$G_i$  est l'ensemble des quatre génotypes possibles construits à partir des allèles parentaux de la famille  $i$ .

La vraisemblance de l'échantillon pour  $N$  familles est :  $L = \prod_{i=1}^N L_i$ . Elle correspond à la vraisemblance d'une régression logistique conditionnelle pour laquelle chaque cas est apparié aux trois témoins fictifs *pseudo-sibs*.

L'approche des *pseudo-sibs* a été introduite dans la littérature épidémiologique par Self et al. (1991). Dans le cas simplifié de deux allèles  $a$  et  $A$  au locus marqueur, résultant en trois génotypes possibles  $aa$ ,  $aA$  et  $AA$ , l'approche estime deux paramètres correspondant aux logarithmes des risques des deux génotypes (par exemple  $aA$  et  $AA$ ), relatifs au génotype de référence ( $aa$ ). Les effets des génotypes peuvent être évalués par des tests basés sur la vraisemblance (test du score proposé par Schaid et Sommers (1993), test du rapport de vraisemblance). L'estimation des paramètres peut être réalisée par la méthode du maximum de vraisemblance.

D'autres méthodes alternatives, plus simples mais moins efficaces, ont été proposées pour estimer les odds-ratios des génotypes. Il s'agit des travaux de Flanders et Khoury (1996), Sun et al. (1998) et Flanders et al. (2001).

## Les *mating-types* ou profils parentaux

Le test de Score développé par Schaid et Sommers (1993) et décrit à nouveau par Schaid (1999a) aborde l'approche génotypique d'une manière différente de celle de Self et al. (1991). Il met ainsi en valeur l'intérêt des *mating-types* parentaux dans la conceptualisation de la modélisation. Les auteurs définissent un *mating-type* par le profil des deux génotypes du couple parental (presque équivalent au nombre de copies d'allèle d'intérêt portées par le couple parental). Nous parlerons par la suite de « profil parental ».

Pour un gène à deux allèles  $a$  et  $A$ , six possibilités de profils parentaux peuvent se produire, aboutissant à dix catégories de génotypes de l'enfant-cas sachant le profil parental.

	Profil parental	Génotype de l'enfant
1	AA x AA	AA
2	AA x aA	aA, AA
3	AA x aa	aA
4	aA x aA	aa, aA, AA
5	aA x aa	aa, aA
6	aa x aa	aa

L'idée des auteurs est d'utiliser en tant qu'unité statistique, la famille décrite par le génotype de l'enfant sachant le profil parental. Ils construisent ainsi la vraisemblance conditionnelle suivante :

$$\ln L(r_1, r_2) = \sum_{i=2,4,5} \sum_{j=0,1,2} n_{ij} \log P_{j/i}(r_1, r_2)$$

$P_{j/i}(r_1, r_2)$  est la probabilité conditionnelle de la  $j^{\text{ème}}$  possibilité pour le génotype de l'enfant sachant le  $i^{\text{ème}}$  profil parental,

$n_{ij}$  correspond au nombre d'enfants de la  $j^{\text{ème}}$  possibilité de génotype de l'enfant pour le  $i^{\text{ème}}$  profil parental,

$r_1$  et  $r_2$  sont les risques de maladie pour l'enfant qui porte une ou deux copies de l'allèle d'intérêt, relativement à aucune copie.

Les profils parentaux 1, 3 et 6 n'apportent aucune information de transmission, seuls les profils parentaux 2, 4 et 5 (c'est-à-dire avec au moins un parent hétérozygote) sont utiles à la modélisation. Le conditionnement du modèle sur le profil génétique des couples parentaux permet d'attribuer à chaque famille un risque de base. Il assure ainsi une robustesse face à une éventuelle stratification de population, ou face à la possibilité de profils parentaux préférentiels dans la population générale, ces deux phénomènes engendrant des dépendances dans les transmissions maternelles et paternelles des allèles.

D'autres méthodes similairement conditionnées sur les génotypes parentaux ont été proposées. Il s'agit des méthodes FBAT (*Family-Based-Association Tests*) non paramétriques développées par Laird et al. (2000), Rabinowitz et Laird (2000), Lange et Laird (2002) (programme disponible sur: <http://www.biostat.harvard.edu/~fbat/default.html>).

## ii.b. METHODOLOGIE LOG-LINEAIRE (WEINBERG ET AL. 1998)

L'ensemble des méthodologies « cas-parents » décrites jusqu'ici, présente une limite inhérente aux études des maladies issues de la vie fœtale. La mère peut jouer un rôle crucial dans la survenue par exemple d'une malformation congénitale chez son enfant. L'allèle maternel peut modifier l'environnement intra-utérin du fœtus et indirectement provoquer une anomalie sur le fœtus qu'il ait hérité ou non cet allèle.

La méthode log-linéaire développée par Weinberg et al. (1998) est la première à intégrer ce rôle potentiel du génotype de la mère sur la survenue de la maladie chez son enfant. Cette méthode a été décrite aussi par Wilcox et al. (1998).

Le modèle proposé est adapté aux gènes bi-alléliques (deux allèles :  $a$  et  $A$ ). Notons  $M$ ,  $P$  et  $C$  le nombre de copies de l'allèle d'intérêt, de 0 à 2, porté par respectivement la mère, le père et l'enfant-cas. Les auteurs s'inspirent du raisonnement de Schaid et Sommers (1993), et à partir des six profils de couples parentaux, ils distinguent 15 possibilités de combinaisons de triades ( $M,P,C$ ) :

Type1 : (2,2,2)

Type2 : (2,1,2) (2,1,1) (1,2,1) (1,2,2)

Type3 : (2,0,1) (0,2,1)

Type4 : (1,1,2) (1,1,1)\* (1,1,0)

\* cette combinaison est ambiguë, car elle peut se produire de deux façons différentes : l'allèle de l'enfant provient de la mère ou l'allèle de l'enfant provient du père.

Type5 : (1,0,1) (1,0,0) (0,1,1) (0,1,0)

Type6 : (0,0,0)

### Scénario A

Différents modèles sont construits selon l'effet génétique testé. Lorsque le génotype hérité est responsable de l'effet, le modèle s'écrit de la façon suivante :

$$\ln(E[n_{M,P,C}]) = \gamma_j + \ln(2) I_{[M=P=C=1]} + \beta_1 I_{[C=1]} + \beta_2 I_{[C=2]} \quad (A)$$

$j$  est l'indice de la strate du profil parental. Le terme  $\ln(2) I_{[M=P=C=1]}$  correspond au multiplicateur de 2 pour la combinaison (1,1,1) et est donc déclaré en tant que variable *offset* avec un coefficient contraint de 1.

De façon identique à l'approche des pseudo-*sibs*, ce modèle permet d'estimer les risques propres à chaque génotype. Plus exactement, le risque relatif d'un enfant portant une copie de l'allèle d'intérêt, par rapport à aucune copie, se calcule par l'exponentielle du coefficient  $\beta_1$ ; le risque relatif d'un enfant avec deux copies de l'allèle d'intérêt se calcule par l'exponentielle du coefficient  $\beta_2$ . Les intervalles de confiance à 95% sont calculés à partir de la procédure usuelle de Wald (cf Annexe 1). L'effet du génotype de l'enfant est évalué par le test usuel du rapport de vraisemblance, avec l'hypothèse nulle suivante,  $H_0 : \beta_1=0$  et  $\beta_2=0$  :

$$2 \{ \log V(\text{modèle complet}) - \log V(\text{modèle contraint sans } I_{[C=1]} \text{ et } I_{[C=2]}) \} \sim \chi^2_{ddl=2}$$

L'écriture du modèle peut être modifiée selon le mode de transmission de l'allèle d'intérêt. Si le mode est dominant, les deux variables indicatrices,  $I_{[C=1]}$  et  $I_{[C=2]}$ , sont sommées, ce qui

revient à poser  $\beta_1=\beta_2$ . Dans le cas d'un mode récessif, nous éliminons la variable indicatrice définissant la présence d'une seule copie du variant chez l'enfant, ce qui signifie que nous posons  $\beta_1=0$ .

De façon équivalente au test du score de Schaid et Sommers (1993), seuls les profils 2, 4 et 5 sont informatifs ; les profils non informatifs jouent un rôle passif et n'ont pas d'influence sur les estimations, les écarts-types ainsi que sur les tests de significativité.

### Scénario B

Cette approche log-linéaire peut être complétée en intégrant les informations maternelles génotypiques.

$$\ln(E[n_{M,P,C}]) = \gamma_j + \ln(2) I_{[M=P=C=1]} + \beta_1 I_{[C=1]} + \beta_2 I_{[C=2]} + \alpha_1 I_{[M=1]} + \alpha_2 I_{[M=2]} \quad (B)$$

Les coefficients liés au génotype de l'enfant s'interprètent de la même façon que dans le scénario A. De plus, le risque relatif de l'effet maternel associé à une seule copie de l'allèle d'intérêt est égal à l'exponentielle du coefficient  $\alpha_1$ , tandis que celui associé à deux copies de l'allèle d'intérêt est égal à l'exponentielle de  $\alpha_2$ . Un test peut être réalisé par la méthode du rapport de vraisemblance pour tester les effets du génotype de l'enfant ou du génotype de la mère, de la même façon que pour le scénario A.

Les tests associés au génotype de l'enfant en présence ou non du génotype de la mère sont équivalents car leurs estimations sont orthogonales. La stratification du modèle par les profils parentaux permet d'estimer un risque de base pour chaque strate, ne nécessitant pas ainsi de poser l'hypothèse de l'équilibre d'Hardy-Weinberg pour l'utilisation du modèle.

#### ii.c. HYPOTHESES, PRINCIPE DE FONCTIONNEMENT ET LIMITES

Les trois méthodes décrites ci-dessus pour évaluer le rôle du génotype de l'enfant sont équivalentes, et traduisent des modes de pensée et de raisonnement différents.

Le test de rapport de vraisemblance pour le génotype de l'enfant proposé par Weinberg et al. (1998) est asymptotiquement équivalent au test du score développé par Schaid et Sommers (1993). Sur de petits échantillons, les résultats sont proches (Weinberg et al. 1998). Les estimations des coefficients et de leurs écarts-types dans les trois types d'analyse sont identiques (Self et al. 1991, Schaid et Sommers 1993, Weinberg et al. 1998). Même en présence de deux degrés de liberté pour le test dans l'approche génotypique, les tests du score et du rapport de vraisemblance offrent une puissance statistique équivalente ou supérieure à celle du TDT et tests assimilés dans la majorité des situations. Les simulations de Weinberg et al. (1998) montrent une performance du TDT statistiquement optimale et supérieure à celle des tests basés sur la vraisemblance uniquement dans le cas d'effets multiplicatifs des allèles. Pour d'autres modes de transmission (allèle dominant ou récessif), les simulations rapportent une puissance toujours inférieure voir faible du TDT. Certains auteurs considèrent le TDT comme un cas particulier de l'approche génotypique.

Les simulations de Weinberg et al. (1998) ont rapporté que 100 triades offrent une puissance de 80% de détecter un risque relatif d'un génotype de 2.5, avec un biais négligeable dans les estimations.

L'estimation du rôle du génotype de l'enfant se base sur l'évaluation d'une distorsion de la transmission observée par rapport aux lois mendéliennes. Dans le cadre du modèle log-linéaire, cette déviation est évaluée, pour chaque strate de profil parental, en comparant l'effectif observé des triades à l'effectif attendu selon ces lois. Par exemple : pour le profil de type 2, où l'un des parents est homozygote pour l'allèle d'intérêt  $AA$  et l'autre est hétérozygote  $aA$ , les lois mendéliennes définissent autant de familles avec un enfant  $aA$  que de familles avec un enfant  $AA$ . Un nombre observé de familles avec un enfant  $AA$  supérieur au nombre observé de familles avec un enfant  $aA$ , indique que le génotype  $AA$  pourrait être à risque de maladie.

L'héritage de type mendélien peut cependant être perturbé par d'autres facteurs qu'il est impératif d'apprécier afin de s'assurer de la validité des modèles. Ces facteurs comme la fertilité des parents et la viabilité du fœtus doivent être indépendants du génotype. Par exemple, si le génotype  $AA$  est associé à une mort fœtale, il est possible d'observer un effet apparent protecteur du génotype  $AA$  pour la maladie étudiée lié au fait que nous comptabilisons moins d'enfants-cas porteur du génotype  $AA$ . La distorsion mendélienne alors observée serait expliquée uniquement par la relation entre le génotype  $AA$  et une mort fœtale.

L'atout supplémentaire du modèle log-linéaire est d'estimer de façon indépendante les effets du génotype de l'enfant et ceux du génotype de la mère.

L'estimation des effets du génotype de la mère se base sur une asymétrie des fréquences parmi les différentes combinaisons des génotypes des couples parentaux à l'intérieur de chaque strate de profil parental. La validité des estimations de ces effets nécessite de poser l'hypothèse supplémentaire avançant que, dans la population en générale, les couples parentaux se constituent de façon indépendante de leur génotype. Plus précisément, nous supposons à l'intérieur de chaque profil parental une symétrie du nombre d'allèles pour les mères et les pères. Par exemple, pour le type 2, nous devons observer, dans la population générale, que la probabilité que la mère soit  $AA$  et le père  $aA$  est égale à la probabilité que le père soit  $AA$  et la mère  $aA$ .

Même si elle est difficilement vérifiable, cette hypothèse supplémentaire mérite de porter une attention particulière dans l'interprétation d'effets maternels apparents. Par exemple, si la survie d'un parent est influencée par le génotype et le sexe, cette situation pourrait provoquer une asymétrie artificielle des génotypes parentaux à l'intérieur des strates de profil parental.

En présence du conditionnement par les profils parentaux, le nombre de copies héritées (portées par l'enfant) et le nombre de copies portées par la mère sont statistiquement indépendants sous l'hypothèse d'un héritage mendélien. Ainsi, les tests et estimations proposés par le modèle log-linéaire sont orthogonaux et permettent au modèle de distinguer les effets du génotype de l'enfant des effets du génotype de la mère. Finalement, une analyse stratifiée sur les profils génétiques des couples parentaux, qui ne considère pas les effets potentiels du génotype de la mère, ne produit pas d'estimations biaisées des effets du génotype de l'enfant.

Enfin, ces modèles ne nécessitent pas de vérifier l'équilibre d'Hardy-Weinberg (HWE). Le conditionnement sur les génotypes parentaux permet d'ajuster et de contrôler une structure potentielle dans l'échantillon d'étude. Dans une situation d'une population bien mélangée et homogène, c'est-à-dire en supposant l'existence de l'HWE, l'écriture du modèle log-linéaire peut être simplifiée et inclut alors deux paramètres au lieu des six paramètres de strates. Il offre alors des tests plus puissants (Weinberg et al. 1998).

#### ii.d. EXTENSIONS LORSQUE LES GENOTYPES DES PARENTS SONT MANQUANTS

La principale limite des méthodes « cas-parents » par rapport à l'approche « cas-témoins » se situe dans la difficulté de recueil et la disponibilité des informations des deux génotypes parentaux simultanément.

En pratique, les parents peuvent être disparus, ils peuvent refuser de participer, ou le père peut être exclu *a posteriori* après avoir été identifié incorrectement comme le père biologique de l'enfant. Ces problèmes pratiques résultent en des triades incomplètes : des diades lorsque qu'un génotype parental est manquant, ou des « monades » (traduit littéralement de *monads*) lorsque les deux génotypes parentaux sont indisponibles. Dans certaines situations, une diade apporte une information sur la transmission. Par exemple, si l'enfant est homozygote  $AA$  et le parent disponible hétérozygote  $aA$ , nous pouvons conclure que le parent manquant a transmis un allèle  $A$  et est donc  $aA$  ou  $AA$ . Exclure ces diades conduit alors à une perte substantielle d'informations et de puissance statistique. D'autres diades ne fournissent que des informations ambiguës. Par exemple, lorsque l'enfant et le parent disponible sont hétérozygotes, nous ne pouvons pas distinguer la source parentale de l'allèle  $A$  hérité.

De nombreux travaux ont été développés pour prendre en compte les données manquantes de génotype parental. Quelques auteurs proposent de reconstruire les génotypes parentaux manquants à partir des diades « informatives » (non ambiguës). Mais, cette idée est controversée et Sham et Curtis (1995) estiment qu'une analyse qui exclut les diades ambiguës mais inclut les diades informatives, est incorrecte. Knapp (1999) propose un test corrigé de l'inflation artificielle apportée par la reconstruction des génotypes, le RC-TDT (*Reconstruction-Combined -TDT*). D'autres travaux utilisent, à travers des méthodes non paramétriques, des informations mixtes des parents et d'autres membres de la famille génotypés et non atteints pour compenser l'absence d'information chez certains parents. Il s'agit des méthodes FBAT avec les travaux de Rabinowitz et Laird (2000) et Rabinowitz (2001). Enfin, certains auteurs énumèrent l'ensemble des génotypes possibles pour les parents manquants, avec leur probabilité de se produire calculée à partir des informations observées des familles, puis les paramètres et les tests sont estimés sous l'hypothèse d'une population homogène et de **panmixie**. Certains travaux choisissent des procédures itératives de traitement de données manquantes affaiblissant la nécessité d'une hypothèse d'homogénéité de la population (Weinberg 1999a, Van den Oord 2000). Ces travaux utilisent l'algorithme EM, « Expectation-Maximisation », pour lequel la convergence est assurée par la théorie statistique entre 20 et 50 itérations (Dempster 1977).

Dans le cadre du modèle log-linéaire, la procédure itérative de l'algorithme EM consiste en :

- un fractionnement et une répartition de l'effectif des triades incomplètes selon les probabilités théoriques des profils parentaux sur la base (des valeurs initiales des paramètres du modèle pour la 1<sup>ère</sup> itération) des estimations courantes des paramètres du modèle (étape E)

- une répétition de la procédure de maximisation de la vraisemblance du modèle sur la base du nouveau jeu de triades intégrant les données pseudo-complètes (étape M).

Par exemple, dans le cas de deux allèles  $a$  et  $A$ , six profils parentaux doivent être considérés ( $AA \times AA$ ,  $aa \times aa$ ), soit  $p_i$  la probabilité du  $i^{\text{ème}}$  profil parental. Si la mère et l'enfant ont tous les deux le génotype  $AA$ , le père manquant peut présenter le génotype  $aA$  ou  $AA$ , impliquant possiblement deux mating-types :  $AA \times AA$ , ou  $AA \times aA$ . Le nombre observé de diades  $n_d$  avec la mère  $AA$  et l'enfant  $AA$  est fractionné en deux effectifs théoriques, selon les probabilités associées aux deux possibilités de profil parental :  $n_d * [p_1/(p_1+p_2)]$  et  $n_d * [p_2/(p_1+p_2)]$ . Le modèle est ensuite réalisé sur le jeu de données des triades complétées ou pseudo-complètes intégrant les effectifs attendus des diades.

Weinberg (1999a) montre que le test du rapport de vraisemblance combiné à l'algorithme EM pour le génotype de l'enfant, offre une puissance statistique supérieure au test réalisé sur les triades complètes seules. Le gain de puissance du test se maintient même dans des cas extrêmes de données manquantes (50% des génotypes parentaux sont manquants). Le gain de puissance du test combiné est plus modéré pour les effets du génotype maternel. Weinberg (1999a) ajoute la possibilité de réaliser ces estimations et tests en présence de monades. Cependant, nous devons reconnaître que, sans les génotypes parentaux, davantage d'hypothèses doivent être vérifiées dans la population d'étude pour obtenir des estimations valides.

De façon générale, la validité de l'ensemble des modèles proposant un traitement de données manquantes repose sur l'hypothèse qu'il n'existe pas d'association entre l'absence ou l'indisponibilité de l'information du génotype parental et le génotype lui-même.

## **C. Evaluation d'une interaction entre un gène et un facteur environnemental**

L'ensemble des méthodologies employées pour l'estimation d'un effet conjoint entre un facteur génétique et un facteur environnemental est basé sur des extensions simples des modèles s'intéressant aux associations génétiques, quelle que soit l'approche cas-témoins ou cas-parents. Un terme d'interaction est inclus dans les régressions logistiques, ou une stratification supplémentaire sur le facteur environnemental du jeu de données est mise en place pour la modélisation log-linéaire.

### **C. 1. Approche cas-témoins**

#### **C.1.i. Modélisation logistique**

Soit la situation simplifiée de deux facteurs dichotomiques génétique  $G$  et environnemental  $E$ , nous notons  $g+$  ( $g-$ ) la présence (l'absence) d'un caractère génétique du sujet (par exemple, présence de l'allèle d'intérêt  $A$  dans le génotype du sujet, qui correspond à un mode de transmission dominante de l'allèle), et  $e+$  ( $e-$ ) la présence (l'absence) d'exposition au facteur environnemental.

Soit le groupe de référence, le groupe de sujets possédant les caractéristiques g- et e-, le modèle logistique s'intéressant à l'interaction entre G et E s'écrit :

$$\text{logit}(p) = \alpha + \beta_g I_{g+} + \beta_e I_{e+} + \beta_{ge} I_{(g+ \text{ et } e+)}$$

où p est la probabilité d'être un cas

$\beta_g$  est le log de l'odds-ratio associé au caractère génétique g+ parmi les non-exposés e- ( $OR_g$ )

$\beta_e$  est le log de l'odds-ratio associé à l'exposition e+ parmi les non porteurs du caractère génétique g- ( $OR_e$ )

$\beta_{ge}$  est le paramètre additionnel lié à la présence conjointe de l'exposition et de la présence du caractère génétique. Son exponentielle est l'odds-ratio de l'interaction gène-environnement. L'exponentielle de  $(\beta_g + \beta_e + \beta_{ge})$  est l'odds-ratio associé à la présence du caractère génétique parmi les exposés au facteur environnemental ( $OR_{ge}$ ).

Les intervalles de confiance à 95% se calculent selon la procédure usuelle de Wald. L'interaction peut être testée par un test du rapport de vraisemblance (à 1 degré de liberté dans cette situation simplifiée).

Ce modèle peut être étendu en considérant les effets de chaque génotype. L'hypothèse nulle du test de l'interaction génotype-environnement est définie alors par la nullité simultanée de deux paramètres (correspondant aux trois génotypes moins un génotype de référence), aboutissant à un test à deux degrés de liberté. Un facteur environnemental à plus de deux classes ou en continu, peut être utilisé ; le nombre de degrés de liberté du test d'interaction doit alors être adapté.

L'interaction gène-environnement estimée et testée par la régression logistique est mesurée, dans ce cas, sous l'échelle multiplicative. L'hypothèse nulle du test est :  $H_0 : \beta_{ge}=0$ , nous obtenons donc sous  $H_0$  :

$$\begin{aligned} OR_{ge} &= \exp(\beta_g + \beta_e + \beta_{ge}) = \exp(\beta_g + \beta_e) \\ &= \exp(\beta_g) \exp(\beta_e) \\ &= OR_g OR_e \end{aligned}$$

L'hypothèse nulle peut être reformulée par :  $H_0$  : Effet conjoint est dit « multiplicatif » ou Absence d'interaction multiplicative ou  $OR_{ge} = OR_g OR_e$ .

Certains auteurs (Khoury et Flanders 1996, Botto et Khoury 2001) suggèrent de présenter par un tableau 2x4 les interactions gène-environnement mesurées par l'approche cas-témoins. Ce tableau a l'avantage de présenter de façon synthétique les effets utilisés, les effets propres et conjoints des facteurs E et G (1<sup>ère</sup> partie du Tableau 15 et de visualiser ainsi simplement l'existence d'une interaction selon les deux échelles de mesure : multiplicative et additive (2<sup>ème</sup> partie du Tableau 15).

**Tableau 15: Tableau 2x4 des effets propres et conjoint des facteurs génétique G et environnemental E, par l'approche cas-témoins**

E	G	Cas	Témoins	Odds-ratio (brut)	Information
-	-	a	b	Ref.	Référence commune
-	+	c	d	$OR_g = bc / ad$	Effet de g+ seul
+	-	e	f	$OR_e = eb / af$	Effet de e+ seul
+	+	g	h	$OR_{ge} = gb / ah$	Effet conjoint de e+ et g+
				<b><math>OR_{ge}</math> attendu</b>	<b>Déviations à l'<math>OR_{ge}</math> attendu (= <math>OR_{interaction}</math>)</b>
Sous l'hypothèse additive				$OR_g + OR_e - 1$	$OR_{ge} - OR_g - OR_e + 1$
Sous l'hypothèse multiplicative				$OR_g \times OR_e$	$OR_{ge} / (OR_g \times OR_e)$

-/+ : absence/présence

Ce mode de présentation peut être extrapolé en un tableau 2x6 lorsque nous nous intéressons aux effets de chaque génotype sans hypothèse particulière de mode de transmission génétique.

### **C.1.ii. Hypothèses, Limite, Puissance, Faisabilité**

L'estimation d'une interaction gène-environnement par l'approche cas-témoins, est sensible aux biais méthodologiques classiques de sélection des cas et des témoins, de classification de l'exposition et du génotype, de confusion et de stratification de population.

Quelques auteurs ont évalué les conséquences de la présence de certains de ces biais sur l'estimation d'une interaction, sa précision et sa puissance (Khoury et al. 1993, Witte et al. 1999, Garcia-Closas et al. 1999, Wacholder et al. 2002b). Garcia-Closas et al. (1999) montrent qu'une erreur de classement (différentielle ou non) générée sur un facteur environnemental dichotomique ou un facteur génétique, produit un biais de la mesure de l'interaction (multiplicative ou additive) qui converge vers une absence d'interaction entre le facteur environnemental et le gène. Wacholder et al. (2002b) se sont intéressés au biais de sélection des témoins hospitaliers sur la mesure d'une interaction gène-environnement. Si le groupe de témoins hospitaliers est recruté à travers une pathologie unique, aucun biais pour l'interaction multiplicative n'est observé si la pathologie des témoins n'est pas liée à cette interaction (quel que soit le lien des deux facteurs propres avec le risque de cette pathologie). En revanche, si le groupe de témoins est constitué de plusieurs pathologies, même si elles ne sont pas liées à l'interaction d'intérêt, ce groupe n'assure pas d'absence de biais dans la mesure de l'interaction.

Il est maintenant bien établi que la puissance de détection de l'effet d'une interaction gène-environnement est très inférieure à celles des effets propres du facteur environnemental et du facteur génétique (Greenland 1983, Smith and Day 1984). De plus, la variance du paramètre estimé de l'interaction est supérieure aux variances des paramètres des effets propres.

Différents auteurs ont estimé la puissance de détection d'une interaction par l'approche cas-témoins et la taille de l'échantillon associée (Hwang et al. 1994, Khoury et al. 1995, Andrieu et Goldstein 1996, Foppa et Spiegelman 1997, Yang et al. 1997, Yang et al. 2003). Les

approches diffèrent dans les définitions de l'hypothèse alternative du test ou plus généralement par leurs propriétés mathématiques, conduisant pour quelques-unes à une sur-estimation de la puissance et à une sous-estimation de la taille de l'échantillon pour des situations très spécifiques (García-Closas et Lubin 1999). L'approche la plus courante se base sur les calculs de taille d'échantillon proposés par Smith et Day (1984) pour détecter une interaction, et convient aux situations pour lesquelles l'effet de l'interaction et les effets propres sont modérés. Les calculs s'appuient sur des informations contenues dans un tableau 2x4 (Tableau 15),  $OR_{ge}$ ,  $OR_e$ ,  $OR_g$ , et sur la prévalence de l'exposition dans la population ( $p_e$ ), sur la prévalence du génotype dans la population ( $p_g$ ), sur le ratio cas-témoins, ainsi que sur l'erreur de type I et la puissance souhaitée. Hwang et al. (1994) montrent que la situation décrite par  $p_g=0.30$ ,  $p_e=0.70$ ,  $OR_e \approx OR_g \approx 2$ , ratio de 2 témoins pour 1 cas, nécessite 200 et 400 témoins pour détecter une interaction multiplicative de 4 (c'est-à-dire une  $OR_{ge} \approx 16$ ) avec une erreur de type I de 5% et une puissance de 80%. Cependant, dans la majorité des situations réelles, les génotypes sont plus rares et l'exposition est moins fréquente. Andrieu et Goldstein (1996) envisage alors la situation suivante,  $p_g=0.01$ ,  $p_e=0.30$ ,  $OR_e \approx OR_g \approx 2$ , ratio cas-témoins de 1:1. Un total de 12000 cas et 12000 témoins est nécessaire pour détecter une interaction multiplicative de 3 (c'est-à-dire une  $OR_{ge} \approx 12$ ) avec une erreur de type I de 5% et une puissance de 80%.

Enfin, la faisabilité d'une étude cas-témoins dans le cas d'un polymorphisme rare (<1%) et d'une exposition peu commune est compromise. Hwang et al. (1994) estiment, de façon synthétique, un effectif minimal de 1000 cas et 1000 témoins pour détecter une interaction dans une situation d'un polymorphisme presque rare (entre 1 et 5 %) et d'odds-ratio de G et E modérés (autour de 2). En revanche, cette approche est appropriée aux études de détection d'interaction si les facteurs E et G sont communs.

### C.1.iii. L'alternative Cas-seuls

#### iii.a. PRINCIPE

Dans certaines situations, les effets propres du facteur environnemental E et du gène G sur le risque de maladie sont négligeables et l'intérêt est porté sur un possible effet de synergie de la présence conjointe des deux facteurs. L'estimation d'une interaction ne requiert pas dans toutes les situations un groupe de témoins, et le schéma d'étude « cas-seuls » a ainsi été proposé dans la littérature épidémiologique (Piegorsch et al. 1994, Khoury et Flanders 1996, Andrieu et Goldstein 1998, Weinberg et Umbach 2000, Umbach et Weinberg 1997, Yang et al. 1997, Botto et Khoury 2001).

Cette méthode d'analyse s'appuie sur la relation suivante : la déviation de l'OR de la présence conjointe du gène et du facteur environnemental à l'hypothèse multiplicative des OR propres de E et G, mesurée par l'approche cas-témoins, est égale à l'OR mesurant l'association entre le gène et le facteur environnemental chez les cas divisé par ce même OR mesuré chez les témoins.

$$\frac{OR_{ge}}{OR_e OR_g} = \frac{OR_{cas}}{OR_{témoins}}$$

où

$$OR_{cas} = ag/ce$$

$$OR_{\text{témoins}} = bh/df$$

(notations du Tableau 15)

Si le génotype et le facteur environnemental sont indépendants dans la population générale (ou dans le groupe des témoins), la valeur de l' $OR_{\text{témoins}}$  s'approche de 1 et l'odds-ratio mesuré chez les cas-seuls mesure la déviation à un effet conjoint multiplicatif du gène et de l'exposition environnementale, c'est l'odds-ratio de l'interaction mesurée sous l'hypothèse de l'échelle multiplicative.

**Tableau 16: Effet conjoint des facteurs génétique G et environnemental E, par l'approche cas-seuls**

		G	
		-	+
E	-	a	C
	+	e	G

(effectifs issus du Tableau 15)

Cette analyse peut se présenter sous la forme d'un tableau 2x2 avec le calcul de l'effet de l'interaction par un OR brut et testé par une statistique du Khi-deux (Tableau 16). Elle peut également être menée par une analyse logistique classique avec ajustement.

### iii.b. HYPOTHESES, LIMITE, PUISSANCE

Si le génotype n'est pas corrélé à l'exposition dans la population et que la maladie est rare, l'alternative « cas-seuls » offre une meilleure précision de l'estimation de l'interaction gène-environnement et une puissance supérieure du test que ne le fournit l'approche cas-témoins traditionnelle. Ces propriétés s'expliquent mathématiquement.

Pour l'approche cas-témoins, la précision de l'estimation du logarithme de l' $OR_{ge}$  s'écrit, à partir des effectifs du tableau 2x4 (Tableau 15), tel que :

$$\text{var}(\hat{\beta}_{ge}) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} + \frac{1}{e} + \frac{1}{f} + \frac{1}{g} + \frac{1}{h} \quad [\text{cas-témoins}]$$

L'approche cas-seuls estime cette variance à partir des effectifs des cas seulement par :

$$\text{var}(\hat{\beta}_{ge}) = \frac{1}{a} + \frac{1}{c} + \frac{1}{e} + \frac{1}{g} \quad [\text{cas-seuls}]$$

Dans une situation asymptotique (si b, d, f et h, les effectifs du groupe des témoins, sont très grands), les variances estimées selon les deux approches sont alors équivalentes. Sinon, la variance estimée par l'approche cas-seuls est inférieure à celle estimée avec les effectifs des témoins et procure ainsi une meilleure précision de l'estimation. De façon similaire, la puissance de détection d'une interaction par l'approche cas-seuls est meilleure que celle de l'approche cas-témoins, et accroît avec l' $OR_{ge}$ , son calcul faisant intervenir des variances du logarithme de l' $OR_{ge}$  plus précises sous les hypothèses nulle et alternative.

Finalement, le gain de performance de l'approche cas-seuls s'explique par le fait que l'on pose une hypothèse d'indépendance entre le facteur génétique et le facteur environnemental.

Cette hypothèse semble raisonnable dans la population générale, cependant il existe des gènes qui conduisent à modifier l'exposition environnementale (par exemple, un comportement d'aversion à l'alcool). Aussi, il peut exister un tiers facteur pour lequel les facteurs génétique et environnemental varient différemment selon ses valeurs, créant ainsi une corrélation entre les deux facteurs (par exemple, le caractère ethnique).

S'appuyant sur l'erreur de 1<sup>ère</sup> espèce à 5% d'un test, Albert et al. (2001) évaluent la sensibilité du test de l'interaction gène-environnement à une déviation de l'hypothèse d'indépendance entre le gène et le facteur environnemental. A partir de 1000 simulations sur un jeu de données de 340 cas et 340 témoins, ils observent 4.1% de rejet du test si l'OR<sub>ge</sub> chez les témoins (noté OR<sub>témoins</sub> plus haut) est égal à 1 (correspondant à l'hypothèse d'indépendance entre G et E), 10% de rejet du test si l'OR<sub>ge</sub> chez les témoins est égal à 1.22, et 60% de rejet pour un OR<sub>ge</sub> chez les témoins de 1.65. Ces résultats indiquent une hypersensibilité du test de l'interaction face à l'hypothèse d'indépendance entre G et E du groupe des témoins.

Finalement, cette alternative n'est pas satisfaisante pour substituer l'approche cas-témoins : elle offre un test d'interaction gène-environnement très sensible à une légère déviation de l'hypothèse d'indépendance entre G et E (elle est encore moins appropriée pour l'étude d'une interaction entre deux facteurs environnementaux, comme des facteurs nutritionnels, la consommation de tabac ou l'exposition professionnelle, l'hypothèse d'indépendance étant encore plus difficile à vérifier). De plus, elle n'estime pas les effets propres des deux facteurs.

#### **C.1.iv. L'alternative mixte : modélisation log-linéaire**

Umbach et Weinberg (1997) ont proposé une méthode mixte permettant d'évaluer à la fois les effets propres du facteur génétique, du facteur environnemental et de l'interaction, en forçant l'hypothèse d'« indépendance » entre le génotype et l'exposition environnementale chez les témoins. L'approche propose trois modèles log-linéaires construits selon les informations disponibles et l'hypothèse d'« indépendance » entre le génotype et l'exposition.

Soit une situation simple, où G note la présence (G=1) ou l'absence (G=0) du caractère génétique d'intérêt, E la présence (E=1) ou l'absence (E=0) de l'exposition, et M la présence (M=1, sujet cas) ou l'absence (M=0, sujet témoin) de la maladie. Nous notons  $u_{mge}$  le nombre de sujets avec M=m, G=g et E=e.

Si l'ensemble des informations sur les témoins est disponible, le modèle complet s'écrit :

$$\ln u_{mge} = \mu_0 + \alpha_0 E + \beta_0 G + \gamma_0 GE + \mu_1 M + \alpha_1 ME + \beta_1 MG + \gamma_1 MGE$$

Ce modèle ne repose sur aucune hypothèse d'« indépendance » entre génotype et exposition. L'approche log-linéaire présentée est équivalente à l'approche logistique classiquement utilisée pour l'analyse des études cas-témoins. Ces modèles fournissent les mêmes descriptions du risque la maladie : les exponentielles des coefficients  $\alpha_1$ ,  $\beta_1$  et  $\gamma_1$  sont les OR lié à, respectivement, l'exposition, le génotype d'intérêt et l'interaction gène-environnement.

Sous l'hypothèse d'indépendance entre le génotype et l'exposition parmi les témoins, le modèle s'écrit tel que l'on pose la contrainte  $\gamma_0 = 0$ :

$$\ln u_{mge} = \mu_0 + \alpha_0 E + \beta_0 G + \mu_1 M + \alpha_1 ME + \beta_1 MG + \gamma_1 MGE$$

Ce modèle offre les mêmes interprétations des coefficients que le modèle précédent, en revanche les valeurs des estimations sont modifiées. En l'absence du terme d'interaction « GE » dans le modèle, l'estimation du coefficient  $\gamma_1$  ne requiert pas les informations du groupe des témoins, il s'agit de l'estimateur « cas-seuls » proposés par Piegorsch et al. (1994). Les paramètres  $\alpha_1$  et  $\beta_1$  dépendent des données de témoins seulement par les sommes marginales. Les variances des coefficients ( $\alpha_1$ ,  $\beta_1$ ,  $\gamma_1$ ) sont plus faibles que celles du modèle complet. Lorsque l'hypothèse d'indépendance est remplie, cette seconde analyse nécessite moins de sujets que l'analyse faite sous le modèle complet. Cependant, l'estimation « cas-seuls » reste sensible à la déviation de l'hypothèse.

Enfin, l'idée initiale des auteurs était le développement d'une analyse qui estimerait les effets propres et l'interaction en prenant en compte la difficulté pratique d'obtention d'information génétique du groupe de témoins. Ainsi, une extension du modèle permet la prise en compte d'informations incomplètes lorsque l'information du génotype chez les témoins n'est pas obtenue. Le modèle s'écrit :

$$\ln u_{mge} = \mu_0 + a_0 E + m_1 M + a_1 ME + b_1 MG + c_1 MGE$$

Les notations de ce troisième modèle sont différentes car  $b_1$  n'estime plus l'effet du génotype correspondant au terme  $\beta_1$ , mais la somme  $\beta_0 + \beta_1$ . Lorsque les témoins ne sont pas génotypés, l'effet propre du génotype G sur le risque de maladie n'est plus estimable, alors que l'interaction gène-environnement le reste sous l'hypothèse d'indépendance des deux facteurs. Le risque de maladie lié à l'interaction gène-environnement est l'exponentielle de  $c_1$ , celui de l'exposition est l'exponentielle du coefficient  $a_1$ .

### **C.1.v. Autres alternatives**

Des stratégies particulières d'appariement des cas et des témoins ont été proposées pour détecter une interaction gène-environnement pour les situations où l'un ou les deux facteurs sont rares. Il s'agit des stratégies de contre appariement (*countermatched design*) ou d'appariement en deux phases décrites par Andrieu et Goldstein (1998) et des stratégies d'appariement flexibles proposées par Stürmer et Brenner (2000). Le principe du contre-appariement est de sélectionner des témoins pour accroître la variation du facteur d'intérêt dans le jeu de données, en maximisant ainsi le nombre de paires discordantes cas-témoins à l'opposé du principe d'appariement classique. L'appariement flexible estime la fréquence de témoins exposés à sélectionner, qui maximise la puissance du test d'interaction.

Ces stratégies sont récentes et la puissance, la performance et la faisabilité de ces études sont encore peu évaluées. La stratégie de contre-appariement sur le(s) facteur(s) rare(s) des cas et des témoins produirait une puissance supérieure de détection d'une interaction gène-environnement par rapport aux analyses traditionnelles cas-témoins. Mais elle nécessiterait encore une taille d'échantillon importante, rendant la faisabilité de ces études toujours difficile. De surcroît, la mise en place de cette stratégie est complexe (Andrieu et al. 2001).

## **C. 2. Approche cas-parents**

La majorité des méthodes d'analyse présentées dans la section précédente permet d'étudier une hétérogénéité de transmission allélique ou de stratifier les modèles pour évaluer des effets des génotypes différents selon l'exposition au facteur environnemental (Khoury et Flanders

1996, Maestri et al. 1997, Waldman et al. 1999, Eaves et Sullivan 2001, Witte et al. 1999, Schaid 1999b, Umbach et Weinberg 2000, Laird 2000, Lake et Laird 2004).

Umbach et Weinberg (2000) soutiennent que les tests d'interaction gène-environnement qui font intervenir un contraste entre les taux de transmission allélique des familles exposées et des familles non exposées, à partir des parents hétérozygotes à leur enfant atteint, sont sensibles à deux sources de biais : la différence entre les taux de transmission des familles exposées et non exposées peut être la conséquence d'une stratification de population même s'il n'y a pas d'interaction ; dans la majorité des situations, il n'existe pas d'indépendance des transmissions alléliques entre les parents hétérozygotes sous l'hypothèse nulle d'absence d'interaction. De tels biais provoquent une inflation du taux d'erreur de première espèce du test de l'interaction. Ils se produisent pour tous les tests basés sur le comptage de transmission allélique à partir des parents hétérozygotes.

Robustes à ces sources de biais, l'approche de Schaid et Sommers (1993), conditionnée sur les génotypes parentaux, s'étend à l'étude d'interaction gène-environnement et est présentée par Schaid (1999b) puis complétée par Cordell et al. (2004). Les méthodes non paramétriques FBAT possèdent l'avantage de permettre l'évaluation d'une interaction entre un génotype et une exposition environnementale quantitative (Laird et al. 2000, Lake et Laird 2004). En raison de sa flexibilité d'implémentation, nous choisissons de développer plus en détails dans la section suivante le modèle log-linéaire de l'approche cas-parents s'intéressant à l'interaction gène-environnement (Umbach et Weinberg 2000).

### **C.2.i. Méthodologie log-linéaire**

Soit un gène G à deux allèles et un facteur environnemental E dichotomique, l'approche log-linéaire peut être étendue à l'étude des interactions gène-environnement, par la construction d'une stratification des familles selon leur exposition à E. Nous nous intéressons aux effets du génotype de l'enfant sur chaque strate d'exposition. Le modèle s'écrit :

$$\ln(E[n_{M,P,C}]) = \ln(2) I_{[M=P=C=1]} + \gamma_j + \delta_j I_{[E=1]} + \beta_1 I_{[C=1]} + \beta_2 I_{[C=2]} + \eta_1 I_{[E=1;C=1]} + \eta_2 I_{[E=1;C=2]}$$

Dans ce modèle,  $\gamma_j$  correspond au risque de base de chaque profil parental j du groupe « non exposé », et  $(\gamma_j + \delta_j)$  à celui des familles « exposées ». Les coefficients  $\beta_c$  ( $c=1,2$ ) sont les logarithmes des risques relatifs associés au nombre c de copies de l'allèle d'intérêt chez l'enfant parmi les triades « non exposées » (par rapport à aucune copie). Pour le groupe des « exposés », le log des risques relatifs associés à  $C=c$ , versus 0 copie, correspond à la somme  $(\beta_c + \eta_c)$ . Enfin, les effets potentiels de l'interaction gène-environnement sont multiplicatifs et sont mesurés par les paramètres  $\eta_1$  et  $\eta_2$ .

Les interactions génotype-environnement sont testées par un test du rapport de vraisemblance ( $H_0 : \eta_1 = \eta_2 = 0$ ) qui suit une loi du Khi-deux à 2 degrés de liberté. Ce modèle peut intégrer les interactions entre le génotype de la mère et le facteur environnemental. Il a été généralisé par Kistner et Weinberg (2004) pour les situations d'un facteur environnemental quantitatif et de triades incomplètes. Cette extension est semi-paramétrique et a l'avantage de ne pas imposer d'hypothèse de normalité du facteur quantitatif.

### **C.2.ii. Hypothèses, Limites et Puissance**

L'étude d'interaction gène-environnement par le modèle log-linéaire sur les triades, nécessite de poser l'hypothèse mendélienne séparément sur chaque niveau de l'exposition, car l'inférence de l'interaction se base sur le degré de la déviation aux proportions attendues entre chaque strate d'exposition. Ainsi, cette analyse est essentiellement applicable aux maladies précoces (*early onset*), tout d'abord pour des raisons inhérentes au schéma d'étude nécessitant la présence des deux parents et la disponibilité des génotypes, puis pour écarter l'influence de facteurs pouvant perturber l'hypothèse mendélienne au cours du temps (exemple, la survie de sujets liée au génotype).

Une seconde hypothèse est nécessaire à la validité de ce modèle : l'exposition d'un sujet doit être indépendante de son génotype conditionnellement aux génotypes des parents. Cette hypothèse est moins forte que l'hypothèse appliquée à la population générale requise pour le schéma d'étude cas-seuls (Witte et al. 1999). Quand l'allèle d'intérêt prédispose à une exposition et crée ainsi une structure dans la population, il en résulte une sur-représentation de sujets-cas porteurs de l'allèle dans chaque famille exposée. Les inférences étant basées sur le nombre de familles, une apparente distorsion mendélienne apparaît, et produit un contraste de distorsion mendélienne entre les familles exposées et non exposées. Par exemple, la source de dépendance entre l'exposition et le génotype pourrait être l'histoire familiale positive d'antécédents de la maladie. Cependant, il est important de distinguer quels membres de la famille sont impliqués : si l'exposition maternelle est influencée par la présence d'antécédents parmi ses frères et sœurs, parents, grands-parents ou cousins, et non par la présence de la maladie chez ses enfants, ceci n'inclut pas de biais dans l'évaluation d'une interaction gène-environnement par l'approche cas-parents (Thomas 2000).

L'approche cas-parents ne permet pas d'estimer l'effet propre du facteur environnemental. Elle examine l'interaction gène-environnement simplement par la comparaison des risques relatifs génétiques estimés parmi les familles exposées et les familles non exposées. Par conséquent, il n'est pas possible de distinguer si l'exposition a un effet neutre en l'absence du génotype d'intérêt et un effet délétère en présence de ce génotype, ou si l'exposition a un effet bénéfique en l'absence du génotype d'intérêt et un effet neutre en présence de ce génotype.

Le schéma d'étude cas-parents présente une puissance du test de l'effet génétique supérieure au schéma cas-témoins dans la majorité des situations (lorsqu'il est indemne de biais de stratification de population) (Witte et al. 1999). En revanche, cette efficacité relative est moins évidente lorsqu'il s'agit d'un test d'interaction gène-environnement.

A un nombre d'enfants-cas fixé, Schaid (1999b) montre que l'approche cas-parents offre une puissance comparable ou légèrement supérieure à l'approche cas-témoins (non appariée) de détecter la présence d'une interaction gène-environnement. Witte et al. (1999), puis Gauderman 2002), comparent l'efficacité d'évaluation d'une interaction entre les schémas d'étude appariés suivants : cas-témoins, cas-*sibs* (le *sib* est un frère ou une sœur), cas-cousins et cas-parents (ou pseudo-*sibs*). A partir de multiples possibilités de modèles, Gauderman (2002) montre que l'approche cas-*sibs* et cas-parents nécessitent un nombre de séries appariées plus faible que l'approche cas-témoins pour atteindre la même puissance. L'approche cas-*sibs* serait particulièrement plus puissante pour un mode de transmission dominant, tandis que le mode récessif convient mieux à l'approche cas-parents. Enfin, l'approche cas-*sibs* serait plus adaptée que l'approche cas-parents aux gènes les plus rares.

## D. Tableau-bilan de l'évaluation d'une interaction gène-environnement

**Tableau 17: Caractéristiques des approches disponibles dans l'étude évaluant une interaction gène-environnement (notée GxE, avec le gène G et le facteur environnemental E)**

	Cas-Témoins	Cas-Seuls	Cas-Parents
Quelle population-témoin ?	Témoins (Externes)	Pas de témoins	- Allèles non transmis des parents, ou - Pseudo-sibs (ou distribution attendue des génotypes selon les génotypes des parents)
Hypothèse associée à la population-témoin	Population des témoins « génétiquement équivalente » à celle de la population-source des cas		Transmission mendélienne
Hypothèse associée à la mesure de l'interaction GxE		Pas de lien entre G et E dans la population générale	- Indépendance entre G et E, conditionnellement aux génotypes parentaux - Hypothèse mendélienne sur chaque strate de E
<b>Sensibilité (impliquant un biais potentiel dans l'estimation d'une interaction GxE) des approches à ...</b>			
... une stratification génétique de la population	Oui	Oui	- Oui, si évaluation de GxE par transmission allélique - Non, si conditionnement sur « profils génétiques des parents »
... une distorsion de l'hypothèse de transmission mendélienne dans la population générale	Equivalent à un biais de sélection. Exemple : la survie des sujets associée à G	Non	Oui, s'il existe un contraste entre les strates du facteur environnemental
... un lien entre le gène G de l'enfant et son exposition E dans la population générale	Non	Oui	Non (Oui, si l'exposition est influencée par le gène de l'enfant : situation impossible dans cette étude, car il s'agit d'une exposition <i>in utero</i> )
<b>Difficulté d'interprétation d'une interaction GxE liée à ...</b>			
... un déséquilibre de liaison entre le locus marqueur et le locus de la maladie	Oui	Oui	Oui
... un déséquilibre de liaison entre le locus marqueur et un locus associé à E	Non	Oui	Non
... la présence d'un effet du génotype maternel	Oui	Oui	- Oui, si évaluation par transmission allélique - Non, si conditionnement sur « profils génétiques » des parents

## E. Mesure des effets d'empreintes parentales

Les effets du génotype maternel, présentés jusqu'à présent dans ce projet, correspondent à des influences génétiques sur l'environnement intra-utérin du fœtus, ou, plus généralement, à des influences génétiques non héritées par l'enfant. Une empreinte parentale est définie par l'influence génétique d'un allèle sur le risque de maladie de l'enfant différente selon l'origine maternelle ou paternelle de cet allèle. Par exemple, il a été observé que l'expression d'un allèle peut être différente selon s'il est hérité de la mère ou du père.

Les premiers tests qui ont proposé de mesurer l'effet d'empreinte parentale sont le TDT<sub>M vs F</sub> (le TDT *Mothers versus Fathers*) et le TAT (*Transmission Asymmetry Test*). Le premier consiste à stratifier l'étude de la transmission allélique selon le sexe du parent hétérozygote et à comparer les taux de transmissions alléliques par un test du Khi-deux standard. Le principe du TAT est équivalent ; il exclut cependant les familles pour lesquelles les deux parents sont hétérozygotes dans le but de vérifier les conditions d'indépendance entre les unités statistiques. Ces deux tests ne sont, toutefois, pas valides en présence d'effets du génotype maternel (Weinberg 1999b).

Une alternative à ces tests, adaptée à la possibilité d'une présence d'effets du génotype maternel, est le PO-LRT (*Parental-of-Origin Likelihood Ratio Test*). Ce test est proposé par Weinberg (1999b). Le calcul de ce test s'appuie sur les trois profils génétiques de couple parental pour lesquels la mère et le père n'ont pas le même génotype. Le modèle proposé exprime la probabilité que la mère porte un nombre de copies de l'allèle d'intérêt supérieur au père (M>P) en fonction de l'effet d'empreinte maternelle et des effets propres du génotype de la mère. Il s'écrit :

$$\text{logit } p(M>P/\text{profil parental}, C) = \alpha I_{(C=1)} + \beta I_{(M+P>1)} + \gamma [I_{(M+P=1)} - I_{(M+P>2)}]$$

où M, P et C sont les nombres de copies de l'allèle d'intérêt porté par, respectivement, la mère, le père et l'enfant, I définit une variable indicatrice.

La présence d'un effet d'empreinte parentale se mesure par la significativité du coefficient  $\alpha$  (testée par le test usuel du rapport de vraisemblance,  $H_0 : \alpha=0$ ). Le conditionnement sur les profils parentaux et le nombre de copies d'allèles portés par l'enfant permet d'évaluer l'effet de l'empreinte parentale de façon statistiquement indépendante des effets potentiels du génotype maternel. Cette robustesse du test lui procure néanmoins une puissance statistique modeste.

D'autres auteurs s'intéressent à un effet possible d'empreinte parentale, les tests proposés sont plus puissants mais ne permettent pas d'évaluer cet effet indépendant d'une influence génétique maternelle (Eaves and Sullivan 2001, Cordell et al. 2004).

# 4

## Objectif du projet

L'objectif général de ce projet est d'évaluer, à partir d'une enquête épidémiologique, le rôle d'interactions gène-environnement dans l'étiologie des fentes orales.

Les études d'interaction gène-environnement en épidémiologie permettent d'évaluer les contributions séparées et simultanées des gènes et de l'environnement sur la maladie. Ces études permettent de clarifier des mécanismes biologiques de la maladie chez l'homme, pertinents au gène et au facteur environnemental étudiés, identifiant ainsi des chemins et des facteurs causaux à la maladie. Elles peuvent aider à déterminer quel composant spécifique dans une situation d'exposition complexe et multiple à laquelle les hommes sont exposés, joue un rôle dans la survenue de la maladie. L'utilisation de ces hypothèses de mécanismes biologiques permet en conséquence de développer des stratégies thérapeutiques et préventives plus ciblées.

La seule identification de gènes de susceptibilité à une maladie fournit des preuves directes que ces gènes et leurs mécanismes biologiques associés sont pertinents à l'explication de la maladie. Cependant, l'identification des interactions gène-environnement permet, de plus, d'individualiser les stratégies de prévention (en modulant par exemple, avant le diagnostic de la maladie, l'exposition environnementale chez les sujets possédant le gène de susceptibilité) et de traitement (en offrant un traitement personnalisé de la maladie).

Une interaction gène-environnement sera acceptée si elle peut être reproduite dans plusieurs études épidémiologiques et si l'effet semble plausible au niveau biologique.

Ce projet permettra de contribuer à la connaissance du rôle d'interactions gène-environnement dans l'étiologie des fentes orales nonsyndromiques. Il s'appuie sur une étude cas-témoins menée en France entre 1998 et 2001, intégrant un schéma d'étude cas-parents.

Il cible plus spécifiquement certains facteurs environnementaux déjà soupçonnés dans la littérature épidémiologique des fentes orales.

Le choix des gènes d'intérêt a été fait *a priori* pour tenir compte d'une part des voies de métabolisme et de détoxification des expositions environnementales identifiées comme facteurs de risque (ou facteurs protecteurs), et d'autre part des gènes récemment soupçonnés d'avoir un rôle dans la formation embryonnaire des fentes orales chez l'homme ou l'animal.

Les effets génétiques maternels indirects (pouvant modifier l'environnement intra-utérin du fœtus) seront évalués et distingués des effets directs du génotype de l'enfant..



# **MATERIEL & METHODES**



# 1

## Population

Le projet s'appuie sur une étude épidémiologique cas-témoins mise en place à partir de services hospitaliers de chirurgie pédiatrique en France, entre 1998 et 2001.

### A. Enquête épidémiologique

#### A. 1. Recrutement des sujets

##### A.1.i. Les cas

Les cas sont des enfants atteints de fente labiopalatine (FL/P) ou de fente palatine isolée (FP), traités par chirurgie maxillo-faciale dans un des services suivants :

- A Paris : Hôpital Trousseau (Pr M-P Vazquez)
- A Lyon : Hôpital de la Croix-Rousse (Pr L-L Béziat), Hôpital Debrousse (Pr C Paulus), Hôpital Edouard Herriot (Pr J-P Chappuis), Clinique du Val d'Ouest (Dr I James-Pangaud)
- A Clermont-Ferrand : CHU (Pr J-M Mondié)
- A Grenoble : CHU (Pr B Raphael)

Entre septembre 1998 et février 2001, 263 mères d'enfants atteints de fentes orales (syndromiques ou non syndromiques) ont été interrogées à l'hôpital. Plus de 50% du recrutement proviennent des hôpitaux Lyonnais et 30% de l'Hôpital Trousseau à Paris (Tableau 20). Un total de 37 enfant-cas éligibles n'a pu être inclus dans l'étude. La cause majeure de non-inclusion est le problème de maîtrise de la langue française ; d'autres raisons peuvent également être évoquées : refus de participation, problèmes pratiques d'emplois du temps avec les mères, et absences d'interview pour cause de décès post-opératoire ou d'abandon de l'enfant.

Les cas se répartissent en 23 cas syndromiques (treize **séquences** Pierre Robin, trois Van der Woude, trois Di George, un Stickler, un Kabuki, un Aarskog et un syndrome orodigitofacial), 76 fentes palatines et 164 cas de fentes labio-palatines. Nous retrouvons les caractéristiques attendues pour cette population d'enfants atteints : deux fois plus de FL/P que de FP, et une prédominance de garçons parmi les FL/P (cf section suivante Description de la population et le Tableau 20).

##### A.1.ii. Les témoins

Le protocole prévoyait le recrutement d'enfants témoins admis dans les mêmes hôpitaux que les cas, pour des raisons autres que malformations, cancers ou maladies génétiques, et pour lesquelles les nécessités du service prévoyaient un prélèvement de sang. Les enfants devaient,

en outre, être appariés aux cas sur le sexe, l'âge ( $\pm 1$  mois), l'éloignement du lieu de résidence au bassin de recrutement de l'hôpital et l'origine géographique de la mère.

De nombreuses difficultés se sont présentées pour le recrutement de témoins hospitalisés et la collecte de leurs prélèvements biologiques. Dans la majorité des services de chirurgie pédiatrique initialement prévus par le protocole, peu d'enfants avec un « diagnostic bénin » sont admis dans les tranches d'âge requises par le protocole d'étude, ce qui a rendu difficile le recrutement d'enfants répondant à l'ensemble des critères attendus de témoin. Il a été nécessaire d'élargir le recrutement des enfants-témoins à d'autres hôpitaux dans l'étude, comme l'hôpital Debrousse à Lyon, et à des services hospitaliers non chirurgicaux. De plus, pour le cas spécifique de l'hôpital Trousseau de Paris spécialisé dans les maladies sévères de l'enfance, tous les enfants-témoins ont été recrutés dans les maternités locales de Tenon et Rothschild. Le choix de recruter des enfants-témoins dans des maternités nous a amené à relaxer le critère d'appariement sur l'âge. Enfin, malgré cet élargissement du recrutement, l'appariement des enfants-témoins avec les enfants-cas sur le lieu de résidence (dans le bassin de recrutement de l'hôpital ou dehors) n'a pas toujours été réalisé.

Nous avons été confrontés au problème d'absence des autorisations nécessaires dans les services pour des prélèvements sanguins sans bénéfice individuel direct pour le sujet, car l'autorisation donnée par le CCPPRB (Comité Consultatif de Protection des Personnes dans la Recherche Biomédicale) exige que le prélèvement ne soit effectué qu'en plus d'un autre déjà prévu pour les besoins du service. Avec l'accord supplémentaire de la CNIL (Commission Nationale de l'Informatique et des Libertés) obtenue en août 2000, nous avons alors pris la décision de demander aux mères d'enfants-témoins l'autorisation d'utiliser le prélèvement de tâches de sang fait à la naissance pour le test de Guthrie. Munis des consentements des mères, nous n'avons eu, dès lors, aucune difficulté à obtenir les buvards des tests de Guthrie, excepté pour Paris où la Fédération Parisienne pour le Dépistage a exigé un second consentement écrit des familles et a refusé l'accès aux buvards lorsque les familles n'avaient pas réaffirmé leur accord.

Finalement, les témoins éligibles à cette étude sont les enfants-témoins appariés avec les enfants-cas sur le centre, le sexe et l'origine géographique des mères (ainsi que l'âge et le lieu de résidence, selon les contraintes pratiques), présents à l'hôpital à la visite de l'enquêteur et contactés pour la collecte d'informations. Parmi les familles de témoins éligibles à cette étude, 25 familles ont refusé de participer. Toutes ces difficultés et les retards qu'elles ont entraînés, font que le recrutement des témoins s'est poursuivi au delà de la fin du recrutement des enfants-cas (interview de janvier 2000 à décembre 2001), et que le nombre de témoins recrutés est légèrement inférieur à celui des cas. Un total de 236 témoins a été inclus dans l'étude, dont 80 enfants sont issus des maternités parisiennes. Les causes d'hospitalisation des 156 témoins hospitaliers sont précisées en Annexe 2.

## **A. 2. Collecte d'informations**

### **A.2.i. Questionnaire à la mère**

L'interview des mères de cas et de témoins s'est tenue à l'hôpital au moyen d'un questionnaire standardisé. Les questions portaient sur les variables socio-démographiques (âge des parents, niveau d'études, pays d'origine, lieu de résidence), l'histoire médicale et obstétricale de la

mère, l'histoire familiale de fentes orales (pour les cas), les maladies infectieuses ou les prises de médicaments (incluant les suppléments vitaminiques), les expositions médicales aux rayons X et les chirurgies, l'alimentation habituelle, la consommation de tabac et d'alcool, l'exposition passive au tabac, les professions et les expositions professionnelles, les expositions domestiques à des produits chimiques (en particulier solvants). Pour l'ensemble des expositions, quatre périodes de temps ont été considérées : période préconceptionnelle (débutant 1 mois avant la grossesse), 1<sup>er</sup>, 2<sup>ème</sup> et 3<sup>ème</sup> trimestres de grossesse.

La construction de ce questionnaire s'est appuyée sur un certain nombre de questionnaires et de méthodes d'évaluation des expositions d'études précédentes : (1) Cordier et al. (1997) pour l'évaluation des expositions professionnelles aux solvants – avec l'aide d'une description complémentaire des tâches professionnelles pour certaines professions des mères, (2) Lorente et al. (2000) pour l'évaluation des consommations d'alcool et de tabac, (3) Frelut et al. (1995) pour le questionnaire alimentaire ciblé sur l'apport en folates. Ce questionnaire alimentaire est un questionnaire de fréquence, il est basé sur la distribution de plusieurs groupes d'aliments contenant des folates et consommés en France. Ces groupes d'aliments incluent les céréales pour petit déjeuner dont le contenu moyen en acide folique les classe parmi les aliments contribuant à l'apport. Ce questionnaire est croisé avec une table de composition alimentaire. Les expositions domestiques aux solvants au cours de la grossesse ont été évaluées en fonction de l'utilisation déclarée de certains produits (peintures, vernis, produits de nettoyage, cosmétiques).

#### **A.2.ii. Evaluation des expositions professionnelles**

Pour l'évaluation des expositions professionnelles, l'ensemble des réponses aux questionnaires a été revu par un expert chimiste en aveugle du statut cas-témoins, lequel a attribué à chaque exposition trois paramètres : fréquence, concentration et fiabilité de l'exposition.

Les expositions principalement évaluées étaient les solvants, les composés du plomb, les gaz propulseurs, les radiations ionisantes et non-ionisantes. Les trois principales classes chimiques de solvants ont été détaillées :

- en 5 sous-classes pour la classe des solvants oxygénés : alcools aliphatiques, aldéhydes aliphatiques, esters aliphatiques, cétones aliphatiques et éthers de glycol,
- en 2 sous-classes pour les solvants chlorés : alcanes et alcènes chlorés,
- en 3 produits pour les solvants pétroliers : essences minérales, hydrocarbures aromatiques mononucléaires, carburants.

## B. Informations génétiques

### B. 1. Prélèvements biologiques

Des prélèvements de sang ont pu être obtenus pour 408 enfants (82% des enfants inclus dans l'étude) ; pour 285 d'entre eux, un prélèvement est également disponible pour la mère. Des échantillons sanguins ont été collectés à la fois chez l'enfant et ses deux parents pour 212 familles.

**Tableau 18: Support de prélèvement sanguin par catégorie de sujets (France, 1998-2001)**

Effectif	Prélèvement veineux	Buvard épais (Isocode card)	Buvard du test de Guthrie	Total de prélèvement disponible	%*
<b>Familles d'enfant atteint de fente orale non syndromique (+ syndromique)</b>					
Enfant	200 (+21)	14	0	214 (+21)	89%
Mère	217 (+21)	0 (+1)	0	217 (+22)	90%
Père	204 (+22)	1	0	205 (+22)	85%
<b>Familles d'enfant témoin</b>					
Enfant	12	18	143	173	73%
Mère	42	14	6†	62	26%
Père	0	0	0	0	-

\* Pourcentage de sujet ayant un prélèvement sanguin disponible, parmi les sujets inclus dans l'étude (240 cas non syndromiques et 236 témoins)

† exceptionnellement pour le prélèvement sanguin de six mères d'enfant-témoin, nous avons eu recours à des buvards utilisés pour les tests de Guthrie

Différents supports de prélèvements sanguins ont été utilisés selon les sujets. Nous avons eu recours à des prélèvements veineux pour la majorité des enfants-cas et leurs parents, et pour les mères des enfants-témoins. Les buvards des tests de Guthrie nous ont permis d'obtenir des prélèvements sanguins pour la majorité des enfants-témoins. Des buvards épais (Isocode cards, Schleicher & Schuell, Keene, NH, USA) adaptés aux techniques d'extraction d'ADN par PCR (*Polymerase Chain Reaction*) ont permis de compléter le recueil de ces informations lorsque le prélèvement veineux n'était pas adapté (Tableau 18).

### B. 2. Génotypage

Les polymorphismes d'intérêt ont été génotypés dans leur majorité par le laboratoire U525 (« Génétique épidémiologique et moléculaire des pathologies cardiovasculaires », Faculté de Médecine de la Pitié Salpêtrière, Paris, France) de l'Inserm, suivant différentes techniques adaptées aux spécificités des polymorphismes d'intérêt.

Pour chaque polymorphisme, l'ADN génomique de chaque individu a été amplifié par PCR, cet ADN génomique étant directement extrait à partir du sang, ou bien à partir de buvards Isocode Card ou Guthrie préalablement amplifié par DOP-PCR (*Degenerate Oligonucleotide Primed amplification/Polymerase Chain Reaction*), technique décrite par Barboux et al (2001) (puis détaillée dans Chevrier et al. 2005).

Le génotypage a été réalisé pour la majorité des polymorphismes étudiés (*ADH1C Ile349Val*, *CYP1A1 I462V*, *CYP1A1 T+1189C*, *MSX1 T+2090G*, *MTHFR C677T*, *TGF $\alpha$  Taq1*, *TGF $\beta$ 3 T-24 in4C*) par la technique des ASO (*Allelic Specific oligonucleotid*) décrite par Saiki et al. (1986) à l'aide de sondes spécifiques de l'allèle muté et de l'allèle sauvage, marquées radioactivement. Cette technique consiste à hybrider, dans un premier temps, la sonde mutée sur le produit de PCR de chaque individu, préalablement blotté sur une membrane. Après déshybridation, l'hybridation de la sonde sauvage est réalisée sur le même produit de PCR. Un signal uniquement avec la sonde sauvage est observé pour un génotype homozygote sauvage, un signal uniquement avec la sonde mutée pour un génotype homozygote muté, et un signal simultané avec les deux sondes pour un génotype hétérozygote.

Les deux polymorphismes du gène *CYP2E1*, *DraI* et *RsaI*, engendrent une disparition d'un site de coupure par les enzymes de restriction respectives *DraI* et *RsaI*. Le génotypage a été réalisé par digestion enzymatique du produit de PCR par ces enzymes, comme décrit par Sarmanova et al. (2001).

Pour les gènes *GSTM1* et *GSTT1*, les polymorphismes d'intérêt se définissent par la délétion d'une région génomique dans laquelle se trouve ces gènes. Lorsqu'un individu porte cette délétion, l'amplification par PCR ne peut pas avoir lieu. Les produits de PCR ont été déposés sur gel d'agarose, puis deux génotypes ont été déterminés directement par l'absence (i.e. délétion à l'état homozygote, appelée génotype *Null*) ou la présence (i.e. présence du gène à l'état hétérozygote ou homozygote) de l'amplification, comme décrit par Arand et al. (1996).

Des difficultés du laboratoire U525 d'amplification d'ADN à partir des buvards de test de Guthrie ont abouti à un grand nombre d'informations génétiques manquantes dans le groupe de témoins, puisque la majorité de ce groupe est concernée par ce type de support (plus de 50% de valeurs manquantes parmi les Guthrie pour l'ensemble des polymorphismes, excepté les gènes *ADH1C* et *TGF $\beta$ 3*). Ces taux élevés d'informations génétiques manquantes dans le groupe de témoins nous ont conduits à ne pas exploiter les analyses cas-témoins, se préservant, de la sorte, d'un risque sévère de biais de sélection et d'un manque certain de puissance statistique.

Nous avons eu recours à deux autres laboratoires (Institute of Biosciences and Technology, Texas A&M University System Health Science Center, Houston, Texas, U.S.A. ; et Birth Defects Research Lab, Children's Hospital Oakland Research Institute (CHORI), Oakland, CA, U.S.A.) réalisant, tous deux, des extractions d'ADN en routine à partir d'une quantité infime de matériel sanguin.

Le polymorphisme *MTHFR C677T* pour le laboratoire de Houston et les polymorphismes *GSTM1* et *GSTT1* pour le laboratoire d'Oakland, ont ainsi été déterminés à nouveau pour l'ensemble des buvards épais ou de Guthrie, utilisés pour les enfants, par des techniques d'enzyme de restriction (Frosst et al. 1995, Arand et al. 1996). Pour chacun de ces polymorphismes, le groupe de sujets génotypés avec succès par l'U525 et par le laboratoire américain a été défini comme échantillon de validation, contenant 76 échantillons communs pour le gène *MTHFR* (32 buvards épais et 44 buvards Guthrie) et 32 échantillons communs pour les deux polymorphismes de *GST* (puisque aucun buvard de Guthrie n'a pu être génotypé par l'U525). La stratégie adoptée était alors de valider le double génotypage si le taux de concordance des résultats de génotypage sur l'échantillon de validation était supérieur à 95%. Cette limite, fixée par Little et al. (2002), permet de satisfaire à une bonne qualité des données génétiques. Une discordance a été observée pour le polymorphisme *C677T* du gène

*MTHFR* (parmi 76 échantillons communs), et aucune discordance (parmi 32) pour les deux polymorphismes de *GST*. Le sujet concerné par la discordance a été exclu de l'analyse portant sur le gène *MTHFR*.

Les personnels des laboratoires de génotypage ont travaillé à l'aveugle du statut cas-témoins et enfant-parent. Le laboratoire U525 de l'Inserm a procédé à une relecture des génotypes et une double saisie des résultats. En revanche, aucun des trois laboratoires n'a réalisé d'échantillon interne de validation, comme il est pourtant recommandé par Little et al. (2002), répertoriant les bonnes pratiques à suivre pour des informations génétiques reproductibles et fiables.

### **B. 3. Effectifs des informations génétiques disponibles**

Quelques incohérences de transmission familiale ont pu être détectées parmi les génotypes déterminés, mais leur effectif reste très limité (n=2 familles pour *ADH1C*, n=3 pour *CYP2E1*, n= 2 pour *CYP1A1*, n=3 pour *MTHFR*, n=2 pour les *GST*, n=2 pour *TGF $\alpha$* , n=2 pour *TGF $\beta$ 3*, n=3 pour *MSX1*). Lorsque des incohérences sont observées pour plusieurs polymorphismes, nous choisissons d'exclure ces familles de l'ensemble des analyses faisant intervenir l'information génétique (n=2 familles). Ces incohérences peuvent, par exemple, révéler *post-hoc* des problèmes de paternité non vérifiée ; ces familles doivent alors être exclues. Lorsque ces incohérences sont isolées à un seul polymorphisme, nous choisissons d'exclure ces familles des analyses pour l'étude de ce polymorphisme. Ces incohérences isolées peuvent traduire des erreurs de génotypage.

L'effectif final des informations génotypiques disponibles est présenté dans le Tableau 19 pour chaque polymorphisme.

Tableau 19 : Effectifs des informations génétiques disponibles

			Témoins n=236		Cas non syndromiques n=240			
			Enfant	Mère	Enfant	Mère	Père	Triade*/Diade**
Nombre de prélèvements sanguins			173	62	214	217	205	191 / 16
<b>Gène</b>	<b>Nom précis des polymorphismes issus du génotypage</b>	<b>Noms des formes sauvage/mutée issus de nomenclature</b>						
<b>Gènes « métaboliseurs »</b>								
<i>ADH1C</i>	<i>Ile349Val</i>	*1 / *2	115	54	205	211	195	174 / 23
<i>Cyp2E1</i>	<i>RsaI</i> ou <i>C-1053T</i>	*1 / *5	38	50	204	208	195	175 / 21
	<i>DraI</i>	*1A / *6	21	40	194	206	191	160 / 27
<i>Cyp1A1</i>	<i>Ile462Val</i>	*1A / *2C	41	51	206	211	195	175 / 23
	<i>T+1189C</i>	*1A / *2A	43	45	200	209	194	171 / 21
<i>GSTM1</i>	Grande délétion (état homozygote)	Wild / Null	165	55	209	211	196	179 / 22
<i>GSTT1</i>	Grande délétion (état homozygote)	Wild / Null	165	55	208	211	196	179 / 22
<i>MTHFR</i>	<i>Ala677Val</i> ou <i>C677T</i> ou <i>Hinfl</i>	C / T	168	59	207	209	195	176 / 23
<b>Gènes du développement embryonnaire</b>								
<i>TGF<math>\alpha</math></i>	<i>TaqI</i>	C1 / C2	23	41	203	207	196	172 / 23
<i>TGF<math>\beta</math>3</i>	<i>T24 in4C</i> (ou <i>X5.1</i> )		90	54	206	210	195	175 / 23
<i>MSX1</i>	<i>T+2090G</i>		30	45	192	205	186	160 / 22

\*Triade : famille avec les génotypes déterminés de l'enfant et des deux parents

\*\*Diade : triade avec un génotype parental manquant

#### B. 4. Nomenclature

Les noms des gènes cités dans ce projet satisfont à la nomenclature des gènes approuvée par le HGNC, *HUGO Gene Nomenclature Committee* (à la date de janvier 2005). Cette base de données de gènes humains (*Human Gene Nomenclature Database*) est disponible à l'adresse suivante : <http://www.gene.ucl.ac.uk/nomenclature/>.

La notation des SNPs (*Single-Nucleotide Polymorphisms*) est fonction de la position de la mutation sur le brin d'ADN et des nucléotides ou acides aminés associés qui y sont substitués. Par exemple, le polymorphisme *Ile349Val* du gène de l'alcool déshydrogénase *ADH1C* correspond à un remplacement des acides aminés de l'isoleucine par la valine à la position 349, le polymorphisme de la *MTHFR C677T* est défini par la substitution du nucléotide C par T à la position 677.

Pour les cytochromes P450, qui constituent une famille importante de gènes, nous adoptons la nomenclature allélique recommandée à la suite d'un consensus établi entre deux systèmes différents de nomenclatures (Antonarakis 1998, Nebert 2000, Ingelman-Sundberg et al. 2001). Ces données sont disponibles à l'adresse suivante : <http://www.imm.ki.se/CYPalleles/>. Il est recommandé d'écrire les génotypes par la combinaison des notations des allèles séparée par « / » (par exemple pour le *CYP2E1*\**DraI*, l'état homozygote-muté est noté *CYP2E1*\*6/\*6, hétérozygote, *CYP2E1*\*1A/\*6, homozygote-sauvage *CYP2E1*\*1A/\*1A, où *CYP2E1*\*1A est l'allèle sauvage).

La notation de l'allèle *CYP2E1*\*5 (ou \**RsaI*) est issue de la nomenclature de Garte et Crosti (1999). Il est inclus dans deux haplotypes, notés \*5A et \*5B, dans la nouvelle nomenclature. Dans notre étude, le polymorphisme d'intérêt est un SNP correspondant à la substitution suivante *C1053T*. Nous l'étudions en tant que tel.

Pour reconnaître l'enzyme du gène, il est convenu de noter les gènes en italique (par exemple l'alcool déshydrogénase ADH est l'enzyme codée par le gène nommé aussi *ADH*).

# 2

## Stratégie d'analyse

La stratégie d'analyse statistique de ce projet de recherche se compose de trois étapes :

### (1) Recherche de facteurs de risque environnementaux

La mesure de l'association entre une exposition à un facteur environnemental et le risque de fentes orales ne peut être évaluée uniquement dans le cadre du schéma d'étude cas-témoins. En effet, l'approche cas-parents dispose d'un groupe de témoins qui est construit de façon à appairer les témoins aux cas sur l'ensemble des facteurs, excepté le gène d'intérêt.

Les aspects méthodologiques d'évaluation d'une association avec un facteur environnemental ne sont pas présentés dans ce projet. Nous utilisons ici les méthodes traditionnelles des études cas-témoins, classiques et performantes, via la régression logistique (Bouyer et al. 1993). Nous n'aborderons pas les aspects méthodologiques visant à améliorer ces inférences (par exemple par une correction d'erreurs potentielles de mesure, une stratégie d'appariement plus élaborée ou une prise en compte de corrélation).

### (2) Evaluation de l'association avec le facteur génétique et (3) Recherche des interactions gène-environnement

L'association entre un gène et le risque de fentes orales et l'évaluation d'une interaction gène-environnement sont évaluées par les deux schémas d'étude cas-témoins et cas-parents, lorsque les effectifs disponibles sont suffisants.

Pour l'approche cas-parents, nous privilégions les modèles log-linéaires de C.Weinberg et son équipe présentés dans la partie précédente [INTRODUCTION GENERALE]Etat de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement, sections B. 2.ii.b et C.2.i, qui permettent de différencier les effets directs du génotype de l'enfant, des effets « indirects » du génotype de la mère, dans la survenue d'une fente orale, quel que soit le facteur environnemental.

# 3

## Description de la population

Le Tableau 20 présente une première description socio-démographique des populations FL/P et FP et des témoins. Les distributions des expositions à des facteurs de risque potentiels seront présentées de façon complémentaire dans chaque application du projet (partie suivante [RESULTATS]).

### **A. Description des populations de cas non syndromiques et de témoins**

En accord avec la stratégie d'appariement, les distributions des populations cas et témoins sont similaires pour le centre de recrutement et l'origine géographique de la mère. La majorité des mères (84%) est d'origine européenne.

L'âge des enfants-témoins est légèrement inférieur à celui des enfants-cas, en raison du recours aux maternités parisiennes pour le recrutement des témoins. L'âge moyen des mères est de 29.5 années pour les deux groupes de population.

Le niveau scolaire des mères d'enfants cas (ou la catégorie socioprofessionnelle des mères) est plus faible que celui des mères d'enfants-témoins. Cette association a déjà été observée dans plusieurs études précédentes (Cordier et al. 1997, Clark et al. 2003). Près de la moitié des mères d'enfants-témoins réside dans une agglomération de plus de 50 000 habitants, tandis qu'une proportion de 26% est observée parmi les mères d'enfants-cas.

Enfin, près de 10% des familles d'enfant-cas connaissent un antécédent de fente orale parmi les frères et sœurs de l'enfant atteint ou ses parents ou ses grands-parents. Cette proportion élevée d'antécédents parmi le groupe des cas est attendue et concordante avec l'histoire familiale identifiée comme facteur de risque des fentes orales.

Tableau 20: Description des populations cas non syndromiques et témoins (France, 1998-2001)

	FL/P n=164		FP n=76		Témoins n=236	
	n	%	n	%	n	%
<b>Centres</b>						
Lyon, Grenoble	90	54.9	50	65.8	126	53.4
Paris	56	34.1	16	21.1	80	33.9
Clermont-Ferrand	18	11.0	10	13.2	30	12.7
<b>Sexe de l'enfant</b>						
Fille	50	30.5	38	50.0	98	41.5
Garçon	114	69.5	38	50.0	138	58.5
<b>Age de l'enfant à l'interview (en mois)</b>						
<i>Moyenne (écart-type de la moyenne)</i>	5.2	(0.2)	7.9	(0.4)	4.7	(0.3)
<i>Médiane</i>	5.2		7.5		4.1	
<i>Min</i>	2		3		0	
<i>Max</i>	18		15		19	
<b>Origine géographique de la mère</b>						
Afrique subsaharienne	5	3.0	2	2.6	7	3.0
Afrique du Nord	12	7.3	6	7.9	17	7.2
Europe	136	82.9	64	84.2	200	84.7
Autres	11	6.7	4	5.3	12	5.1
<b>Age de la mère à la naissance de l'enfant</b>						
<i>Moyenne (écart-type de la moyenne)</i>	29.6	(0.4)	29.1	(0.5)	29.7	(0.3)
<b>Lieu de résidence</b>						
Communes de moins de 5 000 habitants	58	35.4	28	36.8	59	25.0
Communes entre 5 000 et 50 000 habitants	65	39.6	27	35.5	66	28.0
Communes de plus de 50 000 habitants	41	25.0	21	27.6	111	47.0
<b>Niveau d'étude de la mère</b>						
Primaire/Secondaire	45	27.4	19	25.3	45	19.2
Formation technique	59	36.0	29	38.7	57	24.4
Supérieur au baccalauréat/Universitaire	60	36.6	27	36.0	132	56.4
Manquant	0	-	1	-	2	-
<b>Catégorie socio-professionnelle de la mère en début de grossesse</b>						
Profession scientifique, technique, libérale et assimilé	31	26.7	17	29.8	67	35.8
Direction, cadre, administratif supérieur	2	1.7	0	0	3	1.6
Personnel administratif et assimilé	28	24.1	16	28.1	66	35.3
Personnel commercial, vendeur	10	8.6	3	5.3	12	6.4
Services	25	21.6	16	28.1	32	17.1
Agriculteur, éleveur, forestier, pêcheur, chasseur	4	3.4	0	0	2	1.1
Ouvriers et manœuvres non agricoles, Transport	16	13.8	5	8.8	5	2.7
Sans emploi, étudiant	34	-	16	-	15	-
Manquant	14	-	3	-	34	-
<b>Antécédents familiaux de fentes orales (parmi la fratrie, parents, grand-parents)</b>						
Oui	14	8.5	9	11.8	-	

n : effectif ; % : pourcentage

## B. Discussion des biais potentiels de l'enquête épidémiologique

La discussion des biais potentiels d'une enquête épidémiologique est incontournable. Dans le cadre d'une étude cas-témoins par recrutement hospitalier, différentes sources de variabilité et d'erreur peuvent influencer sur la mesure d'une association entre une exposition à un facteur environnemental ou un facteur génétique et le risque de fentes orales.

### B. 1. Biais de sélection

Les biais de sélection interviennent lors de la constitution de l'échantillon d'enquête, l'existence et l'ampleur de ces biais dépendent du processus de choix des cas et de témoins. Lorsque l'on s'intéresse à une mesure d'association (ou de lien, et non de risque) dans une étude cas-témoins, la situation idéale pour éviter les biais de sélection est de privilégier la comparabilité des groupes cas et témoins, excepté sur le facteur d'intérêt, plutôt que de favoriser la représentativité des groupes par rapport à la population-cible. En pratique, la sélection des sujets doit être menée indépendamment du (ou des) facteur(s) d'intérêt. Un exemple simple : la sélection des sujets, cas ou témoins, dépend de la disponibilité pratique des enquêteurs et des lieux d'interviews. Selon les conditions d'enquête, une telle sélection est généralement assimilée à un tirage au sort et ne génère pas de biais.

#### B.1.i. Le recrutement hospitalier

L'inclusion des sujets par recrutement hospitalier dépend de la probabilité d'admission des sujets dans les services participants, probabilité qui résulte d'un grand nombre de paramètres liés au service (spécialisation, disponibilité...) ou liés aux malades (gravité de la maladie, lieu de résidence, catégorie socioprofessionnelle...). Cette situation peut conduire à une sélection d'une sous-population des cas, et sans un choix approprié du groupe de témoins, introduire des biais de sélection.

Dans notre étude, les enfants-cas ont été recrutés lors de la chirurgie initiale traitant la fente orale. Ce choix permet d'exclure, d'une part, les cas polymalformés les plus graves (formes syndromiques), et d'autre part, les formes *a minima* sans nécessité de réparation chirurgicale, comme les luettes bifides, les insuffisances vélaires sans dysraphie réelle ou les encoches labiales. Nous répertorions ainsi la grande majorité des cas de fentes orales non syndromiques (95%) sans sélection de sous-groupes (Saal 2002).

Un biais de sélection peut apparaître lorsque les probabilités d'admission à l'hôpital sont différentes chez les cas et les témoins (atteints d'une autre pathologie que les fentes orales) et lorsqu'elles ont un lien avec le facteur d'intérêt. Cette situation est en pratique inévitable, toutefois, le protocole prévoyait de minimiser le biais de sélection dans cette étude par un appariement des témoins sur les cas selon l'éloignement du lieu de résidence au bassin de recrutement hospitalier et sur leur origine géographique.

Le recrutement hospitalier des cas réalisé dans le centre parisien sous-évalue les FP. Ce biais pourrait s'expliquer par une « orientation préférentielle » de certains actes chirurgicaux au service de chirurgie pédiatrique de l'Hôpital Trousseau.

Les pathologies des témoins hospitalisés ne doivent pas être liées au facteur de risque d'intérêt (ou protecteur) ; ce lien est connu pour sous-estimer (ou sur-estimer) l'association étudiée. Ainsi, il est recommandé de diversifier les causes d'hospitalisation des témoins sélectionnés afin de ne pas être tributaire d'une seule pathologie dans le groupe de témoins et de son lien inconnu avec le facteur d'intérêt. Dans notre étude, le protocole de sélection de témoins prévoyait d'exclure les sujets atteints de malformations congénitales, de cancer, ou de maladie génétique. Puisqu'en pratique très contraignante, cette précaution a conduit, au final, à étendre le recrutement des témoins dans d'autres structures locales. Les principales causes d'hospitalisation des enfants témoins étaient les infections (appareil respiratoire, appareil urinaire,...) ou petites chirurgies (naevus, fractures...), n'ayant pas de causes communes *a priori* avec les fentes orales (Annexe 2).

### **B.1.ii. La non-réponse**

Lors d'une enquête épidémiologique, il n'est pas toujours possible d'inclure l'ensemble des sujets éligibles pour différentes raisons. Dans notre enquête, il s'agit de refus de participation, de problème de maîtrise de la langue française et, plus rarement, d'absence d'interview pour des raisons très diverses (pas de rendez-vous possible avec la mère, enfant abandonné, perte du questionnaire).

Pour s'assurer d'une absence de biais de sélection lié au phénomène de non-réponse, il faudrait comparer les caractéristiques des sujets inclus à celles des sujets exclus, pour lesquels ces informations ne sont pas disponibles. En pratique, nous considérons qu'un tel biais est peu probable ou peu important quantitativement (62 sujets exclus/561 sujets éligibles=11% dans notre étude), c'est-à-dire que nous supposons que l'événement de non-réponse n'est pas lié au facteur d'intérêt. Cette hypothèse est discutable en particulier lorsque les sujets exclus comprennent en majorité des individus ne maîtrisant pas la langue française, ce qui implique alors l'exclusion d'un groupe de sujets très particuliers.

Le taux global de participation dans cette étude ainsi calculé (de 89%) est déformé, puisqu'un recrutement, spécifique au centre parisien, a été mis en place en cours d'enquête pour contrecarrer un manque de participation. Ce recrutement est particulier au groupe de témoins du centre parisien. Il explique alors probablement la non-comparabilité des cas et des témoins pour les facteurs reflétant le statut socio-économique des mères tels que le niveau scolaire, la catégorie socio-professionnelle, ou le lieu de résidence (qui est associé au niveau scolaire de la mère ( $p < 0.0001$  sur l'ensemble des sujets,  $p = 0.10$  sur le groupe des témoins)). L'appariement des enfants-témoins avec les enfants-cas sur l'éloignement du lieu de résidence des mères au bassin de recrutement hospitalier (défini par le département) réalisé dans cette étude est de trop faible qualité pour atténuer ces différences de sélection entre les groupes des cas et des témoins.

Une attention doit aussi être portée sur l'influence de l'information manquante génétique lors de la mesure d'association entre un facteur génétique et le risque de maladie : plusieurs autres processus de « non réponse » peuvent aussi induire un biais de sélection comme l'absence de prélèvement sanguin dû à un refus de consentement à fournir des données biologiques ou à

une absence paternelle (ou maternelle plus rarement), et l'absence d'information du génotype, liée à des supports de prélèvement inadaptés ou à la nécessité de recourir à des techniques de génotypage plus performantes et plus coûteuses.

Nous avons rencontré, dans notre étude, des difficultés techniques d'extraction d'ADN, différentes selon le statut cas ou témoin, générant des résultats indéterminés. Le support de prélèvement sanguin des témoins est, pour la majorité, différent de celui des enfants-cas et a conduit à une absence différentielle d'informations génétiques plus fréquente chez les témoins que chez les cas. Nous avons alors choisi de ne pas exploiter, dans les analyses statistiques, l'information génétique des témoins lorsqu'elle était minoritaire, ou de compléter ces informations par des techniques de génotypage plus pointues.

## **B. 2. Biais de classement**

Dans une enquête cas-témoins, un biais de classement peut se produire lors de la détermination de l'exposition des sujets. Nous avons été amenés à distinguer deux types d'erreur de classement : les erreurs différentielles qui spécifient des degrés d'erreur de classement de l'exposition différents selon les groupes de cas et de témoins et engendrent un biais divergent de la mesure de l'association, et les erreurs non différentielles si les deux groupes sont sujets à autant d'erreurs de classement, qui génèrent un biais convergeant vers une absence d'association.

Pour limiter ces erreurs (et de façon prioritaire les erreurs différentielles), les enquêteurs doivent être formés et conscients de leur rôle au moment de l'interview, et ne pas influencer l'attention, la mémoire et finalement les réponses au questionnaire des mères, différemment selon le statut de leur enfant cas ou témoin. Le questionnaire et les conditions de l'interview doivent être standardisés. Les évaluations *a posteriori* des expositions doivent s'effectuer à l'aveugle du statut cas-témoins. Plus spécifiquement liées aux erreurs non différentielles, les méthodes d'évaluation des expositions doivent être objectives et reproductibles. Ces précautions ont été adoptées dans notre étude.

Puisque le génotype est invariant, la mesure d'association génétique n'est intrinsèquement pas sujette à un biais de classement. En revanche, les sources possibles d'erreur de classement sont alors liées aux erreurs de génotypage, par la reproductibilité et la fiabilité des techniques ou la connaissance du statut cas-témoins par le personnel. Pour prévenir un éventuel biais de sélection dans la mesure de l'association entre un polymorphisme génétique et le risque d'une maladie, il est fréquent d'apparier les groupes de cas et les témoins sur leurs origines géographiques (ou ethniques).

## **B. 3. Biais de confusion**

Contrairement aux deux premiers biais présentés ci-dessus, le biais de confusion peut être généralement pris en compte et corrigé par les analyses statistiques, si les facteurs de confusion ont été identifiés et correctement mesurés. Un tiers facteur  $F$  est défini facteur de confusion, si la relation entre l'exposition environnementale et la maladie est différente de

celle obtenue sur chaque niveau de F, ou plus littéralement, si F est lié à la fois à l'exposition et à la maladie.

Les solutions de prévention de ce type de biais sont multiples et leur principe revient à mesurer l'effet de l'exposition sur la maladie « toutes choses égales par ailleurs ». Au moment de l'enquête, nous pouvons choisir d'apparier sur le facteur F. Selon les effectifs disponibles, il est possible de stratifier sur les niveaux de F, ou encore d'intégrer un ajustement sur le facteur F dans les analyses statistiques multivariées.

Dans cette étude, la stratégie d'ajustement sur des facteurs potentiels de confusion est double : nous choisissons, d'une part, d'ajuster *a priori* toutes nos analyses cas-témoins sur l'âge de la mère. En outre un ajustement « minimal » sur les facteurs liés au phénomène de sélection, sources de confusion résiduelle (tels que le centre de recrutement et l'origine géographique de la mère) est mis en place. Puis, nous choisissons de prendre en compte l'ensemble des facteurs de risque potentiels qui modifient la valeur absolue de l'odds-ratio d'intérêt (de plus de 10% par exemple) et qui ne sont pas facteurs d'interaction.

Enfin, la mesure d'une association génétique dans une étude cas-témoins est sensible à un phénomène non mesuré de confusion, la stratification génétique de population. Ce biais potentiel est défini dans la partie précédente [INTRODUCTION GENERALE] Etat de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement, section B.1.iv.



# RESULTATS



# 1

## Rôle d'un polymorphisme du gène *ADH1C* en interaction avec la consommation maternelle d'alcool sur le risque de fentes orales non syndromiques

De nombreuses études épidémiologiques ont évalué l'association entre la consommation maternelle d'alcool et le risque de fentes orales non syndromiques et rapportent des résultats variables (cf partie [INTRODUCTION GENERALE] L'étiologie complexe des fentes orales, section A. 5). D'autres ont examiné l'existence potentielle de susceptibilité génétique de l'enfant dans cette association (cf partie [INTRODUCTION GENERALE] L'étiologie complexe des fentes orales, section C). Les gènes candidats de ces études d'interaction gène-environnement sont des gènes connus pour être impliqués dans le développement embryonnaire chez l'homme ou l'animal, tels que *TGF $\alpha$* , *TGF $\beta$ 3* et *MSX1* (Romitti et al. 1999, Mitchell et al. 2001, Jugessur et al. 2003). L'hypothèse d'étude est alors une possible altération de l'activité de ces gènes en présence de la consommation maternelle d'alcool.

Cette recherche vise à clarifier les mécanismes plausibles d'action liés à la consommation maternelle d'alcool sur la survenue d'une fente orale. Nous faisons alors l'hypothèse qu'un gène impliqué dans le métabolisme de l'alcool doit modifier l'association entre l'exposition à l'alcool du fœtus et son risque d'être atteint d'une fente orale.

Le métabolisme de l'éthanol se réalise en deux étapes d'oxydation : la première transforme l'éthanol en acétaldéhyde (selon plusieurs voies enzymatiques), puis la seconde l'acétaldéhyde en acétate (par l'aldéhyde déshydrogénase ALDH) (Bosron et Li 1986). Les deux systèmes majeurs et les mieux établis, impliqués dans l'oxydation de l'éthanol en acétaldéhyde, incluent la famille de gènes codant pour l'alcool déshydrogénase (famille de gènes *ADH*) et le système microsomal d'oxydation de l'éthanol (MEOS) faisant intervenir le gène du cytochrome P450-III $\epsilon$ 1. La voie de l'ADH est prépondérante. La famille des gènes *ADH* est aussi connue pour être impliquée dans le métabolisme d'autres alcools : tels que le méthanol, certains éthers de glycol, le rétinol (Expertise Collective Inserm 1999, Crabb et al. 2004).

Parmi les loci de la famille des gènes humains *ADH*, deux sont polymorphiques et appartiennent à la classe I de la famille, il s'agit de *ADH1B* et *ADH1C*. Trois allèles codant pour des enzymes à propriétés cinétiques différentes, ont été identifiés pour chacun des deux gènes, la fréquence des différents allèles variant entre les populations. Ces informations sont présentées dans le Tableau 21.

Tableau 21: Propriétés des formes polymorphiques des gènes humains *ADH1B* et *ADH1C*

Allèle	Propriétés enzymatiques*			Distribution par population (en %)**			
	Sous-unité	Km pour l'éthanol (mM)	V <sub>max</sub> (/min)	Européens Blancs	Américains Blancs	Américains noirs	Japonais
<i>ADH1B*1</i>	β <sub>1</sub>	0.05	9	<b>85</b>	<b>&gt;95</b>	<b>85</b>	15
<i>ADH1B*2</i>	β <sub>2</sub>	0.9	400	15	<5	<5	<b>85</b>
<i>ADH1B*3</i>	β <sub>3</sub>	34	300	<5	<5	<b>15</b>	<5
<i>ADH1C*1</i>	γ <sub>1</sub>	1.0	87	<b>60</b>	<b>50</b>	<b>85</b>	<b>95</b>
<i>ADH1C*2</i>	γ <sub>2</sub>	0.63	35	<b>40</b>	<b>50</b>	15	5
<i>ADH1C*3†</i>	γ <sub>3</sub>	-	-	-	-	-	-

Km : constante cinétique de Mickaelis (homodimères des sous-unités) ; V<sub>max</sub> : vitesse maximale

\* Source : Crabb et al. (2004)

\*\*Source: Bosron et Li (1989)

† Le troisième allèle de l'ADH1C a été découvert récemment par Osier et al. (2002), les propriétés enzymatiques associées ne sont pas encore connues

Notre étude cible sur le gène *ADH1C* (OMIM\*103730), polymorphique dans la population européenne (Tableau 21) et française (Coutelle et al. 1998). Les deux produits enzymatiques gamma1 et gamma2, issus du locus *ADH1C*, se différencient par un polymorphisme découvert par Hoog et al. (1986). Ce dernier est caractérisé par la substitution de l'isoleucine (allèle *ADH1C\*1*) par la valine (allèle *ADH1C\*2*) à la position 349. Les deux enzymes ont une affinité pour l'éthanol similaire et forte (un Km bas). En revanche, l'enzyme gamma2 oxyde l'éthanol moins vite que l'enzyme gamma1. Il n'existe pas de mode connu de transmission du gène. Les génotypes sont notés *ADH1C\*1/\*1* pour le statut homozygote-sauvage, *ADH1C\*1/\*2* pour l'hétérozygotie, et *ADH1C\*2/\*2* pour le statut homozygote pour la mutation.

L'objectif de cette étude est d'évaluer les associations entre la consommation maternelle d'alcool et le polymorphisme *Ile349Val* de l'*ADH1C*, et le risque de fentes orales non syndromiques. Des interactions potentielles entre l'apport en alcool et le polymorphisme *ADH1C* sont testées.

## A. Matériel et Méthodes

Cette étude s'appuie sur les données d'une étude épidémiologique cas-témoins multicentrique (n=164 FL/P, n=76 FP, n=236 témoins) réalisée en France entre 1998 et 2001, décrite précédemment (cf partie [MATÉRIEL & MÉTHODES] Population). Cette étude intègre un schéma d'étude cas-parents (174 triades et 23 diades).

### A. 1. Définition des expositions

La consommation maternelle d'alcool est évaluée sur une période périconceptionnelle, en distinguant les mères qui déclarent avoir consommé de l'alcool durant le mois précédant la conception et qui ont arrêté en début de la grossesse, des mères qui ont continué au premier trimestre de la grossesse. Les niveaux d'alcool ingérés sont calculés à partir de la quantité et la fréquence de consommation par semaine, et de la teneur en alcool pour différents types de

boissons, exprimée en gramme. Nous attribuons 10 grammes d'alcool pour un verre de vin, 12.5 grammes pour une bière ou du cidre et 24 grammes pour un apéritif.

La consommation maternelle de tabac est prise en compte dans l'étude comme facteur potentiel de confusion. Elle est définie par une stratégie similaire : les mères qui déclarent avoir fumé sont considérées comme « ex-fumeuses » si elles ont arrêté avant le mois précédant la grossesse, comme « fumeuse en période préconceptionnelle » si elles ont arrêté avant le premier trimestre, comme « fumeuse » si elles n'ont pas arrêté leur consommation au premier trimestre.

L'apport maternel en folates est évalué à partir des habitudes alimentaires des mères et de l'utilisation de multivitamines contenant de l'acide folique le mois précédant la grossesse (L'apport maternel en folate est défini plus précisément dans la seconde application du projet).

## **A. 2. Analyses statistiques**

### **A.2.i. Comparaison cas-témoins**

Une régression logistique simple est employée pour les analyses statistiques de l'approche cas-témoins. La stratégie minimale d'ajustement intègre les variables du centre (Lyon-Grenoble, Paris, Clermont-Ferrand), de l'âge maternel en classes ( $\leq 30$ , 30-35,  $>35$  ans), de l'origine ethnique de la mère (Afrique subsaharienne, Afrique du Nord, Europe, Autres) et du niveau d'étude de la mère (supérieur au baccalauréat-universitaire, technique, primaire-secondaire).

Des ajustements additionnels sur la consommation maternelle de tabac ou l'apport alimentaire en folate de la mère sont réalisés pour examiner l'association de la consommation maternelle d'alcool avec le risque de fentes orales. Les autres facteurs potentiels de confusion tels que la présence d'une histoire familiale de fentes orales, l'utilisation de suppléments en multivitamines ou la prise de médicaments anti-épileptiques, n'impliquent que trop peu d'effectifs pour être considérés dans la stratégie d'ajustement.

### **A.2.ii. Analyse des triades**

L'association entre le polymorphisme de l'*ADH1C* et le risque de fentes orales ainsi que les interactions gène-environnement, sont évalués par une approche cas-parents. Nous optons pour le modèle log-linéaire proposé par Weinberg et al. (1998, 1999a, avec 40 itérations de l'algorithme) et décrit précédemment (partie [INTRODUCTION GENERALE] Etat de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement, sections B. 2.ii.b et B. 2.ii.d).

La consommation maternelle d'alcool est dichotomisée pour l'étude d'interaction gène-environnement (Umbach et Weinberg 2000), en regroupant les mères qui ont consommé de l'alcool pendant le mois préconceptionnel et le premier trimestre.

Nous examinons la présence d'un possible mécanisme d'empreintes parentales de l'*ADH1C* sur le risque de fentes orales, par le test « PO-LRT » (Weinberg 1999b, décrit dans la partie [INTRODUCTION GENERALE] Etat de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement, section E).

Ces méthodes sont implémentées avec le logiciel SAS (SAS Institute Inc., SAS/STAT® Version 8, Cary, NC, USA).

## B. Résultats<sup>7</sup>

### B. 1. Comparaison cas-témoins

L'ensemble des variables socio-démographiques est décrit dans une section précédente (partie [MATERIEL & METHODES] Description de la population) et présenté dans le Tableau 20 parmi les populations d'enfants cas et témoins.

Le Tableau 22 complémentaire montre un risque augmenté et significatif associé avec la consommation maternelle d'alcool au premier trimestre (OR=1.70, IC95% :1.0,2.9). L'ajustement additionnel sur la consommation de tabac ou sur l'apport alimentaire en folates de la mère augmente l'estimation du risque d'au plus 13% et réduit sa précision, avec respectivement OR=1.83, IC95% :1.0,3.2 et OR=1.91, IC95% :1.1,3.3. Aucun accroissement de risque avec la dose d'alcool ingérée par la mère au premier trimestre n'est suggéré : OR<sub><37.5g/semaine</sub>=1.67, IC95% :0.9,3.1, OR<sub>≥37.5g/semaine</sub>=1.41, IC95% :0.5,3.8. Un risque augmenté de fentes orales est observé parmi les mères ayant consommé de l'alcool le mois précédant la conception puis stoppé leur consommation pendant la grossesse.

L'analyse incluant les 23 formes syndromiques de fentes orales rapporte des risques liés à la consommation maternelle de la mère similaire à ceux observés parmi les fentes orales non syndromiques. Les risques associés à la consommation maternelle d'alcool sont différents pour les deux groupes de FL/P et de FP, avec OR<sub>MoisPréconceptionnel</sub>=1.28, IC95% : 0.7,1.9, OR<sub>1erTrimestre</sub>= 1.87, IC95% : 1.1,3.4 pour les FL/P, et OR<sub>MoisPréconceptionnel</sub>=1.53, IC95% : 0.7,3.3, OR<sub>1erTrimestre</sub> = 1.04, IC95% : 0.4,2.4 pour les FP.

Les statuts « fumeuses » ou « fumeuses en période préconceptionnelle » des mères ne semblent pas être liés au risque de fentes orales non syndromiques. En revanche, un risque augmenté et significatif est observé parmi les mères « ex-fumeuses ». L'apport alimentaire en folates est plus faible parmi les mères de cas que parmi les mères de témoins.

Un total de 85% des enfants-cas et 49% des enfants-témoins a été génotypé pour l'*ADH1C* (Tableau 23). La distribution des génotypes chez les enfants-témoins s'écarte significativement de la distribution attendue sous l'hypothèse de l'équilibre d'Hardy-Weinberg. Cet écart est maintenu parmi les familles dont les parents sont d'origine européenne ( $p_{HWE}=0.004$ ).

Parmi les 205 enfants atteints de fentes orales non syndromiques, 48% sont homozygotes pour l'allèle *ADH1C\*1*, 43% sont hétérozygotes, et 9% sont homozygotes pour l'allèle *ADH1C\*2*; le groupe des 115 enfants-témoins rapportent respectivement 44%, 32% et 24%

<sup>7</sup> Pour l'ensemble des parties Résultats des développements du projet, le terme « significatif » désigne « statistiquement significatif ».

( $OR_{ADH1C*1/*1}=1.00$  ;  $OR_{ADH1C*1/*2}=1.37$ , IC95% : 0.2,2.4;  $OR_{ADH1C*2/*2}=0.41$  ; IC95% : 0.2,0.8 ;  $p_{LRT}=0.007$ ).

Seulement 54 mères d'enfants-témoins ont été génotypées pour l'*ADH1C*, et conduisent aux résultats suivants :  $OR_{ADH1C*1/*1}=1.00$  ;  $OR_{ADH1C*1/*2}=0.93$ , IC95% : 0.5,1.9;  $OR_{ADH1C*2/*2}=0.20$ , IC95% : 0.1,0.5 ;  $p_{LRT}=0.0006$ .

## B. 2. Analyse des triades

### B.2.i. Rôle des génotypes *ADH1C* de l'enfant et de la mère

Le Tableau 24 présente la distribution des génotypes *ADH1C* parmi les 174 triades et 23 diades, construites à partir des familles d'enfants atteints de fente orale non syndromique en fonction du profil parental. Il offre ainsi la possibilité de réaliser le test de déséquilibre de transmission ( $p_{TDT}=0.30$  ;  $RR=0.85$ , IC95%:0.6,1.2 indiquant une transmission de l'allèle *ADH1C\*1* plus élevée chez les enfant-cas, comparée à la transmission de l'allèle *ADH1C\*2*).

L'analyse des triades (Tableau 25) montre que le génotype de l'enfant contribue significativement à l'explication du modèle log-linéaire et indique des risques décroissants pour les génotypes hétérozygote et homozygote pour l'allèle *ADH1C\*2* portés par l'enfant. Le génotype de l'enfant semble jouer un rôle significatif parmi le groupe de FL/P. En revanche, aucun effet significatif du génotype de l'enfant n'est mis en évidence pour le groupe des FP, ce résultat pouvant être expliqué en partie par un nombre modeste de triades. Le génotype de la mère ne semble pas influencer significativement le risque de fentes orales non syndromiques, de FL/P ou de FP. Cependant, les estimations des risques portés par le génotype de la mère décrivent des tendances différentes entre les deux groupes FL/P et FP : un risque réduit de FL/P est observé pour le génotype maternel *ADH1C\*2/\*2*, et le statut hétérozygote *ADH1C\*1/\*2* de la mère est associé à un risque augmenté de FP.

L'analyse limitée aux familles avec les deux parents d'origine européenne considère 145 triades et 16 diades. Elle offre des résultats semblables (mais moins précis) à ceux de l'analyse sur l'ensemble des familles (Tableau 26). Le modèle restreint aux 174 triades complètes (Tableau 27) produit des estimations similaires des effets du génotype de l'enfant à celles présentées dans le Tableau 25, suggérant ainsi un risque décroissant avec l'allèle *ADH1C\*2* porté par l'enfant. La réduction du risque associé au génotype maternel *ADH1C\*2/\*2* est accentuée, conduisant à une contribution significative du génotype de la mère au modèle. Les analyses additionnelles (Tableau 27) incluant les formes syndromiques ou excluant les familles avec un antécédent familial de fente orale (parmi la fratrie, les parents ou les grand-parents) conduisent à des conclusions similaires à celles présentées dans le Tableau 25.

### B.2.ii. Empreinte parentale

Le résultat du test spécifique du rapport de vraisemblance est largement non significatif ( $p_{PO\_LRT}=0.81$ ) et suggère ainsi que l'étiologie des fentes orales non syndromiques n'est pas affectée par des effets d'origine parentale des allèles hérités du gène *ADH1C*.

**B.2.iii. Interaction entre la consommation maternelle d'alcool et les génotypes ADH1C de l'enfant et de la mère**

La Figure 6 ne révèle pas de variations importantes des risques relatifs des génotypes entre les strates des familles de mères ayant ou non consommé de l'alcool en période périconceptionnelle. Ainsi, l'effet conjoint du génotype de l'enfant et de la consommation maternelle d'alcool ne dévie pas de l'hypothèse multiplicative des deux effets propres ( $p_{LRT}=0.77$ ). De même, il n'y a pas d'évidence d'interaction entre le génotype de la mère et sa consommation d'alcool ( $p_{LRT}=0.85$ ).

L'inclusion des formes syndromiques ne modifie pas les résultats. Les résultats d'interaction entre le polymorphisme de l'*ADH1C* et d'autres expositions à l'alcool sont présentés en Annexe 3 : notamment, la consommation d'alcool au premier trimestre de grossesse est étudiée en interaction avec le génotype *ADH1C* et aboutit aux mêmes conclusions.

## B. 3. Tableaux et figures

Tableau 22: Description des populations de cas non syndromiques et de témoins (France, 1998-2001)

	FL/P		FP		Témoins		OR*	95% IC
	n	%	n	%	n	%		
Total	164		76		236			
<b>Age de la mère</b>								
≤30 années†	93	56.7	47	61.8	139	58.9	1.00	
30-35 années	52	31.7	21	27.6	71	30.1	1.04	0.7, 1.6
>35 années	19	11.6	8	10.5	26	11.0	1.02	0.6, 1.9
Moyenne		29.6		29.1		29.7		
<b>Niveau d'étude de la mère</b>								
Manquant	-	-	1	-	2	-		
Primaire-Secondaire	45	27.4	19	25.3	45	19.2	2.36	1.4, 3.9
Formation technique	59	36.0	29	38.7	57	24.4	2.36	1.5, 3.7
Supérieur au baccalauréat-Universitaire†	60	36.6	27	36.0	132	56.4	1.00	
<b>Consommation maternelle d'alcool</b>								
Pas de consommation en période périconceptionnelle†	109	66.5	54	71.1	169	71.6	1.00	
Consommation durant le mois préconceptionnel	23	14.0	13	17.1	37	15.7	1.34	0.8, 2.3
Moyenne de la consommation (médiane) en gramme/semaine		44 (34)		75 (34)		52 (37)		
Consommation au 1 <sup>er</sup> trimestre de la grossesse	32	19.5	9	11.8	30	12.7	1.70	1.0, 2.9
Moyenne de la consommation (médiane) en gramme/semaine		32 (24)		129 (24)		37 (20)		
<b>Statut tabac de la mère</b>								
Non fumeuse†	77	47.0	37	48.7	122	51.7	1.00	
Ex-fumeuse	28	17.1	9	11.8	21	8.9	1.92	1.0, 3.6
Fumeuse en période préconceptionnelle	14	8.5	8	10.5	24	10.2	0.89	0.5, 1.7
Fumeuse	45	27.4	22	28.9	69	29.2	0.77	0.5, 1.2
<b>Apport alimentaire en folate de la mère (µg/jour)</b>								
Moyenne		252		234		270		
<b>Utilisatrice de multivitamines contenant de l'acide folique durant le mois précédant la conception et le 1<sup>er</sup> trimestre</b>								
Oui	3	1.8	4	5.3	2	0.8		
<b>Prise de médicaments anti-épileptiques</b>								
Oui	1		1		1			
<b>Antécédents familiaux de fentes orales (parmi la fratrie, parents, grand-parents)</b>								
Oui	14	8.5	9	11.8	0	0		
<b>Information disponible de l'ADH1C</b>								
Chez les enfants	145	88.4	60	78.9	115	48.7		
Nombre de triades	123	75.0	51	67.1				
Nombre de diades	17	10.4	6	7.9				

Note : Ce tableau est complémentaire au Tableau 20

n: effectif ; %: pourcentage ; OR: Odds-Ratio ; IC: Intervalle de Confiance

\* Ajustement sur le centre (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel en classes (≤ 30, 30-35, >35 ans), l'origine ethnique de la mère (Afrique subsaharienne, Afrique du Nord, Europe, Autres) et le niveau d'étude de la mère (supérieur au baccalauréat-universitaire, technique, primaire-secondaire)

† Catégorie de référence

**Tableau 23: Distribution des génotypes de l'ADH1C parmi les enfants et parents des populations cas et témoins (France, 1998-2001)**

Génotype	ADH1C*1/*1		ADH1C*1/*2		ADH1C*2/*2		Total n	Fréquence allélique ADH1C*2	p HWE
	n	%	n	%	n	%			
<b>FL/P</b>									
Enfant	72	49.7	57	39.3	16	11.0	145	0.31	0.36
Mère	64	43.2	73	49.3	11	7.4	148	0.32	0.11
Père	59	43.7	57	42.2	19	14.1	135	0.35	0.39
<b>FP</b>									
Enfant	26	43.3	32	53.3	2	3.3	60	0.30	0.04
Mère	23	36.5	33	52.4	7	11.1	63	0.37	0.34
Père	31	51.7	21	35.0	8	13.3	60	0.31	0.16
<b>Fentes orales non syndromiques</b>									
Enfant	98	47.8	89	43.4	18	8.8	205	0.30	0.73
Mère	87	41.2	106	50.2	18	8.5	211	0.34	0.07
Père	90	46.2	78	40.0	27	13.8	195	0.34	0.14
<b>Témoins</b>									
Enfant	51	44.3	37	32.2	27	23.5	115	0.40	<0.001
Mère	16	29.	22	40.7	16	29.6	54	0.50	0.17

n: effectif ; %: pourcentage ; p HWE : degré de signification du test de comparaison à la distribution attendue sous l'hypothèse de l'équilibre d'Hardy-Weinberg

**Tableau 24: Distribution des génotypes de l'ADH1C des familles d'enfants atteints de fentes orales non syndromiques, selon le profil parental (France, 1998-2001)**

Profil parental	Génotype ADH1C		Enfant			
	Mère	Père	*1/*1	*1/*2	*2/*2	
6	*1/*1	*1/*1	28	-	-	
5	*1/*1	*1/*2	18 *	13 †	-	
3	*1/*1	*2/*2	-	14	-	
5	*1/*2	*1/*1	25 *	21 †	-	
4	*1/*2	*1/*2	9 **	18 *†	7 ††	
2	*1/*2	*2/*2	-	4 *	4 †	
3	*2/*2	*1/*1	-	4	-	
2	*2/*2	*1/*2	-	3 *	3 †	
1	*2/*2	*2/*2	-	-	3	174 triades

Quand le génotype de la mère est manquant

Père	*1/*1	3	1	-	
	*1/*2	0	1	0	
	*2/*2	-	0	0	5 diades

Quand le génotype du père est manquant

Mère	*1/*1	8	0	-	
	*1/*2	3	2	1	
	*2/*2	-	4	0	18 diades

Paires (enfant, parent hétérozygote) informatives pour le calcul du TDT (cf Annexe 3):

\* L'allèle \*1 est transmis par le parent hétérozygote

† L'allèle \*2 est transmis par le parent hétérozygote

**Tableau 25: Estimations des effets des génotypes de l'ADH1C de l'enfant et de la mère sur le risque de fentes orales nonsyndromiques, par l'approche cas-parents en présence de triades incomplètes (France, 1998-2001)**

	Fentes orales non syndromiques			FL/P			FP		
	RR	IC95% Wald	<i>p</i>	RR	IC95% Wald	<i>p</i>	RR	IC95% Wald	<i>p</i>
Nombre de triades	174			123			51		
Nombre de diades	23			17			6		
<b>Nombre de copies de l'allèle ADH1C*2 porté par l'enfant</b>									
0	<b>1.00</b>			<b>1</b>			<b>1</b>		
1	<b>0.71</b>	0.5, 1.0	<0.05	<b>0.64</b>	0.4, 1.0	0.02	<b>0.95</b>	0.5, 1.9	0.68
2	<b>0.63</b>	0.3, 1.2		<b>0.76</b>	0.3, 1.6		<b>0.33</b>	0.1, 1.7	
<b>Nombre de copies de l'allèle ADH1C*2 porté par la mère</b>									
0	<b>1.00</b>			<b>1</b>			<b>1</b>		
1	<b>1.25</b>	0.8, 1.9	0.20	<b>1.00</b>	0.6, 1.6	0.29	<b>2.57</b>	1.1, 6.1	0.08
2	<b>0.70</b>	0.3, 1.4		<b>0.52</b>	0.2, 1.2		<b>1.15</b>	0.4, 3.5	

RR: Risque Relatif; IC: Intervalle de Confiance. *p* : degré de signification du test de rapport de vraisemblance du génotype

**Tableau 26: Estimations des effets des génotypes de l'ADH1C de l'enfant et de la mère sur le risque de fentes orales nonsyndromiques, par l'approche cas-parents en présence de triades incomplètes, pour les familles d'origine européenne (France, 1998-2001)**

Etude limitée aux familles d'origine européenne									
	Fentes orales non syndromiques			FL/P			FP		
	RR	IC95% Wald	<i>p</i>	RR	IC95% Wald	<i>p</i>	RR	IC95% Wald	<i>p</i>
Nombre de triades	145			100			45		
Nombre de diades	16			13			3		
<b>Nombre de copies de l'allèle ADH1C*2 porté par l'enfant</b>									
0	<b>1.00</b>			<b>1.00</b>			<b>1.00</b>		
1	<b>0.76</b>	0.5, 1.1	0.18	<b>0.68</b>	0.4, 1.1	0.07	<b>1.04</b>	0.5, 2.0	0.94
2	<b>0.65</b>	0.3, 1.3		<b>0.74</b>	0.3, 1.6		<b>0.39</b>	0.5, 2.3	
<b>Nombre de copies de l'allèle ADH1C*2 porté par la mère</b>									
0	<b>1.00</b>			<b>1.00</b>			<b>1.00</b>		
1	<b>1.28</b>	0.8, 2.0	0.11	<b>1.00</b>	0.6, 1.7	0.16	<b>2.79</b>	1.0, 7.7	0.11
2	<b>0.58</b>	0.3, 1.2		<b>0.41</b>	0.1, 1.1		<b>0.99</b>	0.3, 3.2	

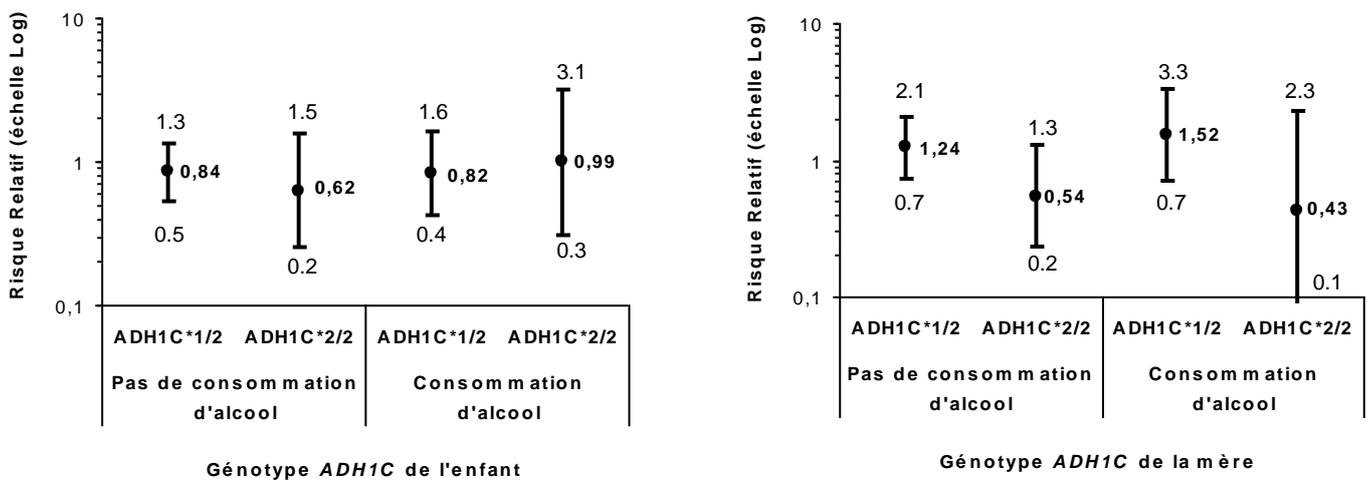
RR: Risque Relatif; IC: Intervalle de Confiance. *p* : degré de signification du test de rapport de vraisemblance du génotype

**Tableau 27: Estimations des effets des génotypes de l'ADH1C de l'enfant et de la mère sur le risque de fentes orales nonsyndromiques, par l'approche cas-parents en présence de triades incomplètes, incluant ou excluant diverses sous-populations (France, 1998-2001)**

	Etude limitée aux triades complètes			Etude incluant les fentes orales syndromiques			Etude excluant les familles avec antécédent familial de fente orale		
	RR	IC95% Wald	p	RR	IC95% Wald	p	RR	IC95% Wald	p
Nombre de triades	174			193			159		
Nombre de diades	0			23			21		
<b>Nombre de copies de l'allèle ADH1C*2 porté par l'enfant</b>									
0	1.00			1.00			1.00		
1	0.83	0.6, 1.2	0.41	0.74	0.5, 1.0	0.06	0.73	0.5, 1.1	0.11
2	0.74	0.4, 1.5		0.70	0.4, 1.3		0.55	0.3, 1.2	
<b>Nombre de copies de l'allèle ADH1C*2 porté par la mère</b>									
0	1.00			1.00			1.00		
1	1.33	0.9, 2.0	0.04	1.19	0.8, 1.7	0.27	1.16	0.8, 1.8	0.37
2	0.51	0.2, 1.1		0.70	0.4, 1.4		0.69	0.3, 1.4	

RR: Risque Relatif; IC: Intervalle de Confiance. p : degré de signification du test de rapport de vraisemblance du génotype

**Figure 6: Risques relatifs de fentes orales non syndromiques des génotypes ADH1C\*1/\*2, ADH1C\*2/\*2 comparé à ADH1C\*1/\*1, porté par l'enfant ou par la mère, pour chaque strate de consommation maternelle d'alcool, estimés par l'approche cas-parents (France, 1998-2001)**



## C. Discussion

Cette étude renforce l'hypothèse du rôle de la consommation maternelle d'alcool au premier trimestre de la grossesse sur le risque de fentes orales non syndromiques. Le génotype de l'enfant paraît jouer un rôle dans la détermination du risque tandis que le rôle du génotype de la mère est mineur et semble plus incertain : l'allèle *ADH1C*\*2 porté par l'enfant a un effet significatif et protecteur dans la survenue d'une fente orale, et seul le statut hétérozygote de la mère est suggéré être associé à un risque augmenté de fentes orales. Finalement, cette étude suggère que les effets des génotypes de l'*ADH1C* sur le risque, s'ils existent, sont indépendants de l'exposition du fœtus à la consommation d'alcool de sa mère. Bien que cette étude soit limitée par une taille modeste de l'échantillon, il s'agit de la première étude à suggérer une association entre le gène de l'*ADH1C* et le risque de fentes orales non syndromiques.

Les résultats des analyses supplémentaires incluant les formes syndromiques de fentes orales, rapportent des conclusions similaires à celles des analyses principales, les syndromes constituant 9% de l'ensemble de la population des cas.

### ECHANTILLONNAGE « EPIDEMIOLOGIQUE ».

La répartition des cas et des témoins parmi les variables socio-démographiques, le choix de la sélection des cas dans des départements hospitaliers de chirurgie, les proportions attendues entre les groupes FL/P et FP et le sex-ratio observé pour chaque groupe ont été discutés précédemment dans la section consacrée à la description générale des populations de l'étude (partie [MATERIEL & METHODES] Description de la population).

### CONSOMMATION MATERNELLE D'ALCOOL ET DE TABAC.

Pour prendre en compte la période précoce de fusion des lèvres et du palais pendant le développement embryonnaire (4-12 semaines), l'étude a ciblé et isolé les expositions se produisant en début de grossesse. La consommation d'alcool de la mère pendant le mois préconceptionnel est associée à un risque accru (non significatif) de fentes orales, suggérant une influence potentielle d'un comportement préconceptionnel de consommation d'alcool. Aussi, il est envisageable que les mères d'enfant-cas aient fourni des déclarations imprécises ou erronées, et aient en réalité continué de consommer de l'alcool en début de grossesse.

Le risque diminué de fentes orales porté par les mères « fumeuses » au premier trimestre est surprenant et contradictoire avec le risque augmenté observé pour les femmes « ex-fumeuses ».

### ECHANTILLONNAGE « GENETIQUE » ET VALIDITE DE L'APPROCHE CAS-TEMOIN.

Nous observons une prévalence de l'allèle *ADH1C*\*2 d'approximativement 40% chez les enfants-témoins et de 50% chez leurs mères. Dans la littérature, cette fréquence est d'environ 40% dans la population blanche européenne (Bosron et Li 1986) et entre 39 et 45% dans la population française (Coutelle et al. 1998, Brennan et al. 2004). Le groupe d'enfants vis à vis du gène *ADH1C* semble constituer un groupe témoin valide et génétiquement similaire à la population française. La différence entre la fréquence observée parmi les mères d'enfant-témoin dans notre étude et les fréquences issues de la littérature suggère que ces mères sont moins représentatives.

L'information génétique est disponible seulement pour 49% des enfants-témoins et 23% des mères de témoins, dû essentiellement à des difficultés pratiques de collecte d'échantillons sanguins ou de support de collecte inadapté aux techniques de génotypage. Ces difficultés pourraient alors être la cause principale de l'écart significatif de la distribution des génotypes des enfants-témoins à l'hypothèse de l'équilibre d'Hardy-Weinberg.

La moitié du génotype de l'enfant provient de celui la mère ; par conséquent, l'estimation de risque de l'un par un modèle logistique traditionnel reflète celle de l'autre.

Pour l'ensemble de ces raisons, l'évaluation de l'association des génotypes *ADHIC* avec le risque de fentes orales ou de l'interaction gène-environnement par l'approche cas-témoins traditionnelle ne paraît pas fiable.

#### APPROCHE CAS-PARENTS ET LES DIADES.

Les auteurs du modèle log-linéaire adopté par la stratégie d'analyse cas-parents dans notre étude, ont montré des tests plus puissants des effets des génotypes lorsque les familles-diades sont prises en compte additionnellement aux triades (partie [INTRODUCTION GENERALE] Etat de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement, section B. 2.ii.d et Weinberg 1999). Peu de génotypes parentaux sont manquants dans notre analyse : 2.5% (5 diades) pour les mères et 9% (18 diades) pour les pères. L'utilisation des diades incorporant l'algorithme « expectation-maximization » (EM) au modèle log-linéaire nécessite l'hypothèse que l'absence d'information du génotype parental se distribue de façon aléatoire. Nous observons une proportion de familles avec une mère portant le génotype *ADHIC\*2/\*2* plus forte parmi les 18 familles-diades sans l'information du génotype paternel (n=4) que parmi les 174 triades complètes (n=13). Ainsi, l'intégration des diades dans l'analyse accroît le nombre de mères *ADHIC\*2/\*2* et pourrait en conséquent réduire l'effet apparent protecteur pour ce génotype maternel.

La différence observée entre les triades et les diades peut se produire par hasard ou s'expliquer par une stratification de population des données observées. L'origine ethnique de la mère diffère significativement parmi les diades et les triades : 30% des mères de diades sont non-européennes contre 12% parmi les triades ( $p=0.04$ ). Or une différence génétique entre les diades et les triades peut biaiser les résultats de l'analyse puisque l'algorithme EM base la « reconstruction » des diades sur les informations observées de triades. L'analyse restreinte aux familles d'origine parentale européenne perd en puissance statistique et prévient alors le génotype de l'enfant de contribuer significativement au modèle. Ces résultats rapportent une évidence de rôle protecteur de l'allèle *ADHIC\*2* porté par l'enfant et du génotype de la mère *ADHIC\*2/\*2* dans la survenue d'une fente orale, et suggèrent un risque augmenté associé au statut hétérozygote de la mère.

#### APPROCHE CAS-PARENTS ET LE GENOTYPE MATERNEL.

Dans notre approche cas-parents, la validité des estimations des risques du génotype maternel repose sur l'hypothèse implicite que, dans la population générale, parmi les couples parentaux portant des génotypes différents, la mère et le père ont autant de chances de porter chaque génotype. Cette hypothèse n'est pas vérifiable dans notre étude, puisque l'information génétique chez les pères des témoins n'est pas disponible. En revanche, nous pouvons vérifier parmi les couples parentaux du groupe témoin d'origine ethnique différente (européenne ou non-européenne), que le nombre de mères d'origine européenne (n=10) est approximativement le même que le nombre de pères européens (n=12).

**DESEQUILIBRE DE LIAISON.**

L'association entre le polymorphisme *ADH1C* et le risque de fentes orales non syndromiques peut refléter un déséquilibre de liaison entre le locus *ADH1C* et le locus causal. Le gène de l'*ADH1C* est localisé dans la région chromosomique 4q22. Des études précédentes soutiennent l'hypothèse d'une association entre le risque de fentes orales et un gène localisé sur le chromosome 4q, bien que ce locus ait été identifié dans la région 4q25-4q31.3 (Beiraghi et al. 1994, Mitchell et al. 1995).

Le site de polymorphisme *ADH1C Ile349Val* a souvent été considéré dans les études d'association, comme suffisant pour différencier et déterminer les enzymes gamma1 et gamma2. Les deux produits enzymatiques diffèrent sur deux acides aminés : 271Arg-349Ile détermine gamma1 codée par *ADH1C\*1*, et 271Gln-349Val détermine gamma2 codée par *ADH1C\*2*. Or récemment, Osier et al. (2002) suggère que l'hypothèse d'un déséquilibre de liaison parfait des deux sites aux positions 349 et 271 de l'*ADH1C* est incorrecte. Nous recommandons que les futures études d'association examinent les formes haplotypiques de l'*ADH1C* définies par les deux sites polymorphiques : Ile349Val et Arg271Gln. Cette information, absente dans notre étude, pourrait aider à la compréhension du résultat inattendu du risque augmenté associé à l'hétérozygotie de la mère.

**INTERACTION *ADH1C*-ALCOOL.**

L'effet conjoint des génotypes *ADH1C* (site *Ile349Val*) et de la consommation maternelle d'alcool sur le risque des fentes orales non syndromiques semble être déterminé par l'hypothèse multiplicative des effets propres. L'estimation de l'interaction implique un petit nombre de familles, et les conclusions doivent être établies avec prudence. C'est pourquoi, nous n'estimons pas les interactions séparément pour les groupes FL/P et FP. Par ailleurs, l'étude de cette interaction considère la consommation maternelle d'alcool quelque soit la fréquence de la consommation et la quantité d'alcool ingérée par les mères. Or nous observons dans cette étude que, parmi les mères consommatrices d'alcool au premier trimestre, la moitié a ingéré au plus deux verres de vin ou de bière par semaine. L'interaction issue de l'oxydation de l'éthanol par l'*ADH1C* sur le risque de fentes orales mériterait d'être étudiée pour des consommations d'alcool de la mère en début de grossesse plus élevées, par de futures études plus puissantes.

**CANCER ET *ADH1C*.**

Le polymorphisme de l'*ADH1C* (site *Ile349Val*) a été examiné en relation avec le risque de cancer : il modifie l'association entre la consommation d'alcool et le risque de cancer de la cavité buccale, du pharynx (Coutelle et al. 1997, Harty et al. 1997, Zavras et al. 2002, Brennan et al 2004) et du sein (Freudenheim et al. 1999). Le risque augmenté de cancer est alors attribué à l'exposition accrue à l'acétaldéhyde, une substance suspectée carcinogène, induit par l'allèle *ADH1C\*1*.

**AUTRES ANOMALIES DE GROSSESSE ET *ADH*.**

Bien que l'acétaldéhyde soit suspecté de jouer un rôle dans la survenue d'anomalies congénitales pour lesquelles la consommation maternelle d'alcool est déterminée comme facteur de risque (Hard 2001), aucune interaction n'a été rapportée, à ce jour et à notre connaissance, entre un polymorphisme de l'*ADH1C* et le risque d'une anomalie congénitale. Plusieurs études ont examiné le deuxième gène (*ADH1B*) de la classe I de la famille des gènes *ADH* en relation avec le risque de certaines anomalies de grossesse, mais les résultats ne sont que peu reproduits ou contradictoires. McCarver et al. (1997) montre que le troisième allèle

(*ADHIB\*3*) exerce un effet protecteur parmi la population afro-américaine contre les anomalies congénitales liées à l'alcool, telles qu'un retard du développement mental. L'efficacité catalytique spécifique de l'enzyme, codée par cet allèle, explique cet effet protecteur : elle oxyde l'éthanol à une vitesse supérieure dans la situation d'une consommation excessive d'alcool (suggérant, selon cette hypothèse, que l'éthanol est l'agent tératogène). Sur une population sud-africaine d'origine mixte, Viljoen et al. (2001) observe une fréquence plus importante de l'allèle *ADHIB\*2*, comparé à l'allèle *ADHIB\*1*, parmi les mères de témoins que les mères d'enfant atteint du syndrome d'alcoolisation fœtal (FAS). Les auteurs suggèrent que l'activité métabolique rapide de l'allèle cause une forte exposition à l'acétaldéhyde (substance soupçonnée dans de nombreuses populations pour favoriser un comportement d'aversion à l'alcool, en particulier lorsqu'elle est accompagnée d'une activité réduite de l'ALDH) et ainsi une diminution de la consommation d'alcool possiblement liée au risque réduit de FAS. Stoler et al. (2002) montre que l'allèle *ADHIB\*3* accroît le risque d'avoir un enfant atteint de caractéristiques similaires au FAS. Ils expliquent le risque augmenté conféré à l'allèle par une plus forte consommation d'alcool des mères portant cet allèle.

#### COMPORTEMENT DE CONSOMMATION D'ALCOOL ET ADH.

La relation entre les gènes *ADH* et risque d'alcoolisme est complexe et très débattue (Crabb et al. 2004). Le comportement de consommation d'alcool semble être influencé par la présence de l'allèle *ADHIB\*2*, en particulier dans la population asiatique (Whitfield 2002). Le polymorphisme de l'*ADHIC* pourrait influencer le comportement de consommation d'alcool indirectement par la présence de déséquilibre de liaison avec d'autres gènes de la famille *ADH* et de la famille de *ALDH* (Chen et al. 1999, Osier et al. 1999, Borrás et al. 2000). Cependant, une étude récente, impliquant une collaboration internationale importante, n'observe aucune association entre l'*ADHIC* et la consommation d'alcool dans leur groupe de témoins sains (Raimondi et al. 2004). Parmi les 54 mères de témoins possédant une information génotypique pour l'*ADHIC* dans notre étude, nous observons toutefois une proportion plus faible de mères « buveuses » parmi les mères portant l'allèle *ADHIC\*1* (21%) comparé aux mères ne portant pas cet allèle (44%).

L'existence potentielle d'un lien entre le comportement de consommation maternelle d'alcool et le gène *ADHIC* n'affecte pas les estimations par l'analyse cas-parents de l'interaction entre le génotype de l'enfant et la consommation d'alcool, puisque l'exposition de l'enfant (la consommation maternelle) et son génotype ne peuvent pas être liés. En revanche, ce lien pourrait influencer les estimations d'interaction entre le génotype de la mère et l'exposition de l'enfant à la consommation maternelle d'alcool.

#### HYPOTHESES DE MECANISMES.

L'oxydation de l'éthanol via la voie de l'*ADH* n'est pas détectée durant le premier trimestre de la vie fœtale (van Ooij et al. 1992) ou dans les tissus placentaires (Andersson et al. 1989). Cependant, le polymorphisme *ADHIC* est connu pour réguler la synthèse de l'acide rétinol via l'oxydation du rétinol (Duester et al. 1991, Yang et al. 1994). L'enzyme de l'*ADHIC*, agissant en tant que rétinol déshydrogénase, est exprimée pendant l'embryogenèse de la souris (Zgombic-Knight et al. 1994). Ni l'activité enzymatique *in vivo* chez l'embryon humain, ni les propriétés cinétiques des deux produits enzymatiques de l'*ADHIC*, en tant que rétinol déshydrogénase, n'ont encore été mesurées. L'acide rétinol est impliqué dans la morphogenèse précoce et est ainsi essentiel au développement normal. Certains auteurs postulent que l'inhibition de l'acide rétinol par l'éthanol peut être le résultat d'une

compétition pour l'ADH, et ce mécanisme est suspecté dans l'étiologie des FAS ou autres maladies liées à l'alcool.

#### CONCLUSION.

Pour conclure, cette étude rapporte un résultat nouveau : l'allèle muté du gène *ADH1C* (site *Ile349Val*) semble jouer un rôle protecteur dans la survenue d'une fente orale non syndromique. Parce que le gène *ADH1C* est impliqué dans diverses activités métaboliques, plusieurs hypothèses de voie biologique causale doivent être considérées, dépendant du cadre spatio-temporel de l'activité enzymatique pendant la grossesse chez l'homme (Clayton et McKeigue 2001). La relation entre la consommation d'alcool et les anomalies congénitales implique des mécanismes biochimiques complexes. Des influences génétiques du fœtus, du placenta ou de la mère peuvent intervenir dans cette relation, elle met en scène probablement plusieurs gènes impliqués dans le métabolisme de l'éthanol ou de l'acétaldéhyde, et elle peut interférer avec le métabolisme d'autres nutriments. De futures études sont nécessaires pour reproduire l'association observée dans cette étude, dans diverses populations afin d'identifier les mécanismes sous-jacents.

## 2

# Les génotypes *MTHFR* C677T de l'enfant et de la mère, l'apport maternel en folates et le risque de fentes orales non syndromiques

La nutrition maternelle pendant la grossesse a longtemps été suspectée de jouer un rôle critique dans l'étiologie des fentes orales non syndromiques. Bien qu'il soit à ce jour bien établi que l'apport maternel en folates ait un rôle spécifique dans la prévention des anomalies du tube neural (ATN) (MRC Vitamin Study Research Group 1991, Czeizel et Dudas 1992, Berry et al. 1999), ce rôle dans l'étiologie des fentes orales est incertain (revue des études épidémiologiques dans la partie [INTRODUCTION GENERALE] L'étiologie complexe des fentes orales section A. 3). Les mécanismes d'action de l'acide folique dans la prévention des ATN (ou des fentes orales s'ils existent), sont inconnus.

Le gène méthylène-tétrahydrofolate réductase (*MTHFR*) joue un rôle majeur dans le métabolisme des folates. Il réduit le 5,10-méthylène THF (dérivé des folates), de manière irréversible, en 5-méthyl THF, qui cède son groupe méthyle pour la synthèse de l'homocystéine en méthionine (par la méthionine synthase, enzyme dépendante de la vitamine B12). L'homocystéine, située à un carrefour métabolique, peut également être transsulfurée en cystéine via la cystathionine, ceci sous la dépendance d'enzymes B6 dépendantes. Notre étude s'intéresse au polymorphisme fonctionnel C677T du gène *MTHFR*, commun dans la population caucasienne (Wilcken et al. 2003). La mutation résulte en une substitution de l'acide aminé alanine (C) en valine (T), et produit une forme caractérisée par une activité réduite à forte température, appelée « thermolabile ».

L'association entre le polymorphisme *MTHFR* C677T et le risque d'ATN est bien documentée : le génotype *TT*, homozygote pour le variant, est établi comme un facteur de risque modéré des ATN dans de nombreuses populations (Botto et Yang 2000). Certaines études suggèrent que le statut nutritionnel de la mère modifie cette association (Christensen et al. 1999). Les études testant l'association du polymorphisme *MTHFR* C677T avec le risque de fentes orales fournissent des résultats conflictuels (Gaspar et al. 1999, Shotelersuk et al. 2003, Vieira et al. 2005, revue pour les populations caucasiennes dans la partie [INTRODUCTION GENERALE] L'étiologie complexe des fentes orales section B.2.v). Aucune association avec le risque de fentes orales n'est rapportée en Norvège (FL/P), en Californie, à Maryland, ou au Brésil. Des études récentes aux Pays-Bas, en Italie, en Irlande et Norvège (pour les FP), suggèrent un risque augmenté pour les sujets porteurs de l'allèle *T*. Généralement, le mécanisme pathogène proposé associe l'activité réduite de l'enzyme issue du génotype *TT* avec un niveau d'homocystéine accrue, qui est considérée tératogène dans certaines expérimentations animales (Rosenquist et Finnell 2001), ou avec un état d'hypométhylation, critique au déroulement d'un développement normal (Botto et Yang 2000, van Rooij et al. 2003).

Cette étude utilise deux schémas d'étude, cas-témoins et cas-parents, pour examiner l'association du polymorphisme *MTHFR C677T* avec le risque de fentes orales non syndromiques. L'apport en folates de la mère est évalué à partir des sources alimentaire et vitaminique, pour mesurer les effets potentiels d'interaction entre cet apport et les génotypes de l'enfant et de la mère. Dans cette étude, nous distinguons les deux groupes de fentes orales : les FL/P et les FP.

## A. Matériel et Méthodes

Cette étude s'appuie sur les données d'une étude épidémiologique cas-témoins multicentrique (n=164 FL/P, n=76 FP, n=236 témoins) réalisée en France entre 1998 et 2001 (décrite dans la partie [MATERIEL & METHODES] Population). Cette étude intègre un schéma d'étude cas-parents (126 triades FL/P et 17 diades FL/P ; 50 triades FP et 6 diades FP).

### A. 1. Définition des apports maternels en folates

L'élaboration du questionnaire alimentaire s'appuie sur un travail précédemment validé et spécifiquement ciblé sur la mesure de l'apport en folates (Frelut et al. 1995). Les questions portent sur l'alimentation habituelle des mères, en considérant qu'elle est équivalente à l'alimentation en période préconceptionnelle et en début de grossesse. Elles portent sur un nombre sélectionné de groupes d'aliments connus pour leur concentration importante en folates (céréales au petit déjeuner, légumes verts, foie, etc.). Le niveau de l'apport est calculé à partir de la concentration en folates (en µg) et de la fréquence quotidienne de consommation pour chaque aliment (Favier et al. 1995).

La variabilité de la définition de l'utilisation maternelle de multivitamines (contenant de l'acide folique) des études du risque de fentes orales non syndromiques (Shaw et al. 2003, van Rooij et al. 2003), nous conduit à proposer deux définitions : la première définit une mère « utilisatrice de multivitamines » si elle a été supplémentée (ou a commencé la prise de multivitamines) durant la période allant du mois préconceptionnel au premier trimestre. La seconde définition est plus restrictive et définit une « utilisatrice de multivitamines » si elle utilise des multivitamines à la fois sur la période préconceptionnelle et au premier trimestre. Ainsi, pour cette seconde définition, une mère est considérée comme « non utilisatrice » si elle commence les suppléments à partir du premier trimestre.

La combinaison des deux sources d'apport en folates par la construction d'une mesure commune, DFE (*Dietary Folate Equivalent*), n'est pas envisageable dans notre étude puisque les informations de la posologie et de la dose réelle ingérée des multivitamines par les mères sont trop imprécises.

### A. 2. Analyse statistique

#### A.2.i. Comparaison cas-témoins

Des tests statistiques classiques de comparaison de deux groupes (test du Khi-deux, T-test d'égalité des moyennes) et une régression logistique, ont été utilisés. Les estimations des

effets du génotype *MTHFR C677T* et de l'interaction génotype-apport en folates sont ajustées sur le centre (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel en classes ( $\leq 30$ , 30-35,  $>35$  ans), et l'origine ethnique de la mère (Afrique subsaharienne, Afrique du Nord, Europe, Autres). Puisque seulement 59 génotypes de mères de témoins sont disponibles, les effets du génotype maternel sur le risque de fentes orales ne sont pas évalués par l'approche cas-témoins.

### **A.2.ii. Analyse des triades**

De façon complémentaire, l'association entre le polymorphisme *MTHFR C677T* et le risque de fentes orales ainsi que les interactions gène-environnement, sont évalués par l'approche cas-parents. Nous optons pour le modèle log-linéaire proposé par Weinberg et al. (1998, 1999a, avec 40 itérations de l'algorithme) et décrit précédemment (partie [INTRODUCTION GENERALE] Etat de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement, sections B. 2.ii.b et B. 2.ii.d). Pour l'étude d'interaction gène-environnement (Umbach et Weinberg 2000), nous considérons les facteurs environnementaux dichotomiques suivants : mère utilisatrice de multivitamines contenant de l'acide folique selon les deux périodes d'intérêt définies plus-haut (oui/non), l'apport alimentaire en folates de la mère (niveau « moyen-fort » ou « faible » pour, respectivement, un niveau supérieur ou inférieur au premier tertile calculé sur le groupe témoin). Nous examinons la présence d'un possible mécanisme d'empreintes parentales du polymorphisme *MTHFR C677T* sur le risque de fentes orales, par le test « PO-LRT » (Weinberg 1999b, décrit dans la partie [INTRODUCTION GENERALE] Etat de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement, section E).

Ces méthodes sont implémentées avec le logiciel SAS (SAS Institute Inc., SAS/STAT® Version 8, Cary, NC, USA).

## **B. Résultats**

L'ensemble des variables socio-démographiques est décrit dans une section précédente (partie [MATRIEL & METHODES] Description de la population) et présenté dans le Tableau 20 parmi les populations d'enfants cas et témoins.

Nous observons, dans le Tableau 28, plus de mères d'enfant-cas que de mères d'enfant-témoin ayant eu recours à des suppléments en multivitamines sur la période périconceptionnelle. La majorité de ces mères (92%) ont commencé les suppléments vitaminiques au cours du premier trimestre de grossesse. Peu de mères ( $n=7$  parmi les cas,  $n=2$  parmi les témoins) utilisent des multivitamines à la fois en périodes pré- et post-conceptionnelles, la plupart de ces femmes ont des antécédents familiaux de fentes orales (fratrie de l'enfant, parents ou grand-parents de l'enfant).

La moyenne d'apport alimentaire en folates des mères d'enfant-témoin (270  $\mu\text{g}/\text{jour}$ ) est significativement supérieure à la moyenne des mères d'enfant-cas (246  $\mu\text{g}/\text{jour}$ ,  $p=0.02$ ), plus particulièrement parmi le groupe de FP (234  $\mu\text{g}/\text{jour}$ ,  $p=0.005$ ).

### B. 1. Génotypes *MTHFR C677T*

La distribution des génotypes *MTHFR C677T* parmi les enfants-témoins est de 32% pour l'homozygotie pour l'allèle *C*, de 48% pour l'hétérozygotie et de 20% pour l'homozygotie pour *T* (Tableau 29). Les estimations de l'approche cas-témoins suggèrent que l'allèle *T* porté par l'enfant est protecteur contre les risques de FL/P et de FP. Des ajustements additionnels sur le niveau d'éducation maternel ou la consommation d'alcool ou de tabac de la mère au premier trimestre de grossesse ne modifient pas les estimations (résultats non présentés).

L'approche cas-parents n'observe pas d'association entre le génotype de l'enfant et le risque de fentes orales, lorsqu'elle distingue les génotypes hétérozygotes et homozygotes pour l'allèle *T* (Tableau 30). Sous l'hypothèse d'un mode d'activité dominant, l'allèle *T* est lié à un risque diminué de FL/P et il n'est pas associé au risque de FP. Il n'existe aucune évidence d'un effet des génotypes de la mère sur les risques de FL/P ou de FP. Le modèle réalisé sur les triades complètes conduit à des conclusions similaires (Tableau 31).

Aucun effet d'origine parentale de l'allèle *T* hérité n'est suggéré ( $p_{LRT}=0.6$  pour le groupe FL/P).

### B. 2. Interaction Génotype *MTHFR C677T* et statut en folates de la mère

L'approche cas-témoins montre que l'association entre l'apport en folate de la mère, via l'alimentation ou les vitamines, et le risque de FL/P n'est pas suffisamment différente entre les génotypes de l'enfant pour indiquer la présence d'une interaction (Tableau 32 et tests d'interaction :  $p_{LRT}=0.57$  pour l'apport alimentaire,  $p_{LRT}=0.35$  pour les suppléments vitaminiques,  $p_{LRT}=0.47$  pour l'apport alimentaire lorsque les femmes supplémentées en multivitamines sont exclues du groupe des femmes ayant un faible apport alimentaire en folates). Toutefois, les enfants porteurs du génotype *CC* dont les mères sont supplémentées sont observés avec un risque doublé de FL/P, tandis qu'une absence d'association entre la supplémentation des mères et le risque de FL/P est observée parmi les enfants porteurs des génotypes *CT* et *TT*.

De même, l'analyse des 124 familles-triades d'enfant atteint d'une FL/P ne révèle pas de déviation significative de l'hypothèse multiplicative pour l'effet conjoint du génotype *MTHFR C677T* et du statut en folates de la mère (Figure 7 et tests d'interaction :  $p_{LRT}$  enfant\*folate alimentation=0.54 ;  $p_{LRT}$  enfant\*folate vitamine=0.99,  $p_{LRT}$  mère\*folate alimentation=0.49,  $p_{LRT}$  mère\*folate vitamine=0.11). L'exclusion des femmes supplémentées en multivitamines sur la période périconceptionnelle parmi le groupe de femmes ayant un faible apport alimentaire en folates produit des conclusions similaires (Figure 7-(3)).

Les résultats sur le groupe de FP ne sont pas présentés puisqu'ils n'impliquent que 59 enfants-cas ou 49 triades. Les résultats sur l'ensemble des fentes orales non syndromiques sont présentés en Annexe 4.

### B. 3. Tableaux et figures

Tableau 28: Description des populations cas non syndromiques et témoins (France, 1998-2001)

	FL/P		FP		Témoins	
	n	%	n	%	n	%
	164		76		236	
<b>Mère utilisatrice de suppléments en multivitamines contenant de l'acide folique (le mois précédant la grossesse OU le premier trimestre de grossesse)</b>						
Manquant	1	-	0	-	0	-
Oui	44	26.8	18	23.7	51	21.6
<b>Mère utilisatrice de suppléments en multivitamines contenant de l'acide folique (le mois précédant la grossesse ET le premier trimestre de grossesse)</b>						
Manquant	1	-		-		-
Oui	3	1.8	4	5.3	2	0.8
<b>Apport alimentaires en folates de la mère</b>						
Manquant	2	-	1	-	1	-
<217 µg/jour	75	46.3	36	48.0	78	33.2
217-301 µg/jour	42	25.9	25	33.3	79	33.6
>301 µg/jour	45	27.8	14	18.7	78	33.2
		*		*		
<i>Moyenne</i>		252		234*		270
<i>(Médiane)</i>		(229)		(224)		(250)
<b>Consommation maternelle d'alcool</b>						
Durant le premier trimestre de la grossesse	32	19.5	9	11.8	30	12.7
<b>Consommation maternelle de tabac</b>						
Durant le premier trimestre de la grossesse	45	27.4	22	28.9	69	29.2
<b>Information génétique de MTHFR</b>						
Nombre d'enfants	148	90.2	59	77.6	168	71.2
Nombre de triades	126	76.8	50	65.8		
Nombre de diades	17	10.4	6	7.9		

Note : Ce tableau est complémentaire au Tableau 20

n: nombre, %: pourcentage

\* Groupe significativement différent du groupe de témoins ( $p < 0.05$ ) (test du Khi-deux pour les variables catégorielles, T-test d'égalité des moyennes)

Tableau 29: Estimations des effets du génotype *MTHFR* C677T de l'enfant sur le risque de fentes orales nonsyndromiques, par l'approche cas-témoins (France, 1998-2001)

	Témoins	Fentes orales non syndromiques				FL/P				FP				
		%	%	OR*	IC95% Wald	<i>p</i>	%	OR*	IC95% Wald	<i>p</i>	%	OR*	IC95% Wald	<i>p</i>
	n=168	n=207				n=148				n=59				
<b>Génotype de l'enfant</b>	<b>CC<sup>†</sup></b>	32.1	44.0	<b>1.00</b>		44.6	<b>1.00</b>			42.4	<b>1.00</b>			
	<b>CT</b>	48.2	43.0	<b>0.69</b>	0.4,1.1	0.07	40.5	<b>0.64</b>	0.4,1.1	0.12	49.2	<b>0.83</b>	0.4,1.6	0.15
	<b>TT</b>	19.6	13.0	<b>0.50</b>	0.3,0.9		14.9	<b>0.56</b>	0.3,1.1		8.5	<b>0.35</b>	0.1,1.0	
	<b>CC<sup>†</sup></b>			<b>1.00</b>			<b>1.00</b>				<b>1.00</b>			
	<b>CT/TT</b>			<b>0.63</b>	0.4,1.0	0.04	<b>0.61</b>	0.4,1.0	0.04		<b>0.69</b>	0.4,1.3	0.25	

n: effectif ; %: pourcentage ; OR: Odds-Ratio ; IC: Intervalle de Confiance ; *p* : degré de signification du test du rapport de vraisemblance pour le génotype

\* Ajustement sur le centre (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel en classes ( $\leq 30$ , 30-35,  $>35$  ans), l'origine ethnique de la mère (Afrique subsaharienne, Afrique du Nord, Europe, Autres)

<sup>†</sup> Catégorie de référence

Tableau 30: Estimations des effets des génotypes *MTHFR* C677T de l'enfant et de la mère sur les risques de fentes orales non syndromiques, par l'approche cas-parents (France, 1998-2001)

		Fentes orales non syndromiques			FL/P			FP		
		RR	IC95% Wald	<i>p</i>	RR	IC95% Wald	<i>p</i>	RR	IC95% Wald	<i>p</i>
Nombre de triades		176			126			50		
Nombre de diades		23			17			6		
<b>Génotype de l'enfant</b>	<b>CC<sup>†</sup></b>	<b>1.00</b>			<b>1.00</b>			<b>1.00</b>		
	<b>CT</b>	<b>0.81</b>	0.6,1.2	0.12	<b>0.73</b>	0.5,1.1	0.04	<b>1.05</b>	0.5,2.2	0.99
	<b>TT</b>	<b>0.87</b>	0.5,1.6		<b>0.96</b>	0.5,1.9		<b>0.69</b>	0.2,2.3	
	<b>CC<sup>†</sup></b>	<b>1.00</b>			<b>1.00</b>			<b>1.00</b>		
	<b>CT/TT</b>	<b>0.81</b>	0.6,1.2	0.07	<b>0.75</b>	0.5,1.2	0.05	<b>1.01</b>	0.5,2.1	0.99
<b>Génotype de la mère</b>	<b>CC<sup>†</sup></b>	<b>1.00</b>			<b>1.00</b>			<b>1.00</b>		
	<b>CT</b>	<b>1.17</b>	0.8,1.8	0.64	<b>1.13</b>	0.7,1.9	0.76	<b>1.31</b>	0.6,3.0	0.78
	<b>TT</b>	<b>0.88</b>	0.5,1.7		<b>0.86</b>	0.4,1.9		<b>0.94</b>	0.3,3.2	
	<b>CC<sup>†</sup></b>	<b>1.00</b>			<b>1.00</b>			<b>1.00</b>		
	<b>CT/TT</b>	<b>1.10</b>	0.7,1.6	0.66	<b>1.07</b>	0.7,1.7	0.81	<b>1.21</b>	0.6,2.7	0.66

RR: Risque Relatif ; 95%IC: Intervalle de Confiance; *p* : degré de signification du test du rapport de vraisemblance pour le génotype

<sup>†</sup> Catégorie de référence

**Tableau 31: Estimations des effets des génotypes *MTHFR C677T* de l'enfant et de la mère sur le risque de fentes orales nonsyndromiques, pour diverses conditions de population, par l'approche cas-parents (France, 1998-2001)**

		Etude limitée aux triades complètes			Etude incluant les formes syndromiques			Etude limitée aux familles d'origine européenne		
		RR	IC95% Wald	<i>p</i>	RR	IC95% Wald	<i>p</i>	RR	IC95% Wald	<i>p</i>
Nombre de triades		176								
Nombre de diades		0								
<b>Génotype de l'enfant</b>	<b>CC<sup>†</sup></b>	<b>1.00</b>			<b>1.00</b>			<b>1.00</b>		
	<b>CT</b>	<b>0.78</b>	0.5,1.2	0.12	<b>0.93</b>	0.7,1.3	0.30	<b>0.88</b>	0.6,1.3	0.46
	<b>TT</b>	<b>0.86</b>	0.5,1.6		<b>1.05</b>	0.6,1.9		<b>0.91</b>	0.5,1.7	
<b>Génotype de la mère</b>	<b>CC<sup>†</sup></b>	<b>1.00</b>			<b>1.00</b>			<b>1.00</b>		
	<b>CT</b>	<b>1.05</b>	0.7,1.7	0.79	<b>1.20</b>	0.8,1.8	0.68	<b>1.24</b>	0.8,2.0	0.40
	<b>TT</b>	<b>0.83</b>	0.4,1.7		<b>1.04</b>	0.6,1.9		<b>0.79</b>	0.4,1.6	

RR: Risque Relatif ; 95%IC: Intervalle de Confiance; *p* : degré de signification du test du rapport de vraisemblance

<sup>†</sup> Catégorie de référence

Tableau 32: Estimations des effets du génotype *MTHFR* C677T de l'enfant sur le risque de FL/P pour chaque strate d'apport en folates de la mère, par l'approche cas-témoins (France, 1998-2001)

		Témoins	FL/P		
		%	%	OR*	IC95% Wald
<b>Génotype <i>MTHFR</i> de l'enfant et Apport alimentaire en folates de la mère</b>					
		n=168	n=146		
<b>CC†</b>	Niveau moyen-fort†	23.8	25.3	<b>1.0</b>	
	Niveau faible (<217 µg/jour)	8.3	19.9	<b>2.1</b>	1.0,4.7
<b>CT</b>	Niveau moyen-fort	33.3	24.0	<b>0.7</b>	0.4,1.3
	Niveau faible (<217 µg/jour)	14.9	15.8	<b>1.0</b>	0.5,2.1
<b>TT</b>	Niveau moyen-fort	13.7	6.8	<b>0.5</b>	0.2,1.1
	Niveau faible (<217 µg/jour)	6.0	8.2	<b>1.3</b>	0.5,3.4
	<i>p</i> interaction				0.57
<b>Génotype <i>MTHFR</i> de l'enfant et Mère utilisatrice de suppléments en multivitamines contenant de l'acide folique (le mois précédant la grossesse OU le premier trimestre de la grossesse)</b>					
		n=168	n=147		
<b>CC†</b>	Oui	4.2	11.6	<b>2.3</b>	0.9,6.1
	Non†	28.0	33.3	<b>1.0</b>	
<b>CT</b>	Oui	11.3	10.9	<b>0.8</b>	0.4,1.8
	Non	36.9	29.3	<b>0.7</b>	0.4,1.2
<b>TT</b>	Oui	6.5	4.1	<b>0.5</b>	0.2,1.6
	Non	13.1	10.9	<b>0.7</b>	0.3,1.5
	<i>p</i> interaction				0.35
<b>Génotype <i>MTHFR</i> de l'enfant et Apport alimentaire en folates de la mère (exclusion des femmes supplémentées en multivitamines qui ont un faible apport alimentaire en folates)</b>					
		n=155	n=127		
<b>CC†</b>	Niveau moyen-fort†	25.8	29.1	<b>1.0</b>	
	Niveau faible (<217 µg/jour)	7.1	18.1	<b>2.0</b>	0.9,4.8
<b>CT</b>	Niveau moyen-fort	36.1	27.6	<b>0.7</b>	0.4,1.3
	Niveau faible (<217 µg/jour)	12.3	11.8	<b>0.9</b>	0.4, 1.9
<b>TT</b>	Niveau moyen-fort	14.8	7.9	<b>0.5</b>	0.2,1.1
	Niveau faible (<217 µg/jour)	3.9	5.5	<b>1.4</b>	0.4,4.5
	<i>p</i> interaction				0.47

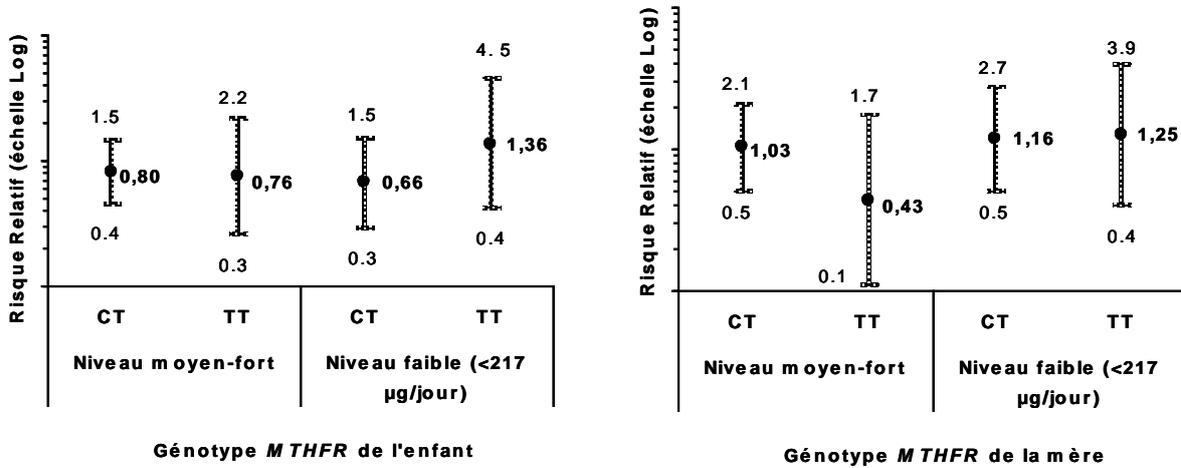
n: effectif ; %: pourcentage ; OR: Odds-Ratio ; IC: Intervalle de Confiance ; *p* : degré de signification du test du rapport de vraisemblance

\* Ajustement sur le centre (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel en classes (≤ 30, 30-35, >35 ans), l'origine ethnique de la mère (Afrique subsaharienne, Afrique du Nord, Europe, Autres)

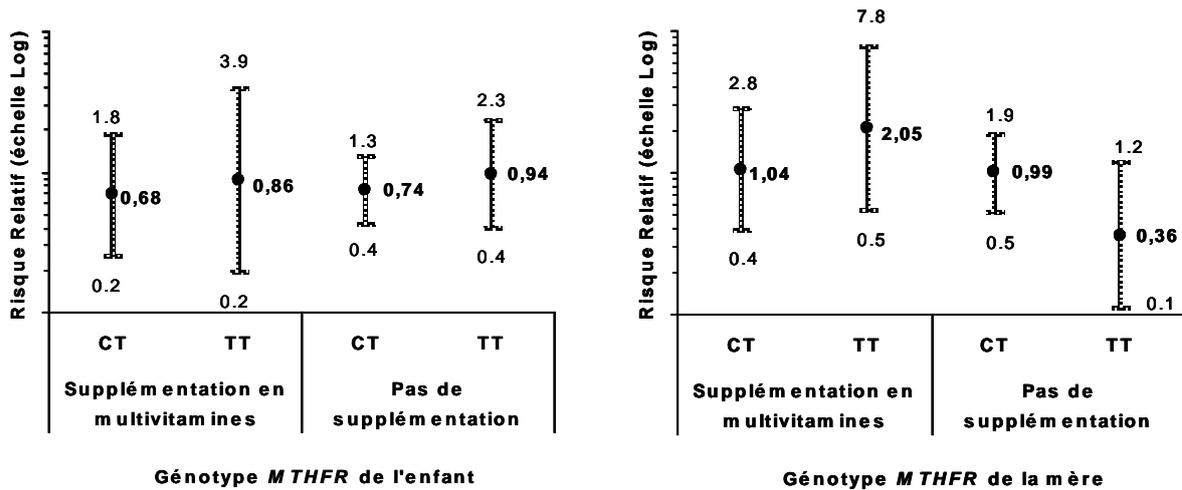
† Catégories de référence

Figure 7: Risques relatifs de FL/P des génotypes *MTHFR* CT et TT, comparé à CC, portés par l'enfant ou la mère, pour chaque strate d'apport en folates (alimentation ou vitamine), estimés par l'approche cas-parents (France, 1998-2001)

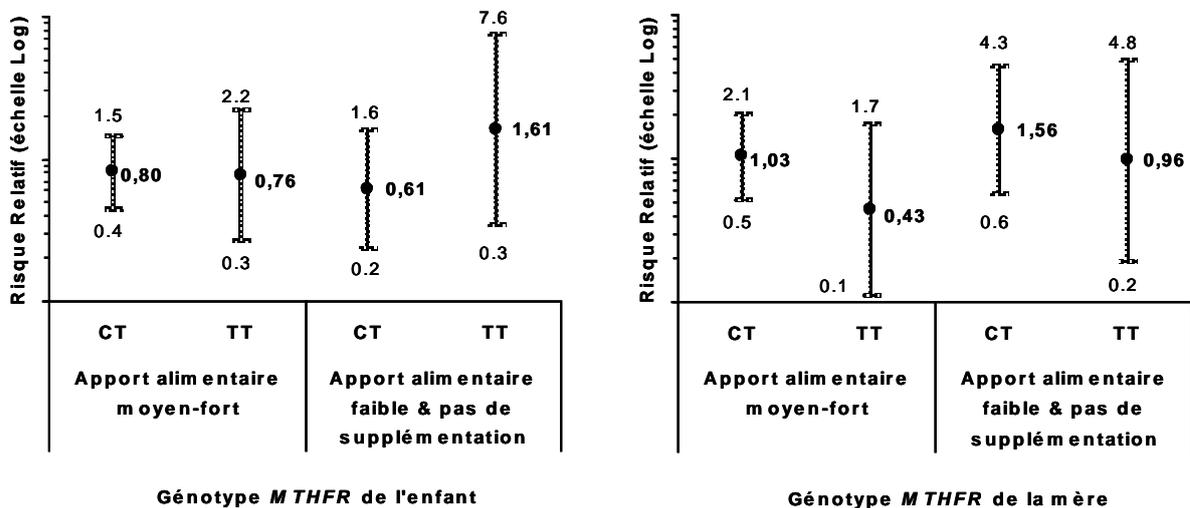
(1) Apport alimentaire de la mère en folates



(2) Supplémentation en multivitamines contenant de l'acide folique le mois précédant la conception OU le premier trimestre de grossesse



(3) Apport alimentaire de la mère en folates (exclusion des femmes supplémentées en multivitamines qui ont un faible apport alimentaire en folates)



## C. Discussion

L'approche cas-témoins suggère que l'allèle *T* du gène *MTHFR* protège contre le risque de FL/P et de FP, alors que l'approche cas-parents ne met pas en évidence de rôle protecteur de l'allèle *T* de façon convaincante selon le nombre de copies portées par l'enfant. Aucun rôle n'est suggéré pour le génotype de la mère. Aucun des deux schémas d'étude ne suggère de présence d'interaction entre le génotype *MTHFR* et l'apport en folates sur le risque de FL/P. Des analyses complémentaires de cette étude suggèrent que l'apport alimentaire en folates des mères joue un rôle dans la prévention des fentes orales (Rouget et al. 2005).

### ECHANTILLONNAGE « EPIDEMIOLOGIQUE ».

La répartition des cas et des témoins parmi les variables socio-démographiques, le choix de la sélection des cas dans des départements hospitaliers de chirurgie, les proportions attendues entre les groupes FL/P et FP et le sex-ratio observé pour chaque groupe ont été discutés précédemment dans la section consacrée à la description générale des populations de l'étude (partie [MATÉRIEL & Méthodes] Description de la population).

### APPORT EN FOLATES DE LA MÈRE.

L'apport alimentaire en folates de la mère est évalué à partir d'un questionnaire standardisé. Le niveau moyen de l'apport dans la population témoin de notre étude (270 µg/jour) est équivalent au niveau estimé dans une étude récente sur un échantillon représentatif des femmes françaises entre 35 et 60 ans (268 µg/jour) (Mennen et al. 2002). Le seuil séparant les apports « faibles » des apports « moyens-forts » en folate alimentaire a été calculé à 217µg/jour à partir du premier tertile observé sur le groupe témoin de notre étude. Cette valeur semble adéquate pour définir un statut de déficience en folates puisque le niveau minimal recommandé pour la population des femmes adultes en Europe est actuellement de 200µg/jour (de Bree et al. 1997). Notons que ce niveau est plus faible que le niveau recommandé en suppléments en acide folique de 400µg/jour, pour les femmes programmant une grossesse, afin de minimiser l'occurrence des ATN.

Deux définitions de l'utilisation vitaminique des mères sont proposées dans cette étude puisqu'il n'existe pas à ce jour de consensus établi sur cette définition en rapport avec l'évaluation des risques de fentes orales. Pour un développement embryonnaire optimal, les tissus qui nourrissent le fœtus les dix premières semaines de gestation (autres que le placenta) doivent être saturés en nutriments dès la conception ou l'implantation de l'œuf. Le statut en folates de ces tissus atteint un niveau optimal au bout d'approximativement 4 semaines. Pour ces raisons pharmacocinétiques, la deuxième option définissant une mère « utilisatrice de multivitamines » si elle est supplémentée à la fois sur la période préconceptionnelle et au premier trimestre de la grossesse, nous paraît la plus convaincante.

### ECHANTILLONNAGE « GENÉTIQUE » ET VALIDITÉ DES APPROCHES CAS-TÉMOIN ET CAS-PARENTS.

Une étude récente a examiné les variations de distribution des génotypes *MTHFR* à travers différents groupes ethniques et zones géographiques (Wilcken et al. 2003). En Europe, la fréquence du génotype *TT* s'observe entre 4 et 26% suivant un gradient nord-sud. Dans cette étude, 19% des enfants-témoins possèdent le génotype *TT* (dont 21% de génotype *TT* parmi les enfants-témoins des centres de Lyon et Grenoble, 20% pour le centre de Paris, 14% pour

Clermont-Ferrand). Cette fréquence est légèrement plus haute que celles rapportées en moyenne par des études précédentes sur la population française (Tableau 33).

**Tableau 33: Fréquence du génotype *MTHFR* TT rapportées dans la population française**

Référence	Population française	Fréquence du génotype TT
Faure-Delaneuf et al. 1997	Population témoin (n=374)	18.5%
Mornet et al. 1997	Population témoin (n=133)	10%
Thuillier et al. 1998	Population témoin Paris (n=107)	15%
Chango et al. 2000	Paris (étude SU.VI.MAX)	
	Total (n=291)	16.8%
	Femmes (n=161)	13.7%
	Hommes (n=130)	20.8%
Wilcken et al. 2003	Strasbourg (France) (n=178)	11.8%

La stratégie de recrutement des témoins de notre étude, imposant les critères du centre et de l'origine ethnique de la mère (neuf groupes) identiques aux enfant-cas, doit limiter les biais possibles de stratification de population. Cependant, nous ne pouvons pas exclure que l'association apparente du génotype *MTHFR* de l'enfant sur le risque de FL/P et de FP estimée par l'approche cas-témoins peut être liée à des résidus d'une stratification de population.

Finalement, puisque le schéma d'étude cas-parents est robuste face à ce biais éventuel, ses résultats nous paraissent plus fiables. Certaines études suggèrent que la présence de l'allèle *T* dans le génotype *MTHFR* de l'enfant ou de la mère accroît le risque d'avortement spontané. Cet effet est parfois observé en interaction avec d'autres gènes ou n'est pas reproduit dans d'autres études (revue dans Zetterberg 2004). L'existence possible de cette relation conduirait à une distorsion de l'hypothèse de transmission mendélienne parmi les naissances vivantes, hypothèse nécessaire à la validité des résultats de l'analyse cas-parents. Les génotypes du groupe d'enfants-témoins ne dévient pas statistiquement de l'équilibre d'Hardy-Weinberg (cf Annexe 4). Nous devons toutefois rester vigilants quant à l'interprétation des résultats de l'analyse cas-parents.

#### ASSOCIATION ENTRE *MTHFR* ET FENTES ORALES.

Le statut en folates pourrait modifier l'association entre le génotype *MTHFR* et le risque de maladie : une quantité « adéquate » en folates supprime les différences d'activité enzymatique des génotypes *MTHFR* (Bailey et Gregory 1999, Finnell et al. 2004). Ces hypothèses pourraient expliquer la variabilité des résultats observés parmi les études du risque de fentes orales associé au gène *MTHFR* à travers différentes populations d'étude (Gaspar et al. 1999, Shotelersuk et al. 2003, Vieira et al. 2005, revue pour les populations caucasiennes dans la partie [Introduction générale] L'étiologie complexe des fentes orales section B.2.v).

L'apport en folates par l'alimentation suit en Europe un gradient nord-sud, atteignant les plus forts niveaux dans la région méditerranéenne, incluant la France (de Bree et al. 1997). Aucune évidence d'un lien entre le gène *MTHFR* et le risque d'ATN n'est rapportée dans la population française (Mornet et al. 1997,  $n_{ATN}=43$ ,  $n_{témoins}=133$ ) ou dans la région sud de l'Europe de l'ouest (Guéant et al. 2003). L'ensemble de ces éléments concorde avec l'absence

d'association croissante, observée dans notre étude, entre le gène *MTHFR* et le risque de fentes orales non syndromiques.

Même si l'intérêt des études épidémiologiques sur le risque d'anomalies congénitales est principalement porté sur le génotype *TT* (Botto et al. 2000), certaines expérimentations *in vitro* montrent que la réduction de l'activité enzymatique de l'allèle *T* progresse avec le nombre de copies de l'allèle : comparé au génotype homozygote pour l'allèle *C*, le statut de l'hétérozygotie aurait un niveau d'activité enzymatique de 65%, et le statut d'homozygotie pour l'allèle *T* aurait un niveau de 30% (Sharp et Little 2004). Ainsi, la stratégie de regroupement des génotypes *CT* et *TT* permet d'offrir une meilleure puissance statistique, cependant l'interprétation de l'association observée par l'approche cas-parents sur le risque de FL/P est rendue difficile par l'absence de cohérence dans les risques associés aux génotypes *CT* et *TT*.

#### INTERACTION ENTRE *MTHFR* ET APPORT EN FOLATES.

Trois études épidémiologiques récentes ont examiné l'interaction entre l'apport en folates de la mère et le génotype *MTHFR* sur le risque de fentes orales. Il s'agit d'une étude cas-témoins sur une population californienne de 1987 à 1989 (Shaw et al. 1998, 1999), d'une étude cas-parents sur une population norvégienne de 1996 à 1998 (Jugessur et al. 2003) et d'une étude cas-témoins (accompagné d'un TDT) sur une population néerlandaise de 1998 à 2000 (van Rooij et al. 2003) (Tableau 34).

Shaw et al. (1998) et Jugessur et al. (2003) ne rapportent aucune évidence statistique d'interaction entre le génotype *MTHFR* et l'apport en folates sur le risque de FL/P. L'étude de van Rooij et al. (2003) observe que le risque de FL/P associé au génotype *TT* porté par la mère, comparé au génotype *CC*, est environ trois fois plus fort parmi les mères non supplémentées en acide folique que les mères supplémentées en période périconceptionnelle. Ils estiment un risque augmenté et significatif de FL/P chez les mères *TT* et n'utilisant pas de suppléments vitaminiques (OR=5.9, IC95% :1.1,30.9), comparé aux mères *CC* et supplémentées. Ce risque est plus élevé pour les mères *TT* n'utilisant pas de suppléments vitaminiques et avec un faible apport alimentaire en folates (OR=10.0, IC95% :1.3,79.1). Parmi les 88 triades, Jugessur et al. (2003) montre un risque de FP augmenté associé à l'allèle *T* porté par l'enfant, et de façon inattendue ce risque est significativement plus important sur la strate des mères supplémentées par rapport aux mères non supplémentées.

**Tableau 34: Résumé des risques de FL/P et de FP portés par les génotypes MTHFR sur chaque strate d'apport en folates, parmi les études épidémiologiques existantes incluant l'étude présente**

Génotype <i>MTHFR</i> C677T RR [IC95%] : TT versus CC	FL/P		FP	
	Enfant	Mère	Enfant	Mère
<b>Mère utilisatrice de suppléments en multivitamines contenant de l'acide folique</b>				
Shaw et al. 1998, 1999 <sup>1</sup> (cas-témoins)				
Oui	0.7 [0.4,1.4]		0.4 [0.2,1.1]	
Non	1.4 [0.5,3.6]		0.9 [0.2,3.3]	
Jugessur et al. 2003 <sup>1</sup> (cas-parents)			CT/TT versus CC	CT/TT versus CC
Oui	RR non reportés ;		4.3 [1.5,12.0]	0.8 [0.3,1.9]
Non	Interactions non significatives		1.4 [0.7,2.8]	1.4 [0.7,2.8]
van Rooij et al. 2003 <sup>2</sup> (cas-témoins)				
Oui	2.4[0.5,12.3]	1.2 [0.4,3.7]		
Non	2.1(=3.5/1.7)	3.5 (=5.9/1.7)		
Notre étude <sup>3</sup> (cas-parents)				
Oui	0.9 [0.2,3.9]	2.1 [0.5,7.8]		
Non	0.9 [0.4,2.3]	0.4 [0.1,1.2]		
<b>Apport alimentaire en folates</b>				
van Rooij et al. 2003 (cas-témoins)				
Niveau moyen-fort	X	1.7 [0.6,4.9]		
Niveau faible	1.3(=1.4/1.1)	1.6 (=2.8/1.8)		
Notre étude (cas-parents)				
Niveau moyen-fort	0.8 [0.3,2.2]	0.4 [0.1,1.7]		
Niveau faible	1.4 [0.4,4.5]	1.2 [0.4,3.9]		

RR : Risque Relatif ; IC : Intervalle de Confiance

<sup>1</sup> Oui : utilisation de suppléments vitaminiques le mois précédant la conception ou les deux mois après la conception

<sup>2</sup> Oui : utilisation de suppléments vitaminiques sur la période entière de 4 semaines avant la conception à 8 semaines après la conception

<sup>3</sup> Oui : utilisation de suppléments vitaminiques le mois précédant la conception ou le 1<sup>er</sup> trimestre de grossesse

nr : résultats non rapportés

**CONCLUSION.**

L'ensemble de ces études, incluant l'étude présente, rapporte des résultats contradictoires quant au rôle du gène *MTHFR* et à l'existence d'une interaction entre le génotype *MTHFR* et l'apport en folates de la mère dans l'étiologie des fentes orales. Ce résultat peut s'expliquer par un manque de puissance statistique ou par l'absence de considération de l'apport alimentaire moyen en folates pour certaines populations d'étude (Shaw et al. 1998, 1999, et Jugessur et al. 2003). Il peut aussi refléter la complexité de la relation entre le polymorphisme *MTHFR* C677T, l'apport maternel en folates, et le risque de fentes orales, pouvant possiblement faire intervenir d'autres polymorphismes du gène *MTHFR*, des gènes du transport des folates, des gènes impliqués dans le métabolisme de l'homocystéine, ou l'action d'autres nutriments tels les vitamines B6 et B12.

# 3

## L'exposition professionnelle de la mère aux solvants organiques en interaction avec des gènes de métabolisme/détoxication, sur le risque de fentes orales non syndromiques

Les solvants sont des substances largement répandues dans notre univers contemporain professionnel et domestique. Si certains sont connus pour être tératogènes chez l'animal, les effets sur la reproduction humaine sont encore incertains. Une expertise Inserm (1999) ciblée sur l'exposition aux éthers de glycol conclut à l'existence potentielle d'un lien entre l'infertilité masculine et l'exposition, tandis que cette même relation vis à vis des femmes est jugée non concluante. Certaines études épidémiologiques rapportent un excès de risque d'avortements spontanés ou de quelques anomalies congénitales, lié à l'exposition professionnelle de la mère aux solvants, mais ces résultats ne sont pas reproduits dans d'autres études. Un risque croissant de fentes orales est rapporté dans de nombreuses études, européennes pour la majorité. Il est associé, en particulier, à l'exposition professionnelle de la mère aux éthers de glycol. Deux autres études observent un risque augmenté de fentes orales associé au travail des mères dans le secteur de l'industrie du cuir et incriminent l'exposition à des solvants tels que les hydrocarbures aliphatiques et/ou le toluène. Ces études sont répertoriées dans le Tableau 35. Les conclusions de l'ensemble de ces résultats observés sont cependant rendus difficiles par la présence simultanée de nombreux solvants dans le métier des femmes, par l'absence d'observation de relation dose-effet pour la majorité des études, et par l'existence possible d'un biais favorisant la publication de résultats positifs.

A partir d'une évaluation de l'exposition individualisée, l'objectif de cette étude est d'examiner l'association entre l'exposition professionnelle de la mère aux solvants organiques et le risque de fentes orales non syndromiques. Diverses combinaisons d'expositions multiples aux solvants sont identifiées. Les relations dose-effet sont évaluées. Des interactions potentielles sont testées avec les polymorphismes codant pour des systèmes enzymatiques de métabolisme de xénobiotiques toxiques : de phase I, pour le cytochrome P450 *CYP2E1*\**RsaI* (nommé aussi *2E1*\*5 par la nomenclature de Garte et Crosti 1999) et le polymorphisme *ADH1C Ile349Val*, et de phase II pour les Glutathion-S-Transférases *GST-M1* et *GST-T1*.

**Tableau 35: Etudes épidémiologiques s'intéressant au risque de fentes orales associé à l'exposition professionnelle des mères aux solvants**

Référence	Population	Evaluation de l'exposition professionnelle	Expositions professionnelles évaluées dans l'étude	Fentes orales : OR [IC95%]
Laumon et al. 1996	France 1985-1989	Interview des mères : description de l'activité professionnelle (2 premiers mois de grossesse)	Halogénés aromatiques, Autres aromatiques, Halogénés aliphatiques, Autres Aliphatiques, Alcools, Glycols, Cétones, Aldéhydes, Esters	Halogénés aliphatiques : FL/P : 4.0 [1.2,14.9]
Cordier et al. 1997	6 registres européens (EUROCAT) 1989-1992	Interview des mères : description des professions et tâches professionnelles (1 <sup>er</sup> trimestre de grossesse) + Questionnaire supplémentaire pour quelques professions + Expertise chimique	314 substances chimiques (dont substances autres que solvants)	Ethers de glycol : FL/P : 2.0 [1.1,3.7] FP : 1.7 [0.7,3.8]
Lorente et al. 2000	Similaire à Cordier et al. 1997			Aldéhydes aliphatiques : FL/P : 2.1 [0.8,5.9] Ethers de glycol : FL/P : 1.7 [0.9,3.3] Trichloroéthylène : FP : 6.7[0.9,49.7]
Cordier et al. 2001	Slovaquie 1995-1996	Interview des mères par un médecin : description de la profession (occupée en début de grossesse) + Expertise chimiste	Ethers de glycol parmi 7 usages (produit-nettoyage, cosmétique, peinture-vernis, colorant-colle, encre, produit-dégraissant, pesticide, autres)	Ethers de glycol : 3.9 [0.6,25]
Shaw et al. 2003	Californie (Etats-Unis) 1987-1989	Interview des mères : description des professions (1 mois avant à 3 mois après la conception) + Expertise chimiste	74 groupes chimiques 9 groupes d'usage	(Résultats avec OR>2) Colorants organiques : FL/P : 2.7[0.9,7.7] Hydrocarbures aliphatiques : FP : 2.2[0.9,5.7] Cétones : FP : 2.3 [07,9.1]
Bianchi et al. 1997	Italie 1982-1989	Interview (non aveugle du statut cas ou témoin)	Personnel infirmier Coiffeur Travailleur dans la teinture des textiles Travailleur du cuir	Travailleur du cuir : Total : 3.7 [1.7,7.6] FL/P : 2.6 [0.7,7.3] FP : 5.4 [1.8,13.4]
García et Fletcher 1998	Espagne 1993-1994	Interview par téléphone	Travail dans l'industrie du cuir	6.2 [1.5,25.7]

OR:Odds-Ratio; IC:Intervalle de Confiance

## A. Matériel et Méthodes

Cette étude s'appuie sur les données d'une étude épidémiologique cas-témoins multicentrique (n=164 FL/P, n=76 FP, n=236 témoins) réalisée en France entre 1998 et 2001 (décrite dans la partie [MATERIEL & METHODES] Population).

### A. 1. Définition de l'exposition

L'évaluation des expositions professionnelles de la mère porte sur l'activité professionnelle occupée au premier trimestre de la grossesse (pour les treize femmes déclarant deux activités professionnelles en début de grossesse, nous considérons les plus forts niveaux d'exposition) et sur les tâches professionnelles décrites par les mères au moment de l'interview. Pour certaines catégories d'emplois (infirmière, aide-soignante, métier de ménage, gardienne d'immeuble, coiffeuse, esthéticienne, cuisinière, personnel de bar, enseignant, emploi dans l'industrie du cuir, travail de la céramique et des matières plastiques), un questionnaire supplémentaire aidait l'enquêteur à préciser les déclarations des mères.

A partir de l'ensemble de ces déclarations, un expert chimiste a évalué les expositions professionnelles de chaque mère, à l'aveugle du statut cas-témoins. L'expertise a attribué une note entre 0 et 2 pour trois paramètres de l'exposition : l'intensité, la fréquence et la probabilité de l'exposition. Une mesure unique de l'exposition est construite à partir de ces informations :

- 0 : pas d'exposition, si intensité, fréquence ou probabilité égales à 0,
- 1 : très faible exposition, si intensité x fréquence = 1 et probabilité = 1,
- 2 : faible exposition, si intensité x fréquence = 1 et probabilité = 2,
- 3 : moyenne ou forte exposition, si intensité x fréquence > 1 et probabilité = 2.

Les expositions professionnelles des mères sont évaluées pour trois classes chimiques de solvants organiques et leurs sous-groupes: les solvants oxygénés (alcools aliphatiques, aldéhydes aliphatiques, esters aliphatiques, cétones aliphatiques, éthers de glycol), les solvants chlorés (les alcanes type chlorure de méthylène, chloroforme, et les alcènes type trichloroéthylène et perchloroéthylène), des produits pétroliers (hydrocarbures aromatiques, essences minérales, fuel).

### A. 2. Analyse statistique

L'analyse des risques de fentes orales liés aux expositions professionnelles des mères aux solvants organiques est limitée au groupe des femmes ayant une activité professionnelle en début de grossesse. Une régression logistique est employée, ajustée sur le centre (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel ( $\leq 30$ ,  $> 30$  ans) et l'origine géographique de la mère (Europe, autres). Le groupe de référence est commun à toutes les expositions étudiées : il est constitué des femmes ayant une activité professionnelle en début de grossesse et exposées à aucun solvant (114 mères de témoins, 55 mères d'enfant atteint d'une FL/P, 28 mères d'enfant FP).

La tendance linéaire entre les niveaux d'exposition et le risque de fentes orales est évaluée par le test du rapport de vraisemblance appliqué à la variable déclarée en continu du niveau d'exposition (selon les codages attribués à chaque niveau, précisés plus hauts).

Les combinaisons les plus fréquentes d'exposition multiple à des solvants sont identifiées et examinées en relation avec le risque de fentes orales.

L'association entre les activités extra-professionnelles des mères, impliquant l'utilisation de produits contenant des solvants, et le risque de fentes orales est évaluée par une régression logistique similaire. Nous examinons l'utilisation des produits de teinture ou décoloration des cheveux et les activités de peinture par les mères au premier trimestre de la grossesse.

Les expositions des mères aux solvants sont examinées par des analyses complémentaires en interaction avec les polymorphismes du *CYP2E1\*Rsai*, *Ile349Val ADH1C*, *GSTM1* et *GSTT1*, impliqués dans le métabolisme de xénobiotiques toxiques. Parce que la majorité de ces polymorphismes sont rares et/ou les expositions peu fréquentes, les interactions gène-environnement sont évaluées sur l'ensemble du groupes de fentes orales non syndromiques, sans distinction des FL/P et des FP. Elles sont estimées par l'approche classique cas-témoins pour les polymorphismes des *GSTM1* et *GSTT1* (209 cas, et 165 témoins) pour lesquels nous nous intéressons à la délétion du gène à l'état homozygote (i.e. le statut hétérozygote n'est pas distingué du génotype homozygote-sauvage), et par l'approche cas-parents pour les polymorphismes *CYP2E1\*Rsai* (175 triades) et *ADH1C Ile349Val* (174 triades), pour lesquels le groupe d'enfants-témoins génotypés est modeste (via une régression logistique conditionnel ou un modèle log-linéaire si nous souhaitons évaluer l'effets du génotype de la mère). De plus, les associations entre chaque polymorphisme et le risque de fentes orales sont évaluées.

## B. Résultats

L'ensemble des variables socio-démographiques est décrit dans une section précédente (partie [MATÉRIEL & MÉTHODES] Description de la population) et présenté dans le Tableau 20 parmi les populations d'enfants cas et témoins.

### B. 1. Expositions maternelles aux solvants

Parmi les mères d'enfants témoins ayant une activité professionnelle au premier trimestre de grossesse, 5% sont exposées à un solvant pétrolier, 2% à un solvant chloré et 38% à un solvant oxygéné. Les proportions observées parmi les mères d'enfants-cas sont respectivement : 15%, 12% et 50% pour les FL/P, 5%, 5% et 49% pour les FP.

Le Tableau 36 montre que, pour chacune des trois classes chimiques de solvants organiques (oxygénés, chlorés ou pétroliers), l'exposition professionnelle des mères est associée à un risque augmenté de FL/P et de FP. Les excès de risque sont maintenus pour chacune des trois sous-classes de solvants oxygénés (alcools aliphatiques, éthers de glycol et autres solvants oxygénés). Ces associations sont statistiquement significatives pour le groupe de FL/P. Des tendances linéaires et croissantes sont observées entre le risque de FL/P et les niveaux d'exposition aux alcools aliphatiques ( $p=0.01$ ), aux éthers de glycol ( $p=0.007$ ), aux autres

solvants oxygénés incluant les aldéhydes, esters et cétones aliphatiques ( $p=0.02$ ), et aux produits pétroliers ( $p=0.004$ ). Lorsque les effectifs nous le permettent, des tendances croissantes similaires sont observées parmi le groupe de FP, mais de façon non significative. Un ajustement supplémentaire sur les activités extra-professionnelles des mères impliquant possiblement l'exposition à des solvants, ne modifie pas les conclusions.

Lorsque les mères exposées à des solvants chlorés ou pétroliers sont exclues de l'analyse ( $n=33$ ), les excès de risques liés à l'exposition à un solvant oxygéné sont maintenus mais la significativité statistique disparaît (Tableau 37). Parmi les 130 mères exposées à la seule classe des solvants oxygénés, 91% sont exposées à un alcool aliphatique, toujours associé avec, au moins, un autre solvant oxygéné. Nous isolons alors les mères exposées aux éthers de glycol (sachant que 76 parmi 78 sont exposées aussi à un alcool aliphatique) et les mères exposées aux autres types de solvants oxygénés (sachant que 42 parmi 52 sont exposées à un alcool aliphatique). Ces deux groupes de mères observent un risque de FL/P ou FP identique et légèrement augmenté. Parmi le groupe des mères exposées aux éthers de glycol, des relations croissantes sont observées entre les risques de FL/P et de FP, les niveaux d'exposition aux éthers de glycol et les niveaux d'exposition aux alcools aliphatiques. La relation entre le risque de fentes orales non syndromiques et les niveaux d'exposition aux éthers de glycol est :  $OR_{TrèsFaible}=0.74$ ,  $IC95\% : 0.3, 1.8$  ;  $OR_{Faible}= 1.47$ ,  $IC95\% : 0.7, 2.9$  ;  $OR_{MoyenFort}= 5.33$ ,  $IC95\%=1.1, 26.1$  ( $p$  tendance = 0.05).

L'utilisation extra-professionnelle par les mères de produits de teinture/décoloration pour les cheveux ou de peinture en début de grossesse semble accroître légèrement le risque de FL/P (Tableau 36). L'utilisation de produits pour les cheveux double significativement le risque de FP.

## B. 2. Interaction gène-environnement

### B.2.i. Interaction statistique

Les résultats des interactions entre les expositions maternelles et les susceptibilités génétiques de l'enfant sur les risques de FL/P et de FP sont présentés dans les tableaux 38 à 41. Les estimations impliquent des effectifs très modestes.

Pour le polymorphisme du *CYP2E1*\**RsaI* (Tableau 38), les estimations d'interaction entre le polymorphisme et les expositions des mères aux solvants sont de l'ordre de 2 et leurs intervalles de confiance sont larges. Par ailleurs, la présence simultanée du variant *CYP2E1*\*5 chez l'enfant et la consommation maternelle d'alcool multiplie par environ 6 le risque de fentes orales associé à l'effet conjoint calculé sous l'hypothèse multiplicative (produit des deux effets propres). Ce résultat est présenté en Annexe 5.

Aucune interaction entre le gène *ADH1C* de l'enfant et les expositions des mères aux solvants n'est suggérée (Tableau 39).

Les risques associés au génotype *Null* du *GSTM1* de l'enfant sont plus élevés parmi le groupe des mères exposées à des solvants organiques (via sa profession ou ses activités extra-professionnelles) que parmi les mères non exposées (Tableau 40). Aucune interaction entre le génotype *Null* du *GSTT1* de l'enfant et les expositions des mères aux solvants n'est suggérée (Tableau 41).

**B.2.ii. Rôles propres des gènes de métabolisme**

La fréquence de l'allèle variant *CYP2E1\*5* est approximativement de 2.9% parmi les enfants-cas, de 4.8% parmi leurs mères et de 4.4% parmi leurs pères. Sous l'hypothèse d'un mode dominant, nous observons un risque diminué de fentes orales non syndromiques associé à la présence de l'allèle variant *CYP2E1\*5* porté par l'enfant, parmi 175 triades et 21 diades (RR=0.26, IC95% :0.1,0.5 ;  $p_{LRT}<0.001$ ). L'allèle *CYP2E1\*5* porté par la mère ne semble jouer aucun rôle.

Un total de 56% des enfants-témoins porte la grande délétion (statut homozygote *Null*) pour le gène *GSTM1*, contre 55% parmi les enfants-cas. Pour le gène *GSTT1*, nous observons les proportions respectives suivantes : 18% parmi les enfants-témoins et 21% parmi les enfants-cas. Il n'existe aucune évidence d'association entre ces gènes et le risque de fentes orales (pour le gène *GSTM1*,  $OR_{Null}=0.93$ , IC95% :0.6,1.4,  $p_{LRT}=0.74$  ; pour le gène *GSTT1* :  $OR_{Null}=1.17$ , IC95% :0.7,2.0,  $p_{LRT}=0.56$ ).

L'estimation de ces associations restreinte aux familles d'origine parentale européenne aboutit aux mêmes conclusions. L'ensemble de ces résultats est présenté en Annexe 5. L'association entre le polymorphisme du gène *ADH1C* et le risque de fentes orales est présentée dans la première application du projet (cf partie [RESULTATS] Rôle d'un polymorphisme du gène *ADH1C* en interaction avec la consommation maternelle d'alcool sur le risque de fentes orales non syndromiques).

**B. 3. Tableaux****Tableau 36: Risques de FL/P et de FP associés à l'exposition professionnelle des mères aux solvants organiques ou à des activités extra-professionnelles impliquant possiblement la présence de solvants (France, 1998-2001)**

	Témoins		FL/P		OR*	IC 95%	FP			
	n	(%)	n	(%)			n	(%)	OR*	IC 95%
<b>Exposition professionnelle de la mère aux solvants organiques, parmi les femmes ayant une activité professionnelle au premier trimestre de la grossesse</b>										
Effectif total	187		116				57			
<b>Non exposées†</b>	<b>114</b>	<b>(61)</b>	<b>55</b>	<b>(47)</b>	<b>1.00</b>		<b>28</b>	<b>(49)</b>	<b>1.00</b>	
<b>Solvants oxygénés</b>	<b>71</b>	<b>(38)</b>	<b>58</b>	<b>(50)</b>	<b>1.81</b>	<b>1.1, 3.0</b>	<b>28</b>	<b>(49)</b>	<b>1.40</b>	<b>0.8, 2.6</b>
Alcools aliphatiques	61		49		1.84	1.1, 3.1	26		1.54	0.8, 2.9
Très faible-Faible	39		27		1.61	0.9, 3.0	13		1.20	0.6, 2.6
Moyen-Fort	22		22		2.22	1.1, 4.4	13		2.15	0.9, 4.9
<i>p</i> tendance						<i>p</i> =0.01				<i>p</i> =0.11
Ethers de glycol	41		37		1.98	1.1, 3.5	18		1.42	0.7, 3.0
Très faible	16		10		1.41	0.6, 3.4	2		0.38	0.1, 1.8
Faible	22		20		1.94	1.0, 4.0	12		1.70	0.7, 4.0
Moyen-Fort	3		7		4.78	1.2, 19.6	4		4.93	1.0, 23.9
<i>p</i> tendance						<i>p</i> <0.01				<i>p</i> =0.06
Aldéhydes, esters and cétones aliphatiques	56		46		1.78	1.1, 3.0	28		1.80	1.0, 3.4
Très faible	54		42		1.69	1.0, 2.9	28			
Faible-Moyen-Fort	2		4		5.12	0.9, 30.1	0			
<i>p</i> tendance						<i>p</i> =0.02				
<b>Solvants chlorés</b>	<b>3</b>	<b>(2)</b>	<b>14</b>	<b>(12)</b>	<b>10.47</b>	<b>2.8, 39.3</b>	<b>3</b>	<b>(5)</b>	<b>3.59</b>	<b>0.7, 19.6</b>
<b>Solvants pétroliers</b>	<b>10</b>	<b>(5)</b>	<b>17</b>	<b>(15)</b>	<b>3.65</b>	<b>1.8, 8.8</b>	<b>3</b>	<b>(5)</b>	<b>1.29</b>	<b>0.3, 5.2</b>
Très faible-Faible	7		10		3.01	1.0, 8.6	3			
Moyen-Fort	3		7		5.16	1.2, 21.4	0			
<i>p</i> tendance						<i>p</i> <0.01				
<b>Activités extra-professionnelles des mères au premier trimestre de grossesse, parmi l'ensemble des femmes</b>										
Effectif total	236		164				76			
Teinture/décoloration de cheveux										
Non†	217		145		1.00		63		1.00	
Oui	19	(8)	17	(10)	1.33	0.7, 2.6	12	(16)	2.30	1.0, 5.1
Activités de peinture										
Non†	223		151		1.00		71		1.00	
Oui	13	(6)	13	(8)	1.51	0.7, 3.4	4	(5)	1.01	0.3, 3.2

n: effectif ; % : pourcentage par rapport aux effectifs indiqués par « Effectif total » ; OR: Odds-Ratio ; IC: Intervalle de Confiance

\* Ajustement sur le centre (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel ( $\leq 30$ ,  $>30$ ), et l'origine géographique de la mère (Europe, autres)

† Catégorie de référence

*p* tendance: degré de signification du test du rapport de vraisemblance pour la variable du niveau d'exposition sous la forme quantitative

**Tableau 37: Risques de FL/P et de FP associés à l'exposition professionnelle des mères aux solvants oxygénés parmi les femmes ayant une activité professionnelle au premier trimestre de la grossesse, sachant que les femmes exposées aux solvants chlorés ou pétroliers sont exclues de l'analyse (France, 1998-2001)**

Exclusion des femmes exposées aux solvants chlorés ou pétroliers	Témoins		FL/P		FP		
	n	n	OR*	IC 95%	n	OR*	IC 95%
Non exposées†	114	55	1.00		28	1.00	
<b>Exposition aux solvants oxygénés<sup>1</sup></b>	<b>63</b>	<b>41</b>	<b>1.49</b>	<b>0.9, 2.5</b>	<b>26</b>	<b>1.45</b>	<b>0.8, 2.8</b>
Exposition aux éthers de glycol <sup>2a</sup>	38	24	1.40	0.7, 2.6	16	1.32	0.6, 2.9
Très faible	16	7	0.99	0.4, 2.6	2	0.37	0.1, 1.8
Faible	20	12	1.29	0.6, 2.9	11	1.66	0.7, 4.1
Moyen-Fort <sup>2b</sup>	2	5	4.84	0.9, 26.1	3	5.87	0.9, 38.7
p tendance				p=0.12			p=0.09
Exposition aux autres solvants oxygénés <sup>3</sup>	25	17	1.47	0.7, 3.0	10	1.50	0.6, 3.6

n: effectif ; OR: Odds-Ratio ; IC: Intervalle de Confiance

\* Ajustement sur le centre (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel ( $\leq 30$ ,  $>30$ ), et l'origine géographique de la mère (Europe, autres)

† Catégorie de référence : femmes ayant une activité professionnelle au premier trimestre de la grossesse et pour lesquelles l'expertise ne rapporte aucune exposition à des solvants organiques

p tendance: degré de signification du test du rapport de vraisemblance pour la variable du niveau d'exposition sous la forme quantitative

<sup>1</sup> 91% des femmes exposées aux solvants oxygénés (excluant les expositions simultanées aux solvants chlorés ou pétroliers) sont exposées aux alcools aliphatiques, toujours associés à au moins un autre solvant oxygéné

<sup>2a</sup> 76 femmes sur 78 sont aussi exposées aux alcools aliphatiques, et 57 d'entre elles sont de plus exposées à un troisième solvant oxygéné

<sup>2b</sup> 9 mères sur 10 sont aussi moyennement-fortement exposées aux alcools aliphatiques

<sup>3</sup> 42 femmes sur 52 sont exposées à la combinaison : alcools aliphatiques, aldéhydes, esters et cétones aliphatiques. 2 sont exposées seulement aux alcools aliphatiques, et 8 aux aldéhydes, esters et cétones aliphatiques.

**Tableau 38: Interaction entre le polymorphisme du *CYP2E1*\**RsaI* de l'enfant et l'exposition professionnelle et extra-professionnelles de la mère aux solvants sur le risque de fentes orales non syndromiques (France, 1998-2001)**

Polymorphisme de l'enfant du <i>CYP2E1</i> * <i>RsaI</i>	n triades* (impliqué dans le calcul de l'interaction)	Interaction entre l'exposition et la présence du variant chez l'enfant †		<i>p</i> interaction
		RR	IC 95%	
<b>Exposition professionnelle de la mère, parmi les mères ayant une activité professionnelle en début de grossesse</b>				
Tous solvants	5	1.9	0.2, 15.3	0.57
Solvants oxygénés	5	1.9	0.2, 15.3	0.57
Ethers de glycol	3	2.1	0.2, 22.3	0.55
Aldéhydes, esters et cétones aliphatiques	5	2.2	0.3, 18.8	0.47
Solvants chlorés	0	-	-	-
Solvants pétroliers	0	-	-	-
<b>Exposition extra-professionnelle des mères au premier trimestre, parmi l'ensemble des femmes</b>				
Teinture/Décoloration de cheveux	2	1.8	0.2, 16.6	0.62
Activités de peinture/vernis	0	-	-	-

\*n triades: nombre de familles exposées et dont l'enfant est porteur du variant *RsaI* ;

RR: Risque Relatif ; IC: Intervalle de Confiance

† Interaction sous l'échelle multiplicative entre le polymorphisme de l'enfant et l'exposition d'intérêt (=1 si l'hypothèse multiplicative est vérifiée (absence d'interaction), ≠ 1 si déviation de l'hypothèse multiplicative (présence d'interaction))

*p* interaction : test du rapport de vraisemblance de l'interaction (régression logistique conditionnelle)

**Tableau 39: Interaction entre le polymorphisme *Ile349Val ADH1C* de l'enfant et l'exposition professionnelle et extra-professionnelles de la mère aux solvants sur le risque de fentes orales non syndromiques (France, 1998-2001)**

Polymorphisme de l'enfant <i>Ile349Val ADH1C</i>	n triades* (impliqué dans le calcul de l'interaction)	Interaction entre l'exposition et les statuts hétérozygote (HZ) et homozygote (HM) pour l'allèle muté chez l'enfant†		p interaction
		RR	IC 95%	
<b>Exposition professionnelle de la mère, parmi les mères ayant une activité professionnelle en début de grossesse</b>				
Tous solvants	HZ 32	1.5	0.6,3.6	0.70
	HM 4	1.7	0.3,11.1	
Solvants oxygénés	HZ 31	1.5	0.6,3.7	0.70
	HM 4	1.7	0.3,11.1	
Ethers de glycol	HZ 19	1.5	0.5,4.4	0.69
	HM 2	0.8	0.1,11.5	
Aldéhydes, esters et cétones aliphatiques	HZ 27	1.6	0.6,4.1	0.62
	HM 3	1.6	0.2,13.0	
Solvants chlorés	HZ 7	1.7	0.4,7.6	-
	HM 1	-	-	
Solvants pétroliers	HZ 7	1.1	0.2,4.8	-
	HM 2	-	-	
<b>Exposition extra-professionnelle des mères au premier trimestre, parmi l'ensemble des femmes</b>				
Teinture/Décoloration de cheveux	HZ 11	3.3	0.8,13.2	0.22
	HM 1	1.7	0.1,2.3	
Activités de peinture/vernis	HZ 5	0.5	0.1,2.8	0.63
	HM 1	1.1	0.0,31.3	

\*n triades: nombre de familles exposées dont l'enfant est porteur du variant *349Val* de *ADH1C*;

RR: Risque Relatif ; IC: Intervalle de Confiance

† Interaction sous l'échelle multiplicative entre le polymorphisme de l'enfant et l'exposition d'intérêt (=1 si l'hypothèse multiplicative est vérifiée (absence d'interaction), ≠ 1 si déviation de l'hypothèse multiplicative (présence d'interaction))

p interaction : test du rapport de vraisemblance de l'interaction (régression logistique conditionnelle)

**Tableau 40: Risques de fentes orales non syndromiques associés à l'exposition professionnelle des mères aux solvants organiques ou à des activités extra-professionnelles impliquant possiblement la présence de solvants, stratifiés sur le génotype *GSTM1* de l'enfant (France, 1998-2001)**

Strate	Référence†	Exposition et <i>GSTM1-présent</i>	Non exposition et <i>GSTM1-Null</i>	Exposition et <i>GSTM1-Null</i>	<i>p</i> interaction
Effectif (cas, témoin) OR* ; IC 95%					
<b>Exposition professionnelle de la mère, parmi les mères ayant une activité professionnelle en début de grossesse</b>					
Tous solvants	(35,35) 1.00	(29,25) 1.21; 0.6,2.5	(40,48) 0.83; 0.4,1.6	(47,29) 1.70; 0.9,3.4	0.29
Solvants oxygénés	(35,35) 1.00	(27,24) 1.18; 0.6,2.5	(40,48) 0.83; 0.4,1.6	(45,29) 1.63; 0.8,3.2	0.31
Ethers de glycol	(35,35) 1.00	(15,14) 1.08; 0.4,2.6	(40,48) 0.85; 0.4,1.6	(45,29) 1.81; 0.8,3.9	0.24
Aldéhydes, esters et cétones aliphatiques	(35,35) 1.00	(24,20) 1.23; 0.6,2.6	(40,48) 0.86; 0.5,1.6	(39,23) 1.57; 0.8,3.2	0.45
Solvants chlorés	(35,35) 1.00	(6,2) 3.17; 0.6,17.2	(40,48) 0.89; 0.5,1.7	(8,1) 8.98; 1.1,76.2	0.40
Solvants pétroliers	(35,35) 1.00	(11,5) 2.16; 0.7,7.0	(40,48) 0.85; 0.4,1.6	(7,2) 4.94; 0.9,26.4	0.34
<b>Exposition extra-professionnelle des mères au premier trimestre, parmi l'ensemble des femmes</b>					
Teinture/Décoloration de cheveux	(86,66) 1.00	(9,6) 1.11; 0.4,3.3	(93,87) 0.81; 0.5,1.2	(18,6) 2.66; 1.0,7.3	0.15
Activités de peinture/vernis	(87,63) 1.00	(8,9) 0.68; 0.2,1.9	(106,91) 0.85; 0.6,1.3	(7,2) 2.60; 0.5,13.1	0.12

OR: Odds-Ratio ; IC: Intervalle de Confiance

\* Régression logistique ajustée sur le centre (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel ( $\leq 30$ ,  $> 30$ ), et l'origine géographique de la mère (Europe, autres)† Non Exposition et *GSTM1-présent*

**Tableau 41: Risques de fentes orales non syndromiques associés à l'exposition professionnelle des mères aux solvants organiques ou à des activités extra-professionnelles impliquant possiblement la présence de solvants, stratifiés sur le génotype *GSTT1* de l'enfant (France, 1998-2001)**

Strate	Référence†	Exposition et <i>GSTT1-présent</i>	Non exposition et <i>GSTT1-Null</i>	Exposition et <i>GSTT1-Null</i>	<i>p</i> interaction
Effectif (cas, témoin) OR* ; IC 95%					
<b>Exposition professionnelle de la mère, parmi les mères ayant une activité professionnelle en début de grossesse</b>					
Tous solvants	(60,69) 1.00	(58,43) 1.60 ; 0.9,2.8	(15,14) 1.25 ; 0.5,2.8	(18,11) 2.05 ; 0.9,4.8	0.97
Solvants oxygénés	(60,69) 1.00	(54,42) 1.54 ; 0.9,2.7	(15,14) 1.25 ; 0.5,2.8	(18,11) 2.05 ; 0.9,4.8	0.91
Ethers de glycol	(60,69) 1.00	(34,26) 1.51 ; 0.8,2.9	(15,14) 1.24 ; 0.5,2.8	(18,11) 2.71 ; 0.9,8.4	0.61
Aldéhydes, esters et cétones aliphatiques	(60,69) 1.00	(47,33) 1.67 ; 0.9,3.0	(15,14) 1.25 ; 0.5,2.8	(14,10) 1.73 ; 0.7,4.3	0.77
Solvants chlorés	(60,69) 1.00	(9,2) 6.17 ; 1.2,31.4	(15,14) 1.22 ; 0.5,2.8	(5,1) 5.87 ; 0.6,53.0	0.86
Solvants pétroliers	(60,69) 1.00	(13,6) 2.59 ; 0.9,7.6	(15,14) 1.25 ; 0.5,2.8	(6,1) 7.05 ; 0.8,61.4	0.54
<b>Exposition extra-professionnelle des mères au premier trimestre, parmi l'ensemble des femmes</b>					
Teinture/Décoloration de cheveux	(139,127) 1.00	(23,8) 2.89 ; 1.2,6.8	(39,26) 1.40 ; 0.8,2.4	(4,4) 0.93 ; 0.2,3.8	0.09
Activités de peinture/vernis	(154,124) 1.00	(9,11) 0.66 ; 0.3,1.6	(38,30) 1.00 ; 0.6,1.7	(6,0) -	-

OR: Odds-Ratio ; IC: Intervalle de Confiance

\* Régression logistique ajustée (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel ( $\leq 30$ ,  $> 30$ ), et l'origine géographique de la mère (Europe, autres)

† Non Exposition et *GSTT1-présent*

## C. Discussion

Cette étude renforce l'hypothèse d'un lien probable entre l'exposition professionnelle aux solvants organiques de la mère en début de grossesse et les risques de FL/P et de FP.

ECHANTILLONNAGE « EPIDEMIOLOGIQUE ».

La répartition des cas et des témoins parmi les variables socio-démographiques, le choix de la sélection des cas dans des départements hospitaliers de chirurgie, les proportions attendues entre les groupes FL/P et FP et le sex-ratio observé pour chaque groupe ont été discutés précédemment dans la section consacrée à la description générale des populations de l'étude (partie [MATÉRIEL & MÉTHODES] Description de la population).

MESURE DE L'EXPOSITION PROFESSIONNELLE.

La détermination des expositions professionnelles des mères aux solvants organiques s'appuie sur les déclarations maternelles concernant leur emploi occupé en période périconceptionnelle et sur une description des tâches professionnelles. A partir de ces déclarations, un expert chimiste a évalué les expositions de chaque emploi à des agents chimiques spécifiques. Cette expertise offre une évaluation standardisée et limite la présence possible de biais de mémoire. Des évaluations similaires ont déjà été menées par les études précédentes de Cordier et al. (1997), Lorente et al. (2000) et Shaw et al. (2003).

Nous observons dans la population des mères-témoins ayant une activité professionnelle en période périconceptionnelle, approximativement 22% de mères exposées aux éthers de glycol et 33% à un type d'alcool aliphatique. Cordier et al. (1997) et Lorente et al. (2000) rapportent pour une population européenne de 1989 à 1992 des fréquences moindres, avec respectivement 18% et 13%. La différence observée entre ces études, en particulier pour les alcools aliphatiques, pourrait s'expliquer par une répartition potentiellement différente des métiers des mères entre les études, ou par une plus faible sensibilité de l'expertise d'évaluation des expositions aux solvants dans l'étude européenne qui s'est intéressée à de nombreuses expositions chimiques professionnelles. Toutefois, chacune des études a évalué l'exposition à l'aveugle du statut cas-témoin.

EXPOSITIONS PROFESSIONNELLES AUX SOLVANTS ET RISQUE DE FENTES ORALES.

Les expositions professionnelles aux solvants chlorés ou pétroliers sont associées à des risques élevés de FL/P et de FP. Ces résultats concordent avec les résultats de quelques études épidémiologiques précédentes. Lorente et al. (2000) suggèrent des risques élevés de FP associés à une exposition des mères au trichloréthylène. Wennborg et al. 2005 rapportent un excès de risque d'anomalies craniofaciales et conotruncales associé à une exposition professionnelle des mères au benzène. Bove et al. (1995) montrent un lien entre l'ingestion de l'eau par les mères contaminée par les solvants organiques (tels que le tétrachloroéthylène, le dichloroéthane et le benzène) et la survenue de fentes orales.

Dans notre étude, une exposition à un solvant chloré se produit toujours en présence d'une autre exposition à un solvant pétrolier et/ou oxygéné. Seulement 3 mères sont exposées uniquement à la classe des solvants pétroliers. Les estimations s'appuient sur des effectifs faibles : seulement 20 mères sont exposées à un solvant chloré et 30 à un solvant pétrolier (17 mères sont exposées à la fois aux solvants chlorés et pétroliers).

L'interprétation de ces résultats doit être finalement menée avec prudence et, si un risque existe, il pourrait être attribué de préférence à la présence d'expositions multiples incluant des solvants chlorés et pétroliers.

Les études précédentes suggèrent un lien entre l'exposition professionnelle des mères aux solvants oxygénés et la survenue de fentes orales, mais cette exposition est rarement isolée des autres classes chimiques de solvants organiques (Cordier et al. 1997, Lorente et al. 2000, Cordier et al. 2001, Shaw et al. 2003). Les effectifs de notre étude ont permis d'examiner un groupe de mères exposées à des substances de la seule classe chimique des solvants oxygénés. Elles sont associées modérément (et de façon non significative) au risque d'avoir un enfant atteint d'une FL/P ou d'une FP. La majorité (91%) des mères de ce groupe est exposée à un type d'alcool aliphatique. A partir de 96 produits ou substances chimiques, Lorente et al. (2000) prennent en compte la simultanéité des expositions par la construction de modèles multivariés. Comparées aux estimations réalisées individuellement pour chaque exposition, la majorité des risques estimés par ces modèles perd la significativité statistique. Les auteurs observent finalement des risques élevés de FL/P associés à deux solvants oxygénés : les aldéhydes aliphatiques (OR=2.1, IC95% :0.8,5.9) et les éthers de glycol (OR=1.7, IC95% :0.9,3.3). Aucune association ( $p \leq 0.20$ ) n'est observée entre l'exposition maternelle à un solvant oxygéné et le risque de FP. Une relation dose-effet croissante est observée par Cordier et al. (1997) entre le risque de FL/P et les niveaux d'exposition aux éthers de glycol. Ce résultat est reproduit dans notre étude pour les deux groupes de FL/P et de FP, et parmi les mères qui ne sont exposées ni aux solvants chlorés, ni aux solvants pétroliers. Parmi les 10 femmes exposées fortement aux éthers de glycol, 9 sont aussi fortement exposées à un type d'alcool aliphatique. L'exposition aux éthers de glycol est finalement indissociable d'une exposition simultanée à un autre type de solvant oxygéné tel que les alcools aliphatiques ou les aldéhydes, esters et cétones aliphatiques.

#### ACTIVITES EXTRA-PROFESSIONNELLES DES MERES EN DEBUT DE GROSSESSE ET RISQUE DE FENTES ORALES.

Les résultats de cette étude associent, de façon statistiquement significative, l'utilisation de produits de teinture/décoloration de cheveux par la mère en début de grossesse, à un risque augmenté de FP. La mesure de cette exposition domestique est toutefois grossière : la fréquence d'utilisation des produits n'est pas disponible, les produits et ses composés ne sont pas identifiés. Parmi les 12 mères d'enfant atteint d'une FP et utilisant ces produits, quatre (33%) sont exposées à des solvants organiques dans leur univers professionnel (aucune n'est coiffeuse). Parmi les 19 mères d'enfant-témoin et utilisant ces produits, neuf (47%) sont exposées à des solvants organiques dans leur univers professionnel (dont 2 coiffeuses).

Malgré un effectif modeste, ce résultat mérite une attention particulière : les métiers de coiffure sont exposés à une variété d'agents chimiques potentiellement toxiques pour la reproduction et constituent une préoccupation majeure de certaines études épidémiologiques. Des études montrent que les coiffeuses ont des risques légèrement élevés d'avortements spontanés (Kersemackers et al. 1997), et d'avoir un enfant atteint d'un retard de croissance intra-utérin ou de malformations majeures (Rylander et al. 2002), en particulier de fentes palatines (Lorente et al. 2000).

#### TERATOGENICITE DES SOLVANTS.

De nombreux solvants ont été montrés tératogènes chez l'animal, de façon variable selon les modes d'exposition (inhalalé, ingéré ou cutané), les concentrations, les espèces animales et les substances testées (Schardein 1993). Notre étude ne permet pas d'identifier les agents

spécifiques potentiellement toxiques pour la reproduction (se préservant ainsi du problème de multiplicité des tests) et produit alors des estimations moyennes des risques associés à chaque classe chimique de solvants organiques. Les solvants chlorés, en particulier le trichloroéthylène et le dichloroéthylène, ont été associés chez l'animal à un risque augmenté de malformations cardiaques spécifiquement (Johnson et al. 2003). Certains solvants pétroliers, incluant le benzène, le toluène, le xylène, le fuel et quelques additifs pour carburant, sont identifiés tératogènes chez l'animal, et l'exposition à certains d'entre eux provoque la survenue de fentes palatines chez la souris (Schardein 1993). Parmi les solvants oxygénés, les éthers de glycol constituent une famille de solvants universellement répandue depuis plusieurs décennies, et particulièrement préoccupante en terme de conséquences sanitaires sur la reproduction et le développement : certains éthers de glycol ont été répertoriés possiblement tératogènes chez l'homme et agissent sur le développement de l'enfant sans affecter l'état de santé de la mère (Expertise Inserm 1999). Notre étude ne distingue pas les différentes familles d'éthers de glycol. Quelques alcools aliphatiques sont montrés tératogènes chez l'animal, suite à une exposition par inhalation (c'est-à-dire pour les alcools capables de générer des vapeurs) (Schardein 1993).

#### PROFESSIONS EXPOSEES AUX SOLVANTS ORGANIQUES.

Dans notre étude, une exposition simultanée aux trois classes chimiques de solvants organiques a été attribuée aux professions maternelles suivantes : ingénieur ou technicien de laboratoire, employé de bureau ou caissier ou secrétaire, commerçant ou vendeur, esthéticien, ouvrier du cuir, autres ouvriers industriels ou empaqueteur. L'exposition aux solvants oxygénés concerne principalement les métiers de ménage et d'infirmière ou d'aide-soignante. Une exposition aux éthers de glycol est rapportée essentiellement chez les femmes de ménage et la catégorie des commerçants-vendeurs. La distribution des professions pour chaque exposition est présentée dans le tableau (Tableau 42).

**Tableau 42:Professions des mères exposées aux solvants organiques (France, 1998-2001)**

Exposition (% colonne)	Solvants oxygénés				Solvants chlorés	Solvants pétroliers
	Alcools aliphatiques	Ethers de glycol	Aldehy., est.,cét. aliph.*	Au moins un solvant oxygéné		
<b>Profession</b>	n=136	n=96	n=130	n=157	n=20	n=30
Médecin, dentiste, vétérinaire, pharmacien	10.3	0	10.8	8.9	0	0
Infirmière, aide-soignant	19.1	4.2	20.0	16.6	0	0
Personnel enseignant et animateur	2.9	5.2	5.4	5.1	0	3.3
Ingénieur, technicien de laboratoire	4.4	5.2	6.9	7.0	10.0	16.7
Employé de bureau, caissier et secrétaire	9.6	9.4	3.8	9.6	15.0	16.7
Commerçant et vendeur	11.0	16.7	7.7	10.2	5.0	13.3
Femme de ménage	19.9	27.1	20.8	17.2	0	3.3
Nourrice	3.7	5.2	0	3.2	0	0
Autres employés de maison	2.2	3.1	2.3	1.9	0	0
Coiffeur	3.7	5.2	3.8	3.2	5.0	0
Serveur, cuisinier	5.1	7.3	5.4	4.5	0	0
Esthéticien	2.2	3.1	2.3	1.9	15.0	13.3
Ouvrier du cuir	0	0	0.8	0.6	5.0	3.3
Autres ouvriers industriels, empaqueteur	5.9	8.3	7.7	8.3	45.0	30.0
Ouvrier agricole	0	0	2.3	1.9	0	0

\* Aldéhyde, esters et cétones aliphatiques  
Tableau des effectifs en Annexe 6

## CONCLUSION.

Finalment, même si l'évaluation de ces associations peut être influencée par d'autres facteurs de confusion chimiques potentiels (tels que des expositions professionnelles ou domestiques à d'autres agents chimiques des mères, les conditions de travail des mères, l'histoire maternelle des expositions professionnelles), cette étude apporte des preuves solides pour conclure à la tératogénicité potentielle des solvants dans les environnements de l'homme professionnel et domestique. Les solvants sont largement et trop communément utilisés, et ce fait nécessite de sérieuses considérations puisqu'une exposition courte pendant la période critique de l'embryogenèse pourrait apparaître suffisante pour engendrer une anomalie de reproduction.

Parce que de nombreux solvants sont souvent présents simultanément dans nos univers professionnels et domestiques, l'identification d'un agent responsable d'une issue de grossesse défavorable est difficile. Certaines variations dans le métabolisme de telles substances chimiques durant la vie embryonnaire peuvent modifier l'association entre l'exposition aux solvants et le risque de fentes orales, et ainsi aider à identifier certaines voies métaboliques spécifiques.

## ANALYSE COMPLEMENTAIRE : INTERACTION GENE-ENVIRONNEMENT.

Cette étude examine les effets potentiels de susceptibilité génétique de l'enfant pouvant altérer le développement de l'embryon à travers un métabolisme modifié d'agents toxiques exogènes. Les systèmes enzymatiques du *CYP2E1* et de l'*ADH1C* peuvent transformer une substance en un métabolite actif et possiblement toxique (Phase I d'activation), puis les enzymes de Phase II, telles que les GST, convertissent leur substrat en un métabolite moins toxique que la substance initiale (Phase de détoxification).

Le gène *CYP2E1* possède un grand nombre de substrats qui en majorité sont caractérisés par un petit poids moléculaire (éthanol, aldéhydes, cétones, composés aromatiques, éthers, alcanes et alcènes, benzène, éthylène glycol, styrène, trichloroéthylène, etc (Wormhoudt 1998, Smith et al. 1995)). Le polymorphisme défini par l'enzyme de restriction *RsaI* est fonctionnel et se caractérise par une altération de l'activité transcriptionnelle du gène (Smith et al. 1995). Le fœtus humain possède une capacité métabolique des xénobiotiques bien développée à travers le système enzymatique des cytochromes P450. Cependant, l'activité enzymatique embryonnaire est qualitativement et quantitativement différente de celle de l'âge adulte, et elle est encore très controversée dans la littérature (Hakkola et al. 1998). Certaines études montrent une expression et une activité métabolique du *CYP2E1* à partir de la 16<sup>ème</sup> semaine de gestation seulement (Carpenter et al. 1996, Hakkola et al. 1998).

Les résultats de cette étude ne suggèrent aucune interaction entre le polymorphisme *CYP2E1*\**RsaI* de l'enfant et l'exposition professionnelle des mères aux solvants dans la survenue de fentes orales. Cette interaction n'a encore jamais été examinée dans le domaine des anomalies de la reproduction.

Un résultat complémentaire montre que l'interaction entre la consommation maternelle d'alcool et le polymorphisme *CYP2E1*\**RsaI* s'approche de la significativité statistique. Ce résultat implique un faible effectif et peut être un résultat du hasard. L'étude de Carpenter et al. (1996) montrent une induction du gène humain fœtal *CYP2E1* par l'éthanol, suggérant un rôle potentiel du gène dans la survenue d'anomalies de grossesse liées à la consommation maternelle d'alcool (à partir du second trimestre de grossesse). Hakkola et al. (1998) suggèrent qu'une expression élevée du système enzymatique du *CYP2E1* observée dans les

tissus placentaires après une consommation maternelle d'alcool pourrait indiquer un rôle du système enzymatique propre au placenta dans le métabolisme de xénobiotiques toxiques.

Enfin, les résultats de ce projet observent une association protectrice entre l'allèle *CYP2E1*\*5 porté par l'enfant et le risque de fentes orales non syndromiques. Aucune évidence d'un rôle de l'allèle maternel *CYP2E1*\*5 n'est observée dans cette étude (cf Annexe 5). De nombreux substrats connus au gène *CYP2E1*, autres que l'éthanol et certains solvants organiques, pourraient être considérés pour expliquer ce résultat. Cependant, l'activité catalytique du polymorphisme *CYP2E1*\**RsaI* lors de la vie embryonnaire n'étant pas clairement identifiée, le résultat de cette étude s'explique difficilement. De même, les associations entre le gène *CYP2E1* et les risques de pathologie cancéreuse sont encore très mal déterminées (Stubbins and Wolf 1999).

L'ADH est connue pour être impliquée dans le métabolisme de l'éthanol et de certains solvants, en particulier les éthers de glycol qui possèdent une fonction alcool primaire (Multigner et al. 2005). L'activité embryonnaire du gène *ADH1C* n'est pas encore bien établie (cf partie [RESULTATS] Rôle d'un polymorphisme du gène *ADH1C* en interaction avec la consommation maternelle d'alcool sur le risque de fentes orales non syndromiques). Les résultats de cette étude ne suggèrent aucune interaction entre le polymorphisme *Ile349Val* *ADH1C* de l'enfant et l'exposition des mères aux solvants dans la survenue de fentes orales.

Les *GST* jouent un rôle majeur dans la détoxication d'une grande variété de substances mutagènes, exogènes ou endogènes (Smith et al. 1994). La délétion observée à l'état homozygote, *Null*, des gènes *GSTM1* et *GSTT1* implique une absence d'activité de détoxication, et elle est généralement attendue à risque en présence d'exposition à des xénobiotiques toxiques. Les fréquences de génotype *Null* observées dans cette étude pour les deux gènes sont concordantes avec celles observées dans la littérature parmi les populations européenne ou caucasienne (Wormhoudt 1998, Garte et al. 2001). L'activité enzymatique des *GST* a été très étudiée dans les tissus humains adultes, mais cette activité dans les tissus de l'embryon ou du fœtus est encore très peu connue (Raijmakers et al. 2001). Lors de la vie embryonnaire, le gène *GSTM1* aurait une expression faible et essentiellement localisée dans le foie, tandis que l'activité enzymatique issue du gène *GSTT1* est encore indéterminée (Raijmakers et al. 2001).

Shaw et al. (2003) suggèrent des risques élevés de FL/P ou de FP avec la présence conjointe du génotype *GSTM1*\**Null* ou *GSTT1*\**Null* de l'enfant et de l'exposition maternelle à certains solvants organiques et autres substances chimiques, mais ne rapportent aucun test d'interaction. L'ensemble des résultats d'interaction observés dans notre étude et celle de Shaw et al. (2003) implique des effectifs très modestes et ne peut être que suggestifs.



# **CONCLUSION GENERALE**



# Résumé des résultats

Les principaux biais de l'enquête épidémiologique sur laquelle s'appuie ce projet, ont été discutés dans la partie [MATERIEL & METHODES] Description de la population, section B. La principale difficulté de cette enquête réside dans la forte proportion d'informations génétiques absentes parmi le groupe des témoins. La raison majeure de cette absence différentielle d'information s'explique par un support de prélèvement sanguin spécifiquement utilisé pour le groupe de témoins à des fins de dépistage, et inadapté aux techniques classiques de génotypage. La solution adoptée dans ce projet a été de compléter la détermination des polymorphismes par des techniques plus pointues (sous condition de la validation d'un échantillon interne commun aux laboratoires de génotypage), sous peine de renoncer à l'analyse cas-témoins.

Nous rappelons que cette étude n'inclut pas les fentes orales diagnostiquées pendant la grossesse et qui ont abouti à un avortement. La restriction aux cas prévalents à la naissance pourrait conduire à sous-évaluer la mesure de certaines associations.

Enfin, bien que les FL/P et les FP soient considérées comme étiologiquement distinctes, nous avons choisi pour certaines analyses de rassembler ces deux groupes lorsqu'ils présentaient des estimations similaires, pour assurer une meilleure puissance statistique. Par conséquent, parce que le rôle de facteurs environnementaux (tels que la consommation maternelle d'alcool) n'est pas bien établi dans l'étiologie des fentes orales syndromiques, l'exclusion des formes syndromiques ne semblait pas justifiée pour garantir un groupe homogène. Nous avons donc choisi de présenter des résultats supplémentaires incluant les formes syndromiques.

La présence d'un biais potentiel de stratification de population est discutée dans la partie suivante. Les autres biais de l'étude sont mineurs ou corrigés par la stratégie d'analyse.

Trois applications sont présentées dans ce travail intégrant des facteurs environnementaux déjà soupçonnés dans la littérature et des gènes impliqués dans les voies de métabolisme/détoxication de ces expositions environnementales : (1) un polymorphisme du gène *ADH1C* (alcool déshydrogénase, site Ile349Val) en interaction avec la consommation maternelle d'alcool, (2) le polymorphisme commun *C677T* du gène *MTHFR* en interaction avec l'apport en folates de la mère, et (3) l'exposition professionnelle de la mère aux solvants organiques en interaction avec les polymorphismes d'un cytochrome P450 (*2E1\*Rsai*), de l'*ADH1C*, et de gènes de glutathion-s-transférase (*M1*) et (*T1*).

Les résultats de la première analyse renforcent l'hypothèse d'un lien entre la consommation maternelle d'alcool en début de grossesse et le risque de fentes orales non syndromiques, et en particulier de FL/P. Cette étude ne met pas en évidence d'accroissement du risque avec la dose ingérée. Cette première application rapporte un résultat nouveau : elle suggère l'influence possible du polymorphisme de l'*ADH1C* porté par l'enfant, indépendamment de l'exposition à l'alcool (aucune interaction gène-environnement statistiquement significative n'est observée). Parce que le gène est impliqué dans l'activité métabolique d'autres substances (telles que le rétinol et certains éthers de glycol), le rôle possible de ces expositions doit être considéré dans l'étiologie des fentes orales.

La seconde application apporte des résultats conflictuels quant au rôle du polymorphisme *C677T* du gène *MTHFR* sur le risque de fentes orales non syndromiques. Sous réserve de puissance statistique, aucune évidence d'interaction entre le polymorphisme *C677T* du gène *MTHFR* et l'apport en folates de la mère par la prise de multivitamines ou par l'alimentation n'est observée. Certaines expérimentations montrent pourtant que l'apport en folates influe sur l'activité enzymatique du polymorphisme. La littérature épidémiologique concernant le rôle du polymorphisme *MTHFR C677T*, quel que soit l'apport en folates de la mère, dans l'étiologie des fentes orales, est encore peu développée. Les résultats de ces études, incluant l'étude présente, sont conflictuels et reflètent un mécanisme complexe intervenant entre le gène *MTHFR*, l'apport en folates et le risque de fentes orales. D'autres analyses de ce projet montrent le rôle des apports alimentaires en folates dans la prévention des fentes orales.

La troisième application participe à une littérature épidémiologique très débattue concernant l'exposition professionnelle des mères aux solvants organiques et le risque de malformations congénitales, en particulier de fentes orales. Elle rapporte une association entre l'exposition professionnelle aux solvants et le risque de fentes orales, bien qu'il soit difficile d'attribuer ce risque à un agent spécifique dû aux expositions multiples à différents solvants. Un risque protecteur est observé entre l'allèle muté du *CYP2E1 \*RsaI* porté par l'enfant et le risque de fentes orales non syndromiques. Aucune interaction gène-environnement statistiquement significative n'est observée. De façon générale, peu d'études se sont intéressées aux interactions entre l'exposition professionnelle des mères à des substances chimiques et des gènes impliqués dans le métabolisme de xénobiotiques toxiques, dans la survenue d'une fente orale ou, plus généralement, de malformation congénitale. Ces quelques études, incluant l'étude présente, fournissent des estimations d'interaction très peu puissantes, dûes à une prévalence faible de l'exposition et des polymorphismes d'étude rare.

Les perspectives d'application de ce projet sont encore nombreuses et permettront de contribuer à l'amélioration des connaissances des interactions gène-environnement dans l'étiologie des fentes orales. Nous nous intéresserons à la consommation maternelle de tabac en interaction avec les polymorphismes de *GST* et de *CYP1A1*. Les rôles des polymorphismes du développement embryonnaire déjà soupçonnés dans la littérature (tels que *TGF $\alpha$* , *TGF $\beta$ 3* et *MSX1*) seront évalués, seuls et en interaction avec la consommation maternelle de tabac et l'apport alimentaire en folates des mères pendant la grossesse.

# Discussion du biais potentiel de stratification de population

Les résultats des analyses cas-témoins et cas-parents sont conflictuels dans la seconde application. Les puissances statistiques des tests d'évaluation du rôle du polymorphisme *MTHFR C677T* sur le risque de FL/P sont comparables selon les deux approches (Les calculs *a posteriori* sont présentés en Annexe 7). Même si la population française, et plus généralement, la population européenne caucasienne, est identifiée comme une population génétiquement homogène, la forte variabilité géographique de la fréquence du polymorphisme *MTHFR C677T* semble pouvoir expliquer ces résultats contradictoires. Cela nous a conduits par conséquent à ne pas exclure la possibilité d'une présence d'un biais résiduel de stratification de population dans nos analyses.

Nous rappelons qu'un biais de stratification de population peut affecter la mesure des associations des études cas-témoins lorsque le risque de maladie et la fréquence du génotype d'intérêt diffèrent dans plusieurs sous-populations de l'étude. Ce biais peut être réduit par un appariement ou un ajustement des analyses sur des critères ethniques, cependant il n'est pas toujours possible de l'éliminer si la structure de la population n'est pas clairement identifiée par ces critères. A partir de travaux réalisés sur diverses populations, ce biais paraît négligeable lorsque les principes épidémiologiques de schéma d'étude, de protocole et d'analyse sont rigoureusement appliqués.

La présence d'une stratification de population est probablement la raison la plus citée pour expliquer la non-reproduction de résultats d'association génétique entre les études, ou entre les schémas d'étude. Cependant, cette raison est évoquée de façon plus systématique qu'elle n'est justifiée et argumentée. Les résultats de la seconde application de ce projet nous amène à nous poser les deux questions suivantes :

- (1) La contradiction des résultats entre les deux schémas d'étude, cas-témoins et cas-parents, reflète-t-elle nécessairement la présence d'une stratification de population ?
- (2) La variabilité observée des fréquences alléliques du polymorphisme *MTHFR C677T* est-elle suffisante pour engendrer une stratification de population ?

(1) L'approche cas-parents s'appuie sur un groupe de témoins fictifs de population génétiquement identique à celle des enfants-cas. Elle est ainsi robuste au biais de stratification de population. Elle est considérée par certains comme la solution idéale à adopter pour la mesure d'une association génétique. Un éditorial du journal *Nature Genetics* en 1999 présente, dans ce sens, la position du journal face à la publication de résultats d'association génétique : privilégier les études présentant et reproduisant leur résultat par le schéma d'étude cas-témoins et un schéma d'étude basé sur des données familiales (par exemple, cas-parents).

Le schéma d'étude basé sur des données familiales est, cependant, sensible à d'autres biais tels que la distorsion à l'hypothèse de transmission mendélienne pour une approche cas-parents. La contradiction dans les résultats des deux approches n'implique pas nécessairement la présence d'un biais de stratification de population dans l'étude cas-témoins, elle doit être discutée aussi selon les hypothèses requises par l'utilisation du schéma d'étude cas-parents.

Dans notre étude qui inclut seulement les cas prévalents à la naissance, la validité de cette hypothèse pourrait être compromise par le lien possible entre l'allèle *T* et un risque augmenté de mort fœtale. L'association entre le polymorphisme *MTHFR C677T* et le risque de fentes orales estimée par l'approche cas-parents pourrait être alors sous-évaluée.

(2) De façon générale, la stratification de population se produit si la variation des fréquences alléliques est corrélée à la variation des risques de base de la maladie (non attribuable au gène candidat) dans les sous-populations. La prévalence des fentes orales varie entre les pays européens mais aussi à l'intérieur des pays. En France, les trois registres de malformations congénitales rapportent une variabilité géographique évidente de la prévalence des fentes orales, en particulier de FL/P. Les variabilités observées dans cette étude, de la fréquence allélique *MTHFR 677T* et de la prévalence des fentes orales sont modérées. Si une stratification de population existe dans notre étude, l'effet devrait être modeste.

Le recours à des témoins génomiques ou un ajustement des analyses statistiques sur les critères ethniques des populations sont quelques solutions pour détecter et éliminer les biais de stratification de population. Cependant, certains auteurs défendent l'idée que l'ethnicité n'est pas seulement une fonction de gènes, mais que les critères ethniques, déterminants d'un risque de maladie, sont probablement mieux expliqués par de multiples facteurs tels que les modes de vie, les facteurs sociaux, les pratiques culturelles, linguistiques et religieuses. Cette définition d'appartenance ethnique est, en pratique, difficile à collecter dans une étude épidémiologique, et elle dépend probablement de plusieurs mesures.

L'appariement et l'ajustement réalisé dans notre étude à la fois sur le centre de recrutement et sur l'origine géographique des mères (France, Europe du Nord, Europe du Sud, Europe de l'Est, Afrique du nord, Autre Afrique, Antilles, Asie, Autres) pourraient alors ne pas suffire pour éliminer un biais potentiel de stratification de population.

# Puissance statistique des tests d'interactions gène-environnement

La puissance statistique est modeste dans la majorité des études d'interaction gène-environnement sur le risque de maladie d'origine multifactorielle, en particulier lorsque le polymorphisme est rare et/ou la fréquence de l'exposition est modeste. Sensibles à ce problème majeur des études d'interaction gène-environnement, certains auteurs ont développé des outils de calcul de puissance statistique (et de taille d'échantillon), appliqués aux études épidémiologiques, disponibles gratuitement et faciles à mettre en place. Nous devons noter que le calcul de puissance *a posteriori* est considéré par certains auteurs comme non informatif (Goodman et Berlin, 1994).

## CALCUL DE LA PUISSANCE STATISTIQUE *A POSTERIORI* DANS LE PROJET.

Différentes méthodologies de calcul existent, nous choisissons d'appliquer la méthode proposée par Gauderman (2002) aux trois applications de ce projet. Nous utilisons le logiciel QUANTO Version 1.0 (<http://hydra.usc.edu.gxe>). Il nous permet de plus de calculer les puissances statistiques des tests d'association pour chaque facteur environnemental et chaque gène. Nous présentons ces résultats en Annexe 7.

Les résultats de calculs de la puissance des tests d'interaction sont exposés dans le Tableau 43. Ils indiquent, pour chaque application, un manque évident de puissance statistique des tests de détection d'une interaction gène-environnement, lorsqu'elle est modeste. Par conséquent, il est possible que ce projet présente des résultats faux-négatifs d'interaction statistique gène-environnement, caractéristiques des études sous-puissantes. La majorité des applications de ce projet dispose cependant d'une puissance statistique de détection d'une interaction gène-environnement convenable lorsque l'interaction à détecter est forte (supérieure à 4).

D'un point de vue de l'interprétation, les résultats des tests d'interaction de ce projet semblent particulièrement incertains lorsque les polymorphismes d'intérêt, impliqués dans le métabolisme de facteur environnemental, sont possiblement associés au risque de fentes orales, quelque soit le facteur environnemental. Il s'agit des polymorphismes du gène *ADH1C* et du gène *CYP2E1*. Les effets propres de ces polymorphismes détectent alors la présence d'une interaction biologique gène-environnement. L'identification du mécanisme biologique sous-jacent est difficile, en particulier lorsque le gène d'intérêt est pléiotropique et qu'aucune interaction statistique entre le polymorphisme étudié de ce gène et ses substrats environnementaux n'est détectée.

**Tableau 43: Puissance statistique calculée *a posteriori* d'un test d'interaction gène-environnement, pour les trois applications du projet**

Paramètres requis par QUANTO	ADH1C et Alcool	MTHFR et Folates alimentaires (pour les FL/P)	CYP2E1 et Solvants	GSTM1 et Solvants	
Schéma d'étude*	Cas-Parents	Cas-témoins (non appariés)	Cas-Parents	Cas-Parents	Cas-témoins (non appariés)
Effectif	174 triades	146 témoins	124 triades	127 triades	136 témoins
Mode de transmission de l'allèle†	Dominant	Récessif	Récessif	Dominant	Récessif
Fréquence de l'allèle d'intérêt dans la population‡	0.40	0.44	0.44	0.05	0.75§
Prévalence de l'exposition dans la population‡	0.29	0.33	0.33	0.40#	0.40#
Risque de base de la maladie dans la population	0.0014	0.0008	0.0008	0.0014	0.0014
Risque relatif de maladie associé à l'exposition, parmi les non porteurs de l'allèle	1.5	2.1	2.1	1.5	1.5
Risque relatif de maladie associé à l'allèle d'intérêt, parmi les non exposés	0.65	0.5	0.5	0.5	0.8
<b>Puissance statistique estimée pour détecter une interaction gène-environnement de :</b>					
2	38%	19%	19%	12%	27%
3	71%	41%	42%	24%	52%
4	86%	60%	61%	35%	67%
Estimation de l'interaction dans l'application	1.6	1.2	1.8	≈2	≈2

Note : Compte-tenu des contraintes de QUANTO, les valeurs choisies des paramètres sont approximées par rapport aux valeurs observées dans chaque application. Nous considérons ici le gène de l'enfant. Le risque d'erreur de première espèce est fixé à 0.05.

\* Pour un schéma d'étude cas-témoins non apparié, le ratio cas :témoin est 1 :1

† Dominant, récessif ou hypothèse multiplicative

‡ Calculée à partir de la population de témoins

§ La fréquence du génotype homozygote Null est de 0.56 dans la population témoins, la fréquence de la délétion est calculée à partir de l'équilibre  $p^2+2pq+q^2$

# Sur la population des femmes ayant une activité professionnelle au 1er trimestre de grossesse

## QUELLE TAILLE D'ÉCHANTILLON ?

De façon générale, les maladies d'origine complexe et multifactorielle se caractérisent par des effets modestes de multiples facteurs (environnemental, gène et interaction gène-environnement). Dans ce contexte, la taille d'échantillon adéquate pour la détection d'une interaction gène-environnement avec une puissance statistique raisonnable, atteint aisément les 1 000 sujets-cas (mode de transmission dominant) ou les 10 000 sujets-cas (mode de transmission récessif *a priori* le plus défavorable) dans la situation d'un polymorphisme et d'une exposition peu communs (<10%) (auxquels doivent être ajoutés autant de sujets-témoins pour le schéma d'étude cas-témoins, et autant de couples parentaux pour le schéma d'étude cas-parents, Tableau 44).

**Tableau 44: Simulations par QUANTO de la taille de l'échantillon requise pour un test d'interaction gène-environnement par les schémas d'étude cas-témoins et cas-parents**

Prévalence de l'exposition dans la population	0.1	0.1	0.5	0.5	0.1	0.5
Fréquence de l'allèle d'intérêt dans la population	0.1	0.5	0.1	0.5	0.1	0.5
Risque relatif de maladie associé à l'exposition, parmi les non porteurs de l'allèle	1.5	1.5	1.5	1.5	4	4
Risque relatif de maladie associé à l'allèle d'intérêt parmi les non exposés	1.5	1.5	1.5	1.5	4	4
Interaction	2	2	2	2	2	2
(Risque relatif de maladie associé à l'effet conjoint)	(4.5)	(4.5)	(4.5)	(4.5)	(32)	(32)

### Taille de l'échantillon nécessaire (pour une puissance de détection à 80% et un risque d'erreur de 1<sup>ère</sup> espèce à 5%) : Nombre de sujets-cas

Hypothèse d'un mode dominant de transmission du gène						
Cas-témoins	1 593	1 894	767	1 021	1 394	3 050
Cas-parents	1 251	1 285	858	1 063	985	4 006
Hypothèse d'un mode récessif de transmission du gène						
Cas-témoins	21 512	1 367	10 282	663	17 363	933
Cas-parents	10 621	856	6 399	594	4 330	1 294

Calcul réalisé par le logiciel QUANTO avec les hypothèses suivantes : risque de base de la maladie dans la population = 0.001 ; Interaction multiplicative

## VERS DE NOUVEAUX SCHEMAS D'ETUDE ?

L'intérêt pour l'étude des interactions gène-environnement dans l'étiologie des maladies complexes est croissant. Des recherches récentes portent sur la construction de schémas d'étude plus performants et sur l'amélioration des méthodes d'analyse statistique associées, dans le but de maximiser la puissance statistique de détection d'une interaction gène-environnement.

Les stratégies alternatives au schéma d'étude cas-témoins traditionnel, permettant d'évaluer à la fois les effets propres et l'interaction entre un facteur environnemental et un facteur génétique, telles que les stratégies de contre-appariement ou d'appariement flexible, s'adaptent particulièrement aux situations de facteurs rares. Elles offrent théoriquement des puissances de test d'interaction gène-environnement plus séduisantes que le schéma traditionnel cas-témoins, mais requièrent encore une taille d'échantillon importante, en particulier lorsque l'interaction est modérée (<5).

Plus récemment, Andrieu et Goldstein (2004) proposent un schéma d'étude qui exploite et combine l'information de témoins traditionnels et de témoins familiaux. Son principe est d'apparier un sujet cas à ces deux types de témoins simultanément : un témoin de population et un frère ou une sœur (*sib*). L'idée majeure est d'accroître la fréquence du gène d'intérêt et du facteur environnemental (si les expositions du sujet-cas et de son témoin-*sib* sont corrélées). Malgré la violation à une hypothèse de validité de modélisation logistique conditionnelle (J.Little, communication personnelle), les auteurs montrent que ce schéma d'étude combiné est plus efficace pour l'étude d'une interaction gène-environnement que les schémas traditionnels cas-témoins (appariés), seulement lorsque les deux facteurs sont peu fréquents ( $<0.20$ ). Son efficacité relative décroît avec l'accroissement de la fréquence du facteur génétique. La taille de l'échantillon nécessaire est cependant prohibitive dans le cas de gènes rares ( $<0.01$ ), et devient raisonnable seulement lorsque le facteur environnemental est commun et l'interaction est forte ( $>5$ ).

D'autres schémas d'étude hybrides sont proposés dans la littérature. Ils s'intéressent essentiellement à la mesure d'une association génétique, et n'intègrent pas actuellement d'hétérogénéité environnementale. Il s'agit des travaux de Nagelkerke et al. (2004) et Epstein et al. (2005). Certains de ces travaux incorporent des données familiales étendues (avec des familles multiplexes) et utilisent la méthode des équations estimantes généralisées GEE considérant ainsi la corrélation familiale (Slager et al. 2003).

D'autres schémas d'étude peuvent être exploités pour examiner une interaction gène-environnement ; il s'agit des études cas-témoins appariées pour lesquelles les témoins sont construits à partir de données familiales telles que les parents, les cousins ou les frères et sœurs (*sibs*), non atteints. Les approches cas-*sibs* et cas-parents sont plus performants que l'approche cas-témoins appariée traditionnelle. De façon équivalente au schéma d'étude combiné, leur efficacité est semblable à l'approche cas-témoins traditionnelle lorsque le polymorphisme n'est pas rare.

De façon générale, la majorité de ces développements est récente et encore très peu appliquée. Lorsque l'efficacité est évaluée, elle est montrée supérieure à celle du schéma d'étude traditionnel cas-témoins, seulement dans certaines conditions spécifiques. Leur faisabilité est compromise et elle nécessite une taille d'échantillon élevée pour atteindre une puissance statistique raisonnable. Le recrutement des sujets peut devenir en pratique très complexe (stratégie de contre-appariement) et contraignant (schéma d'étude hybride).

#### VERS DES COORDINATIONS ENTRE LES ETUDES ?

Les méta-analyses d'études publiées ont la possibilité de surmonter le problème du manque de puissance statistique et peuvent aider à clarifier de l'existence d'une association et de sa force. Plusieurs méta-analyses ont déjà été réalisées dans le domaine de l'épidémiologie-génétique (Ioannidis et al. 2001, Munafò et Flint 2004). Ces outils statistiques sont cependant fortement débattus en épidémiologie, les principales critiques étant la précision sur-évaluées des résultats, la présence de biais de publication, et l'insuffisance d'ajustement sur des facteurs de confusion potentiels.

La stratégie d'étude d'interaction gène-environnement semble se situer davantage à travers des collaborations internationales par le consortium de plusieurs centres d'études déjà existantes ou par la planification d'une étude commune. Les analyses *poolées* s'appuient sur les données d'études individuelles (publiées ou non) et offrent ainsi plusieurs avantages comparé aux méta-analyses : une standardisation des informations et des définitions de cas et

des variables, un meilleur contrôle des facteurs de confusion, une prise en compte possible de l'évolution de la maladie dans le temps, une flexibilité des effets testés, une prise en compte de l'existence possible de déséquilibre de liaison entre plusieurs loci. Ces analyses nécessitent cependant des investissements considérables de temps et de ressources, pour la collection de l'ensemble des données et la coordination complète du projet (Ioannidis et al. 2002, Little et al. 2002). Actuellement, les méta-analyses et les analyses *poolées* ne permettent pas de considérer simultanément différents schémas d'étude, tels que cas-témoins et cas-parents.

#### PUISSANCE DES ETUDES D'INTERACTION GENE-ENVIRONNEMENT DANS L'ETIOLOGIE DES FENTES ORALES.

L'ensemble de la littérature existante des études d'interaction gène-environnement dans l'étiologie des fentes orales, présente des tailles d'échantillon équivalentes à celle de notre étude, pour les études européennes, ou légèrement plus importantes pour les études américaines (cf Tableau 10). Cette littérature est encore peu développée et s'intéresse à une variété d'interactions potentielles. Une unique méta-analyse, à ce jour, a été réalisée à partir de 5 études cas-témoins pour examiner l'association entre le polymorphisme *TGF $\alpha$  TaqI*, la consommation maternelle de tabac, et l'interaction des deux facteurs avec le risque de fentes orales non syndromiques (Zeiger et al. 2005).

Le consortium international de la génétique des fentes orales (*International Consortium of Oral Clefts Genetics*, ICOCG) présente, en 2002, des lignes directrices précises et rigoureuses pour les études de l'étiologie des fentes orales non syndromiques, dans le but de faciliter les comparaisons des études et le *pool* de ces données *a priori* et *a posteriori* (Mitchell et al. 2002).

## Discussion du choix *a priori* des polymorphismes

Lors de la mise en place de ce projet, le choix des gènes s'est réalisé *a priori*, et a porté sur deux catégories de gènes polymorphiques susceptibles d'influer sur la survenue d'une fente orale. Il s'agit des gènes du développement embryonnaire de la face, et des gènes impliqués dans le métabolisme de facteurs exogènes d'expositions maternelles soupçonnées de jouer un rôle dans l'étiologie des fentes orales.

### L'ETUDE DE GENES DE METABOLISME ET EXPLICATION INCERTAINE DES RESULTATS.

Pour ce deuxième groupe de gènes, l'idée initiale de l'étude était d'examiner les susceptibilités génétiques pouvant modifier l'association entre l'exposition maternelle et le risque de fentes orales (interaction gène-environnement statistique), et ainsi d'identifier des sous-groupes de population à risque. Egalement, une association entre ces gènes de métabolisme et le risque de fentes orales permet de clarifier des mécanismes biologiques pouvant expliquer la survenue de la malformation (interaction gène-environnement biologique).

Ce projet nous montre que cette seconde idée, défendue par Clayton et McKeigue (2001), est difficilement exploitable dans certaines conditions. La première difficulté se situe dans les propriétés spécifiques des gènes de l'étude. Le gène *ADH1C* s'avère être un gène impliqué dans le métabolisme de multiples substances exogènes. L'identification du (ou des) mécanisme(s) expliquant l'association potentielle entre ce gène et le risque de fentes orales n'apparaît pas évidente. Les différences d'activité enzymatique du polymorphisme *C677T* du gène *MTHFR* semblent s'estomper lorsque le niveau d'apport en folates (substrat indirect de la *MTHFR*) est adéquat. Ainsi, une absence d'association entre le gène *MTHFR* et le risque de fentes orales n'indique pas nécessairement que le mécanisme de métabolisme des folates n'a pas de rôle dans la survenue d'une fente orale. La seconde difficulté concerne l'interprétation d'une association génétique dans la survenue d'une malformation congénitale, en particulier lorsqu'elle est précoce lors de l'embryogenèse. Les gènes à l'âge embryonnaire possèdent des activités et des expressions spécifiques, qui peuvent être très différentes de celles des gènes à l'âge adulte. Peu d'expérimentations examinent l'existence d'une activité enzymatique de gènes de métabolisme, *in vivo* lors de la vie embryonnaire chez l'humain. Ainsi, le mécanisme d'un résultat d'association génétique dans ce contexte est décrit par les connaissances provenant d'expériences chez l'homme adulte, ou chez l'animal. L'interprétation peut alors mettre en évidence un mécanisme biologique inexistant dans le cadre spatio-temporel de la formation d'une fente orale.

### PROBLEMATIQUE DU CHOIX DES POLYMORPHISMES D'ETUDE.

La majorité des polymorphismes étudiés dans ce projet sont des SNPs (*Single Nucleotide Polymorphism*). Un SNP est une mutation caractérisée par la substitution d'un seul nucléotide. Le génome humain contiendrait près de 10 millions de SNPs, dont au moins 5 millions seraient déjà répertoriés dans des bases de données publiques (Rebeck et al. 2004).

La performance et la disponibilité des technologies de la génomique ont rapidement progressé depuis la mise en place de ce projet (1998). La détermination d'un SNP peut être aujourd'hui

rapidement effectuée et n'est pas chère ; en revanche, l'interprétation causale d'une association entre un marqueur génétique décrit par un SNP et une maladie est toujours sujette à un possible déséquilibre de liaison entre ce marqueur et le vrai locus de la maladie. Il semble alors plus approprié d'utiliser l'haplotype formé par différents SNPs. L'étiologie des maladies multifactorielles fait intervenir des mécanismes complexes impliquant possiblement plusieurs gènes, des expositions environnementales multiples, des interactions gène-environnement, et des interactions gène-gène. La recherche d'une association entre une unique exposition environnementale ou un unique gène et le risque d'une maladie peut être aujourd'hui décevante (Editorial de *Cancer Epidemiology, Biomarkers & Prevention* 2004), à juste titre si l'on considère l'énorme volume de données disponibles.

En conclusion, les mesures d'association génétique et d'interaction gène-environnement confrontent les futures études épidémiologiques à de nombreux challenges.

#### PRENDRE EN COMPTE LA MULTIPLICITE DE TESTS.

Avec une erreur de première espèce fixée à 5% pour 1000 associations génétiques testées, nous pouvons nous attendre à observer 50 associations génétiques (statistiquement significatives avec  $p < 0.05$ ), alors qu'aucun des gènes n'est associé au risque de maladie. L'adoption d'un critère plus strict d'exclusion du hasard comme explication des résultats est recommandée, même si aucune stratégie n'est communément proposée. La correction du seuil du degré de signification (*p-value*) par la méthode de Bonferroni peut être envisagée. Cette correction divise le seuil classique fixé à 5% par le nombre d'hypothèses testées ; cependant, aucun consensus n'est établi sur le nombre d'hypothèses qui doit être pris en compte, et cette solution pointe sur les problèmes pratiques de l'utilisation du *p* comme une preuve statistique. Des approches bayésiennes sont proposées, pour lesquelles l'interprétation des résultats d'une association est un problème de décision. Elles sont plus subjectives, car la décision dépend des probabilités que les auteurs accordent à chaque résultat possible (absence de détection d'une vraie association, détection d'une fausse association...) et à l'importance (ou l'utilité) que les auteurs placent dans leurs résultats. Enfin, basé sur l'observation (fréquentiste) parmi 1000 études que le critère du degré de signification doit être 400 fois plus strict, pour assurer un réel ratio d'associations vraie-positives: fausse-positives de 20:1, certains auteurs fixent le seuil du *p* à  $5 \times 10^{-5}$  (Colhoun et al. 2003).

Cette réduction importante du risque de première espèce, même si elle semble incontournable, a un impact non négligeable sur la taille de l'échantillon nécessaire pour maintenir une puissance statistique raisonnable. D'autres recherches s'orientent vers de nouvelles priorités qui, indirectement, permettent d'assouplir cette décision.

#### PRENDRE EN COMPTE LA COMPLEXITE DES DONNEES A ANALYSER.

De nouvelles approches d'analyse statistique émergent pour analyser la complexité de ces informations dans le cadre d'une étude épidémiologique. Nous présentons ici une liste non exhaustive de ces travaux.

Certains auteurs étendent les méthodes d'analyse statistique de mesure d'association appliquée aux SNPs (marqueur biallélique), aux études de marqueurs multi-alléliques, d'interaction gène-gène, d'haplotypes, et d'interactions haplotype-environnement (Clayton et Jones 1999, Schaid et al. 2004, Cordell et al. 2004, Lake et Laird 2004, revue dans Zhao et al. 2003). La valeur des approches basées sur l'information des haplotypes est aussi intensivement débattue (Carlson et al. 2004, Neale et Sham 2004, Zhang et al. 2004).

D'autres travaux proposent d'utiliser des méthodologies statistiques plus innovantes pour les études épidémiologiques. Des modèles à effets mixtes sont proposés pour examiner l'association entre plusieurs polymorphismes, des facteurs environnementaux et le risque d'une maladie (ou un phénotype) (Foulkes et al. 2005). Ces modèles permettent d'intégrer des facteurs de confusion potentiels. Les réseaux de neurones définissent une autre approche qui peut être considérée comme une généralisation des modèles logistiques et permet de prendre en compte de nombreuses dimensions. Ils sont proposés par Curtis et al. (2001) et Ritchie et al. (2003). Les techniques de réduction de dimension peuvent permettre de « synthétiser » les nombreuses informations en une unique dimension (information), dont la relation est ensuite évaluée avec la maladie. Elles sont appliquées pour examiner des interactions gène-gène et gène-environnement (jusqu'à l'ordre 15) dans le cadre d'une étude cas-témoins, par Hahn et al. (2003) (*Multifactor dimensionality reduction*). La méthodologie non-paramétrique des arbres de classification permet d'identifier des structures complexes d'un jeu de données. La structure est imposée par l'arbre mais elle n'est pas limitée par les contraintes classiques de régression linéaire. Elle est adaptée au champ d'application de l'épidémiologie génétique par Zhang et Bonney (2000) et permet d'intégrer des interactions d'ordre élevé.

#### REPOUDRE A L'OBJECTIF DE RECHERCHE ETIOLOGIQUE EN CONSIDERANT LES MECANISMES BIOLOGIQUES.

Les méthodologies modélisant la complexité des relations entre facteurs environnementaux, facteurs génétiques, et le risque d'une maladie, sont utiles à la recherche étiologique seulement s'ils participent à la compréhension des mécanismes biologiques sous-jacents. Dans le contexte d'identification du génome humain, les études épidémiologiques, intégrant des informations génétiques, doivent privilégier l'évaluation des polymorphismes fonctionnels. Nous appelons un polymorphisme fonctionnel s'il altère la séquence d'acides aminés, et si les différentes formes de ce polymorphisme produisent des protéines de fonction ou structure différentes.

Contrairement aux études de liaison ou aux techniques de balayage du génome, les études d'association doivent s'appuyer sur des hypothèses *a priori* de plausibilité biologique. La construction de ces hypothèses se fait à partir des connaissances actuelles de biochimie et de biologie moléculaire, chez l'homme et l'animal, et elle est ainsi influencée par le manque de connaissances pour certains mécanismes (par exemple lors de la vie embryonnaire, cf plus-haut). La compréhension des systèmes biologiques faisant intervenir des facteurs génétiques et environnementaux et leurs interactions, est cruciale pour l'étude de maladie à étiologie complexe. Les recherches biologiques, dans ce sens, doivent être développées, en particulier pour la période embryonnaire.

Finalement, les études épidémiologiques d'association et d'interaction gène-environnement, génératrices de sources de données, occupent une position-clé à l'intérieur d'un travail de collaboration entre des disciplines très diverses (statistique, génétique, biochimie, biologie moléculaire, épidémiologie).

**BIBLIOGRAPHIE**

**& AUTRES REFERENCES**



## [INTRODUCTION GENERALE] Introduction / L'étiologie complexe des fentes orales / Objectif du projet

- Abrishamchian AR, Khoury MJ, Calle EE. The contribution of maternal epilepsy and its treatment to the etiology of oral clefts: a population based case-control study. *Genet Epidemiol.* 1994;11:343-51
- Ardinger HH, Buetow KH, Bell GI, Bardach J, VanDemark DR, Murray JC. Association of genetic variation of the transforming growth factor-alpha gene with cleft lip and palate. *Am J Hum Genet.* 1989;45:348-53
- Arpino C, Brescianini S, Robert E, Castilla EE, Cocchi G, Cornel MC, de Vigan C, Lancaster PA, Merlob P, Sumiyoshi Y, Zampino G, Renzi C, Rosano A, Mastroiacovo P. Teratogenic effects of antiepileptic drugs: use of an international database on malformations and drug exposure. *Epilepsia* 2000;41:1436-1443
- Beaty TH, Maestri NE, Hetmanski JB, Wyszynski DF, Vanderkolk CA, Simpson JC, McIntosh I, Smith EA, Zeiger JS, Raymond GV, Panny SR, Tiftt CJ, Lewanda AF, Cristion CA, Wulfsberg EA. Testing for interaction between maternal smoking and TGFA genotype among oral cleft cases born in Maryland 1992-1996. *Cleft Palate Craniofac J.* 1997;34:447-54
- Beaty TH, Wang H, Hetmanski JB, Fan YT, Zeiger JS, Liang KY, Chiu YF, Vanderkolk CA, Seifert KC, Wulfsberg EA, Raymond G, Panny SR, McIntosh I. A case-control study of nonsyndromic oral clefts in Maryland. *Ann Epidemiol* 2001;11:434-42
- Beaty TH, Hetmanski JB, Zeiger JS, Fan YT, Liang KY, Vanderkolk CA, McIntosh I. Testing candidate genes for non-syndromic oral clefts using a case-parent trio design. *Genet Epidemiol.* 2002;22:1-11
- Bianchi F, Cianciulli D, Pierini A, Seniori Costantini A. Congenital malformations and maternal occupation: a registry based case-control study. *Occup Environ Med.* 1997;54:223-8
- Bille C, Skytthe A, Vach W, Knudsen LB, Andersen AM, Murray JC, Christensen K. Parent's age and the risk of oral clefts. *Epidemiology.* 2005a;16:311-6
- Bille C, Knudsen LB, Christensen K. Changing Lifestyles and Oral Clefts Occurrence in Denmark. *Cleft Palate Craniofac J.* 2005b;42:255-59
- Blanton SH, Bertin T, Patel S, Stal S, Mulliken JB, Hecht JT. Nonsyndromic cleft lip and palate: four chromosomal regions of interest. *Am J Med Genet A.* 2004a;125:28-37
- Blanton SH, Bertin T, Serna ME, Stal S, Mulliken JB, Hecht JT. Association of chromosomal regions 3p21.2, 10p13, and 16p13.3 with nonsyndromic cleft lip and palate. *Am J Med Genet A.* 2004b;125:23-7
- Botto LD, Olney RS, Erickson JD. Vitamin supplements and the risk for congenital anomalies other than neural tube defects. *Am J Med Genet C Semin Med Genet.* 2004;125:12-21
- Bove FJ, Fulcomer MC, Klotz JB, Esmart J, Dufficy EM, Savrin JE. Public drinking water contamination and birth outcomes. *Am J Epidemiol.* 1995;141:850-62
- Bove F, Shim Y, Zeitz P. Drinking water contaminants and adverse pregnancy outcomes: a review. *Environ Health Perspect.* 2002;110 Suppl 1:61-74
- CDC - Centers for Disease Control and Prevention. Economic costs of birth defects and cerebral palsy – United States, 1992. *MMWR* 1992; 44(37) (<http://www.cdc.gov/mmwr/preview/mmwrhtml/00038946.htm>).
- Chenevix-Trench G, Jones K, Green AC, Duffy DL, Martin NG. Cleft lip with or without cleft palate: associations with transforming growth factor alpha and retinoic acid receptor loci. *Am J Hum Genet.* 1992;51:1377-85
- Chong SS, Cheah FSH, Wang Jabs E. Genes implicated in lip and palate development. Chapter 3 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment.* Oxford, Oxford University Press, 2002, p25-39
- Christensen K. The 20th century Danish facial cleft population--epidemiological and genetic-epidemiological studies. *Cleft Palate Craniofac J.* 1999a;36:96-104
- Christensen K, Olsen J, Norgaard-Pedersen B, Basso O, Stovring H, Milhollin-Johnson L, Murray JC. Oral clefts, transforming growth factor alpha gene variants, and maternal smoking: a population-based case-control study in Denmark, 1991-1994. *Am J Epidemiol* 1999b;149:248-55
- Chung KC, Kowalski CP, Kim HM, Buchman SR. Maternal cigarette smoking during pregnancy and the risk of having a child with cleft lip/palate. *Plast Reconstr Surg.* 2000;105:485-91
- Clark JD, Mossey PA, Sharp L, Little J. Socioeconomic status and orofacial clefts in Scotland, 1989 to 1998. *Cleft Palate Craniofac J.* 2003;40:481-5
- Cohen MM. Syndromes with orofacial clefting. Chapter 6 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment.* Oxford, Oxford University Press, 2002, p53-65
- Cordier S, Ha MC, Ayme S, Goujard J. Maternal occupational exposure and congenital malformations. *Scand J Work Environ Health.* 1992;18:11-7
- Cordier S, Bergeret A, Goujard J, Ha MC, Ayme S, Bianchi F, Calzolari E, De Walle HE, Knill-Jones R, Candela S, Dale I, Dananche B, de Vigan C, Fevotte J, Kiel G, Mandereau L. Congenital malformation and maternal occupational exposure to glycol ethers. Occupational Exposure and Congenital Malformations Working Group. *Epidemiology.* 1997;8:355-63
- Cordier S, Szabova E, Fevotte J, Bergeret A, Plackova S, Mandereau L. Congenital malformations and maternal exposure to glycol ethers in the Slovak Republic. *Epidemiology.* 2001;12:592-3
- Czeizel AE, Timar L, Sarkozi A. Dose-dependent effect of folic acid on the prevention of orofacial clefts. *Pediatrics* 1999;104:e66
- Czeizel AE, Dobo M, Vargha P. Hungarian cohort-controlled trial of periconceptional multivitamin supplementation shows a reduction in certain congenital abnormalities. *Birth Defects Research (Part A).* 2004;70:853-61
- Dodds L, King WD. Relation between trihalomethane compounds and birth defects. *Occup Environ Med.* 2001;58:443-6
- Dravet C, Julian C, Legras C, Magaouda A, Guerrini R, Genton P, Soulayrol S, Giraud N, Mesdjian E, Trentin G, et al. Epilepsy, antiepileptic drugs, and malformations in children of women with epilepsy: a French prospective cohort study. *Neurology* 1992;42:75-82

- Ericson A, Kallen B, Westerholm P. Cigarette smoking as an etiologic factor in cleft lip and palate. *Am J Obstet Gynecol* 1979;135:346-51 (abstract)
- Etheredge AJ, Christensen K, Del Junco D, Murray JC, Mitchell LE. Evaluation of two methods for assessing gene-environment interactions using data from the Danish case-control study of facial clefts. *Birth Defects Res A Clin Mol Teratol*. 2005 [Sous presse]
- EUROCAT Special Report. EUROCAT and Orofacial Clefts: the Epidemiology of Orofacial clefts in 30 European regions. 2000 (<http://www.eurocat.ulster.ac.uk/pubdata/Publications.html>).
- Evans DR, Newcombe RG, Campbell H. Maternal smoking habits and congenital malformations: a population study. *Br Med J*. 1979;2(6183):171-3 (abstract)
- Fallin MD, Hetmanski JB, Park J, Scott AF, Ingersoll R, Fuernkranz HA, McIntosh I, Beaty TH. Family-based analysis of MSX1 haplotypes for association with oral clefts. *Genet Epidemiol* 2003;25:168-75
- Feng H, Sassani R, Bartlett SP, Lee A, Hecht JT, Malcolm S, Winter RM, Vintiner GM, Buetow KH, Gasser DL. Evidence, from family studies, for linkage disequilibrium between TGFA and a gene for nonsyndromic cleft lip with or without cleft palate. *Am J Hum Genet*. 1994;55:932-6 (abstract)
- Garcia AM, Fletcher T. Maternal occupation in the leather industry and selected congenital malformations. *Occup Environ Med*. 1998;55:284-6
- Garcia AM, Fletcher T, Benavides FG, Orts E. Parental agricultural work and selected congenital malformations. *Am J Epidemiol*. 1999;149:64-74
- Gordon JE, Shy CM. Agricultural chemical use and congenital cleft lip and/or palate. *Arch Environ Health*. 1981;36:213-21
- Hanke W, Jurewicz J. The risk of adverse reproductive and developmental disorders due to occupational pesticide exposure: an overview of current epidemiological evidence. *Int J Occup Med Environ Health*. 2004;17:223-43
- Hayes C, Werler MM, Willett WC, Mitchell AA. Case-control study of periconceptional folic acid supplementation and oral clefts. *Am J Epidemiol*. 1996;143:1229-34
- Hayes C. Environmental risk factors and oral clefts. Chapter 13 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment*. Oxford, Oxford University Press, 2002, p159-69
- Hecht JT, Annegers JF. Familial aggregation of epilepsy and clefting disorders: a review of the literature. *Epilepsia*. 1990;31:574-7
- Hernandez-Diaz S, Werler MM, Walker AM, Mitchell AA. Folic acid antagonists during pregnancy and the risk of birth defects. *N Engl J Med* 2000;343:1608-1614
- Hill L, Murphy M, McDowall M, Paul AH. Maternal drug histories and congenital malformations: limb reduction defects and oral clefts. *J Epidemiol Community Health*. 1988;42:1-7 (abstract)
- Holder SE, Vintiner GM, Farren B, Malcolm S, Winter RM. Confirmation of an association between RFLPs at the transforming growth factor-alpha locus and non-syndromic cleft lip and palate. *J Med Genet*. 1992;29:390-2
- Holmberg PC, Herberg S, Kurppa K, Rantala K, Riala R. Oral clefts and organic solvent exposure during pregnancy. *Int Arch Occup Environ Health*. 1982;50:371-6
- Hwang SJ, Beaty TH, Panny SR, Street NA, Joseph JM, Gordon S, McIntosh I, Francomano CA. Association study of transforming growth factor alpha (TGF-alpha) TaqI polymorphism and oral clefts: Indication of gene-environment interaction in a population-based sample of infants with birth defects. *Am J Epidemiol* 1995; 141:629-36
- ICBDMS - International Clearinghouse for Birth Defects Monitoring Systems (<http://www.icbd.org/>)
- Itikala PR, Watkins ML, Mulinare J, Moore CA, Liu Y. Maternal multivitamin use and orofacial clefts in offspring. *Teratology*. 2001;63:79-86 (abstract)
- Jugessur A, Lie RT, Wilcox AJ, Murray JC, Taylor JA, Saugstad OD, Vindenes HA, Abyholm F. Variants of developmental genes (TGFA, TGFB3, and MSX1) and their associations with orofacial clefts: a case-parent triad analysis. *Genet Epidemiol*. 2003a;24:230-9
- Jugessur A, Wilcox AJ, Lie RT, Murray JC, Taylor JA, Ulvik A, Drevon CA, Vindenes HA, Abyholm FE. Exploring the effects of methylenetetrahydrofolate reductase gene variants C677T and A1298C on the risk of orofacial clefts in 261 Norwegian case-parent triads. *Am J Epidemiol*. 2003b;157:1083-91
- Jugessur A, Lie RT, Wilcox AJ, Murray JC, Taylor JA, Saugstad OD, Vindenes HA, Abyholm FE. Cleft palate, transforming growth factor alpha gene variants, and maternal exposures: assessing gene-environment interactions in case-parent triads. *Genet Epidemiol*. 2003c;25:367-74
- Juriloff DM. Mapping studies in animal models. Chapter 22 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment*. Oxford, Oxford University Press, 2002, p265-82
- Kallen K. Maternal smoking and orofacial clefts. *Cleft Palate Craniofac J*. 1997; 34:11-6 (abstract)
- Kelsey JL, Dwyer T, Holford TR, Bracken MB. Maternal smoking and congenital malformations: an epidemiological study. *J Epidemiol Community Health*. 1978;32:102-7 (abstract)
- Khoury MJ, Weinstein A, Panny S, Holtzman NA, Lindsay PK, Farrel K, Eisenberg M. Maternal cigarette smoking and oral clefts: A population-based study. *Am J Public Health* 1987;77:623-25 (abstract)
- Khoury MJ, Gomez-Farias M, Mulinare J. Does maternal cigarette smoking during pregnancy cause cleft lip and palate in offspring? *Am J Dis Child*. 1989;143:333-7 (abstract)
- Kristensen P, Irgens LM, Daltveit AK, Andersen A. Perinatal outcome among children of men exposed to lead and organic solvents in the printing industry. *Am J Epidemiol*. 1993;137:134-44 (abstract)
- Kristensen P, Irgens LM, Andersen A, Bye AS, Sundheim L. Birth defects among offspring of Norwegian farmers, 1967-1991. *Epidemiology*. 1997;8:537-44
- Lammer EJ, Shaw GM, Iovannisci DM, Van Waes J, Finnell RH. Maternal smoking and the risk of orofacial clefts: Susceptibility with NAT1 and NAT2 polymorphisms. *Epidemiology*. 2004a;15(2):150-6
- Lammer EJ, Shaw GM, Iovannisci DM, Finnell RH. Periconceptional multivitamin intake during early pregnancy, genetic variation of acetyl-N-transferase 1 (NAT1), and risk for orofacial clefts. *Birth Defects Res A Clin Mol Teratol*. 2004b;70:846-52

- Lammer EJ, Shaw GM, Iovannisci DM, Finnell RH. Maternal smoking, genetic variation of glutathione-s-transferases, and risk of orofacial clefts. *Epidemiology*. 2005 Sous presse
- Laumon B, Martin JL, Collet P, Bertucat I, Verney MP, Robert E. Exposure to organic solvents during pregnancy and oral clefts: a case-control study. *Reprod Toxicol*. 1996;10:15-9
- Leck I. The etiology of human malformations: insights from epidemiology. *Teratology*. 1972;5:303-14
- Lidral AC, Romitti PA, Basart AM, Doetschman T, Leysens NJ, Daack-Hirsch S, Semina EV, Johnson LR, Machida J, Burds A, Parnell TJ, Rubenstein JL, Murray JC. Association of MSX1 and TGFB3 with nonsyndromic clefting in humans. *Am J Hum Genet*. 1998;63:557-68
- Lieff S, Olshan AF, Werler M, Strauss RP, Smith J, Mitchell A. Maternal cigarette smoking during pregnancy and risk of oral clefts in newborns. *Am J Epidemiol* 1999;150:683-94
- Little J. Which genes, exposures and interactions? Presented to European Science Foundation (ESF) Conference on Gene-Environment interaction in early human development: demonstration project on orofacial clefts. Strasbourg, France, 1999
- Little J, Cardy A, Arslan MT, Gilmour M, Mossey PA. Smoking and orofacial clefts: a United Kingdom-based case-control study. *Cleft Palate Craniofac J*. 2004a;41:381-6
- Little J, Cardy A, Munger RG. Tobacco smoking and oral clefts : a meta-analysis. *Bulletin of World Health Organization*. 2004b; 82:213-8 (A-E)
- Loffredo LC, Souza JM, Freitas JA, Mossey PA. Oral clefts and vitamin supplementation. *Cleft Palate Craniofac J*. 2001;38:76-83
- Lorente C, Cordier S, Goujard J, Ayme S, Bianchi F, Calzolari E, De Walle HE, Knill-Jones R. Tobacco and alcohol use during pregnancy and risk of oral clefts. *Am J Public Health* 2000a;90:415-19
- Lorente C, Cordier S, Bergeret A, De Walle HE, Goujard J, Ayme S, Knill-Jones R, Calzolari E, Bianchi F. Maternal occupational risk factors for oral clefts. *Occupational Exposure and Congenital Malformation Working Group. Scand J Work Environ Health*. 2000b;26:137-45
- Lorente C. Rôle des facteurs de risque environnementaux dans l'étiologie des fentes orales. Thèse, Université de Paris XI, 2001
- Maestri NE, Beaty TH, Hetmanski J, Smith EA, McIntosh I, Wyszynski DF, Liang KY, Duffy DL, VanderKolk C. Application of transmission disequilibrium tests to nonsyndromic oral clefts: including candidate genes and environmental exposures in the models. *Am J Med Genet*. 1997;73:337-44
- Malloy MH, Kleinman JC, Bakewell JM, Schramm WF, Land GH. Maternal smoking during pregnancy: no association with congenital malformations in Missouri 1980-83. *Am J Public Health*. 1989;79:1243-6 (abstract)
- Marazita ML, Neiswanger K. Association studies. Chapter 20 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment*. Oxford, Oxford University Press, 2002, p240-54
- Martinelli M, Scapoli L, Pezzetti F, Carinci F, Carinci P, Stabellini G, Bisceglia L, Gombos F, Tognon M. C677T variant form at the MTHFR gene and CL/P: a risk factor for mothers? *Am J Med Genet*. 2001;98:357-60
- Matalon S, Schechtman S, Goldzweig G, Ornoy A. The teratogenic effect of carbamazepine: a meta-analysis of 1255 exposures. *Reprod Toxicol*. 2002;16:9-17
- Meyer KA, Werler MM, Hayes C, Mitchell AA. Low maternal alcohol consumption during pregnancy and oral clefts in offspring: the Slone Birth Defects Study. *Birth Defects Res A Clin Mol Teratol*. 2003;67:509-14
- Meyer KA, Williams P, Hernandez-Diaz S, Crattingius S. Smoking and the risk of oral clefts: exploring the impact of study designs. *Epidemiology*. 2004;15:671-8
- Mills JL, Graubard BI. Is moderate drinking during pregnancy associated with an increased risk for malformations? *Pediatrics*. 1987;80:309-14
- Mills JL, Kirke PN, Molloy AM, Burke H, Conley MR, Lee YJ, Mayne PD, Weir DG, Scott JM. Methylenetetrahydrofolate reductase thermolabile variant and oral clefts. *Am J Med Genet*. 1999;86:71-4
- Mitchell LE, Healey SC, Chenevix-Trench G. Evidence for an association between nonsyndromic cleft lip with or without cleft palate and a gene located on the long arm of chromosome 4. *Am J Hum Genet*. 1995;57:1130-6
- Mitchell LE. Transforming growth factor alpha locus and nonsyndromic cleft lip with or without cleft palate: a reappraisal. *Genet Epidemiol*. 1997;14:231-40
- Mitchell LE, Murray JC, O'Brien S, Christensen K. Evaluation of two putative susceptibility loci for oral clefts in the Danish population. *Am J Epidemiol* 2001;153:1007-15
- Mitchell LE. Mode of inheritance of oral clefts. Chapter 19 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment*. Oxford, Oxford University Press, 2002a, p234-9
- Mitchell LE. Twin studies in oral cleft research. Chapter 17 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment*. Oxford, Oxford University Press, 2002b, p214-21
- Mitchell LE, Murray JC, O'Brien S, Christensen K. Retinoic acid receptor alpha gene variants, multivitamin use, and liver intake as risk factors for oral clefts: a population-based case-control study in Denmark, 1991-1994. *Am J Epidemiol*. 2003;158:69-76
- Mossey PA, Little J. Epidemiology of oral clefts: an international perspective. Chapter 12 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment*. Oxford, Oxford University Press, 2002: 127-158.
- Munger RG, Romitti PA, Daack-Hirsch S, Burns TL, Murray JC, Hanson J. Maternal alcohol use and risk of orofacial cleft birth defects. *Teratology*. 1996;54:27-33
- Munger RG. Maternal nutrition and oral clefts. Chapter 14 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment*. Oxford, Oxford University Press, 2002, p170-92
- Munger RG, Sauberlich HE, Corcoran C, Nepomuceno B, Daack-Hirsch S, Solon FS. Maternal vitamin B-6 and folate status and risk of oral cleft birth defects in the Philippines. *Birth Defects Res A Clin Mol Teratol*. 2004;70:464-71
- Natsume N, Kawai T, Suzuki T. Preference for vegetables rich in beta-carotene and manifestation of cleft lip and/or palate. *Plast Reconstr Surg*. 1995;95:934-5
- Nurminen T. Maternal pesticide exposure and pregnancy outcome. *J Occup Environ Med*. 1995;37:935-40
- Olshan AF, Teschke K, Baird PA. Paternal occupation and congenital anomalies in offspring. *Am J Ind Med*. 1991;20:447-75
- OMIM - Online Mendelian Inheritance in Man (<http://www.ncbi.nlm.nih.gov/omim/>).

- Pezzetti F, Martinelli M, Scapoli L, Carinci F, Palmieri A, Marchesini J, Carinci P, Caramelli E, Rullo R, Gombos F, Tognon M. Maternal MTHFR variant forms increase the risk in offspring of isolated nonsyndromic cleft lip with or without cleft palate. *Hum Mutat.* 2004;24:104-5
- Prescott NJ, Winter RM, Malcolm S. Maternal MTHFR genotype contributes to the risk of non-syndromic cleft lip and palate. *J Med Genet.* 2002;39:368-9
- Romitti PA, Lidral AC, Munger RG, Daack-Hirsch S, Burns TL, Murray JC. Candidate genes for nonsyndromic cleft lip and palate and maternal cigarette smoking and alcohol consumption: evaluation of genotype-environment interactions from a population-based case-control study of orofacial clefts. *Teratology* 1999;59:39-50
- Saal HM. Classification and Description of nonsyndromic Clefts. Chapter 5 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment.* Oxford, Oxford University Press, 2002, p47-65
- Sassani R, Bartlett SP, Feng H, Goldner-Sauve A, Haq AK, Buetow KH, Gasser DL. Association between alleles of the transforming growth factor-alpha locus and the occurrence of cleft lip. *Am J Med Genet.* 1993;45:565-9
- Saxen I. Associations between oral clefts and drugs taken during pregnancy. *Int J Epidemiol.* 1975;4:37-44 (abstract)
- Scapoli L, Pezzetti F, Carinci F, Martinelli M, Carinci P, Tognon M. Lack of linkage disequilibrium between transforming growth factor alpha Taq I polymorphism and cleft lip with or without cleft palate in families from Northeastern Italy. *Am J Med Genet.* 1998;75:203-6
- Schnitzer PG, Olshan AF, Erickson JD. Paternal occupation and risk of birth defects in offspring. *Epidemiology.* 1995;6:577-83
- Shaw GM, O'Malley CD, Wasserman CR, Tolarova MM, Lammer EJ. Maternal periconceptional use of multivitamins and reduced risk for conotruncal heart defects and limb deficiencies among offspring. *Am J Med Genet.* 1995;59:536-45
- Shaw GM, Wasserman CR, Block G, Lammer EJ. High maternal vitamin A intake and risk of anomalies of structures with a cranial neural crest cell contribution. *Lancet.* 1996a;347:899-900.
- Shaw GM, Wasserman CR, Lammer EJ, O'Malley CD, Murray JC, Basart AM, Tolarova MM. Orofacial clefts, parental cigarette smoking, and transforming growth factor-alpha gene variants. *Am J Hum Genet* 1996b;58:551-561
- Shaw GM, Rozen R, Finnell RH, Todoroff K, Lammer EJ. Infant C677T mutation in MTHFR, maternal periconceptional vitamin use, and cleft lip. *Am J Med Genet.* 1998a;80:196-8
- Shaw GM, Wasserman CR, Murray JC, Lammer EJ. Infant TGF-alpha genotype, orofacial clefts, and maternal periconceptional multivitamin use. *Cleft Palate Craniofac J.* 1998b;35:366-70
- Shaw GM, Lammer EJ. Maternal periconceptional alcohol consumption and risk for orofacial clefts. *J Pediatr.* 1999a;134:298-303
- Shaw GM, Wasserman CR, O'Malley CD, Nelson V, Jackson RJ. Maternal pesticide exposure from multiple sources and selected congenital anomalies. *Epidemiology.* 1999b;10:60-6
- Shaw GM, Todoroff K, Finnell RH, Rozen R, Lammer EJ. Maternal vitamin use, infant C677T mutation in MTHFR, and isolated cleft palate risk. *Am J Med Genet.* 1999c;85:84-5
- Shaw GM, Nelson V, Iovannisci DM, Finnell RH, Lammer EJ. Maternal occupational chemical exposures and biotransformation genotypes as risk factors for selected congenital anomalies. *Am J Epidemiol.* 2003a;157:475-84
- Shaw GM, Ranatunga D, Quach T, Neri E, Correa A, Neutra RR. Trihalomethane exposures from municipal water supplies and selected congenital malformations. *Epidemiology.* 2003b;14:191-9
- Shaw GM, Zhu H, Lammer EJ, Yang W, Finnell RH. Genetic variation of infant reduced folate carrier (A80G) and risk of orofacial and conotruncal heart defects. *Am J Epidemiol.* 2003c;158:747-52
- Shiono PH, Kelbanoff MA, Berendes HW. Congenital malformation and maternal smoking during pregnancy. *Teratology* 1986;34:65-71 (abstract)
- Thomas DC, Petitti DB, Goldhaber M, Swan SH, Rappaport EB, Hertz-Picciotto I. Reproductive outcomes in relation to malathion spraying in the San Francisco Bay Area, 1981-1982. *Epidemiology.* 1992;3:32-9
- van den Eeden SK, Karagas MR, Daling JR, Vaughan TL. A case-control study of maternal smoking and congenital malformations. *Ped Perinatal Epidemiol* 1990;4:147-55 (abstract)
- van Rooij IA, Wegerif MJ, Roelofs HM, Peters WH, Kuijpers-Jagtman AM, Zielhuis GA, Merkus HM, Steegers-Theunissen RP. Smoking, genetic polymorphisms in biotransformation enzymes, and nonsyndromic oral clefting: a gene-environment interaction. *Epidemiology* 2001;12:502-507
- van Rooij IA, Groenen PM, van Drongelen M, Te Morsche RH, Peters WH, Steegers-Theunissen RP. Orofacial clefts and spina bifida: N-acetyltransferase phenotype, maternal smoking, and medication use. *Teratology.* 2002; 66:260-6
- van Rooij IA, Swinkels DW, Blom HJ, Merkus HM, Steegers-Theunissen RP. Vitamin and homocysteine status of mothers and infants and the risk of nonsyndromic orofacial clefts. *Am J Obstet Gynecol.* 2003a;189:1155-60
- van Rooij IA, Vermeij-Keers C, Kluijtmans LA, Ocke MC, Zielhuis GA, Goorhuis-Brouwer SM, van der Biezen JJ, Kuijpers-Jagtman AM, Steegers-Theunissen RP. Does the interaction between maternal folate intake and the methylenetetrahydrofolate reductase polymorphisms affect the risk of cleft lip with or without cleft palate? *Am J Epidemiol.* 2003b;157:583-91
- van Rooij IA, Ocke MC, Straatman H, Zielhuis GA, Merkus HM, Steegers-Theunissen RP. Periconceptional folate intake by supplement and food reduces the risk of nonsyndromic cleft lip with or without cleft palate. *Prev Med* 2004; 39:689-94
- Vieira AR, Orioli IM, Murray JC. Maternal age and oral clefts: A reappraisal. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2002; 94:530-5
- Wennborg H, Magnusson LL, Bonde JP, Olsen J. Congenital malformations related to maternal exposure to specific agents in biomedical research laboratories. *J Occup Environ Med.* 2005;47:11-9
- Werler MM, Lammer EJ, Rosenberg L, Mitchell AA. Maternal vitamin A supplementation in relation to selected birth defects. *Teratology.* 1990a;42:497-503
- Werler MM, Lammer EJ, Rosenberg L, Mitchell AA. Maternal cigarette smoking during pregnancy in relation to oral clefts. *Am J Epidemiol* 1990b;132:926-32 (abstract)
- Werler MM, Lammer EJ, Rosenberg L, Mitchell AA. Maternal alcohol use in relation to selected birth defects. *Am J Epidemiol.* 1991 1;134:691-8
- Werler MM. Teratogen Update: Smoking and Reproductive Outcomes. *Teratology* 1997; 55:382-8

- Werler MM, Hayes C, Louik C, Shapiro S, Mitchell AA. Multivitamin supplementation and risk of birth defects. *Am J Epidemiol.* 1999;150:675-82 (abstract)
- WHO - World Health Organization. Global registry and database on craniofacial anomalies, Report of a WHO Registry Meeting on Craniofacial Anomalies, Brazil 2001 (<http://www.who.int/genomics/publications/en/>).
- Wong WY, Eskes TK, Kuijpers-Jagtman AM, Spauwen PH, Steegers EA, Thomas CM, Hamel BC, Blom HJ, Steegers-Theunissen RP. Nonsyndromic orofacial clefts: association with maternal hyperhomocysteinemia. *Teratology* 1999;60:253-7
- Wyszynski DF, Beaty TH. Review of the role of potential teratogens in the origin of human nonsyndromic oral clefts. *Teratology.* 1996a;53:309-17
- Wyszynski DF, Beaty TH, Maestri NE. Genetics of nonsyndromic oral clefts revisited. *Cleft Palate Craniofac J.* 1996b;33:406-17
- Wyszynski DF, Duffy DL, Beaty TH. Maternal cigarette smoking and oral clefts: a meta-analysis. *Cleft Palate Craniofac J.* 1997a;34:206-10
- Wyszynski DF, Maestri N, Lewanda AF, McIntosh I, Smith EA, Garcia-Delgado C, Vinageras-Guarneros E, Wulfsberg E, Beaty TH. No evidence of linkage for cleft lip with or without cleft palate to a marker near the transforming growth factor alpha locus in two populations. *Hum Hered.* 1997b;47:101-9 (abstract)
- Wyszynski DF, Diehl SR. Infant C677T mutation in MTHFR, maternal periconceptional vitamin use, and risk of nonsyndromic cleft lip. *Am J Med Genet.* 2000;92:79-80
- Wyszynski DF, Wu T. Use of US birth certificate data to estimate the risk of maternal cigarette smoking for oral clefting. *Cleft Palate Craniofac J.* 2002;39:188-92
- Wyszynski DF, Sarkozi A, Vargha P, Czeizel AE. Birth weight and gestational age of newborns with cleft lip with or without cleft palate and with isolated cleft palate. *J Clin Pediatr Dent.* 2003;27:185-90
- Zeiger JS, Beaty TH, Liang KY. Oral Clefts, Maternal Smoking, and TGFA: A Meta-Analysis of Gene-Environment Interaction. *Cleft Palate Craniofac J.* 2005;42:58-63
- Zuccherro TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S, Schutte BC, Marazita ML, Murray JC. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *N Engl J Med.* 2004;351:769-80

## [INTRODUCTION GENERALE] L'état de l'art des méthodes statistiques d'évaluation d'une interaction gène-environnement

- Albert PS, Ratnasinghe D, Tangrea J, Wacholder S. Limitations of the case-only design for identifying gene-environment interactions. *Am J Epidemiol.* 2001;154:687-93
- Allison DB. Transmission-Disequilibrium tests for quantitative traits. *Am J Hum Genet.* 1997; 60: 676-90
- Andrieu N, Goldstein AM. Use of relatives of cases as controls to identify risk factors when an interaction between environmental and genetic factors exists. *Int J Epidemiol.* 1996;25:649-57
- Andrieu N, Goldstein AM. Epidemiologic and genetic approaches in the study of gene-environment interaction: an overview of available methods. *Epidemiol Rev.* 1998;20:137-47
- Andrieu N, Goldstein AM, Thomas DC, Langholz B. Counter-matching in studies of gene-environment interaction: efficiency and feasibility. *Am J Epidemiol.* 2001;153:265-74
- Bacanu SA, Devlin B, Roeder K. The power of genomic control. *Am J Hum Genet.* 2000;66: 1933-44
- Bickeboller H, Clerget-Darpoux F. Statistical properties of the allelic and genotypic transmission/disequilibrium test for multiallelic markers. *Genet Epidemiol.* 1995;12:865-70
- Boehnke M, Langefeld CD. Genetic association mapping based on discordant pairs: the discordant-alleles test. *Am J Hum Genet.* 1998 ;62 :950-61
- Botto LD, Khoury MJ. Commentary: facing the challenge of gene-environment interaction: the two-by-four table and beyond. *Am J Epidemiol.* 2001; 153:1016-20
- Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet.* 2003;361:598-604
- Clayton D, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet.* 2001;358:1356-60
- Cordell HJ, Barratt BJ, Clayton DG. Case/pseudocontrol analysis in genetic association studies: A unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. *Genet Epidemiol.* 2004;26:167-85
- Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Ser. B* 1977 ; 39:1-38
- Deng HW, Chen WM, Recker RR. Population admixture: detection by Hardy-Weinberg test and its quantitative effects on linkage-disequilibrium methods for localizing genes underlying complex traits. *Genetics.* 2001;157:885-97
- Devlin B, Roeder K. Genomic control for association studies. *Biometrics.* 1999; 55:997-1004. (abstract)
- Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol.* 2001;60:155-66
- Eaves LJ, Sullivan P. Genotype-environment interaction in transmission disequilibrium tests. *Adv Genet.* 2001 ; 42 :223-40
- Ewens WJ, Spielman RS. The transmission/disequilibrium test: history, subdivision, and admixture. *Am J Hum Genet.* 1995;57:455-64
- Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet.* 1987; 51:227-33
- Flanders WD, Khoury MJ. Analysis of case-parental control studies: method for the study of associations between disease and genetic markers. *Am J Epidemiol.* 1996;144:696-703
- Flanders WD, Sun F, Yang Q. New estimator of the genotype risk ratio for use in case-parental control studies. *Am J Epidemiol.* 2001;154:259-63
- Foppa I, Spiegelman D. Power and sample size calculations for case-control studies of gene-environment interactions with a polytomous exposure variable. *Am J Epidemiol.* 1997 ;146:596-604
- García-Closas M, Lubin JH. Power and sample size calculations in case-control studies of gene-environment interactions: comments on different approaches. *Am J Epidemiol.* 1999;149:689-92
- García-Closas M, Rothman N, Lubin J. Misclassification in case-control studies of gene-environment interactions: assessment of bias and sample size. *Cancer Epidemiol Biomarkers Prev.* 1999;8:1043-50
- Gauderman WJ. Sample size requirements for association studies of gene-gene interaction. *Am J Epidemiol.* 2002 ;155:478-84
- Greenland S. Tests for interaction in epidemiologic studies: a review and a study of power. *Stat Med.* 1983;2:243-51 (abstract)
- Greenland S, Rothman KJ. Concepts of Interaction. Chapter 18 in Rothman KJ, Greenland S: *Modern Epidemiology*, Second Edition. Philadelphia, Lippincott-Raven, 1998, p330
- Hoggart CJ, Parra EJ, Shriver MD, Bonilla C, Kittles RA, Clayton DG, McKeigue PM. Control of confounding of genetic associations in stratified populations. *Am J Hum Genet.* 2003;72(6):1492-1504
- Hwang SJ, Beaty TH, Liang KY, Coresh J, Khoury MJ. Minimum sample size estimation to detect gene-environment interaction in case-control designs. *Am J Epidemiol.* 1994;140:1029-37. (abstract)
- Khoury MJ, Beaty TH, Cohen BH. *Fundamentals of Genetic Epidemiology*. New York, Oxford University Press, 1993
- Khoury MJ, Beaty TH, Hwang SJ. Detection of genotype-environment interaction in case-control studies of birth defects: how big a sample size? *Teratology.* 1995 ;51:336-43
- Khoury MJ, Flanders WD. Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! *Am J Epidemiol.* 1996; 144:207-13
- Kistner EO, Weinberg CR. Method for using complete and incomplete trios to identify genes related to a quantitative trait. *Genet Epidemiol.* 2004;27:33-42
- Knapp M. The transmission/disequilibrium test and parental-genotype reconstruction: the reconstruction-combined transmission/disequilibrium test. *Am J Hum Genet.* 1999; 64:861-70

- Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. *Genet Epidemiol.* 2000;19 Suppl 1:S36-42
- Lake SL, Laird NM. Tests of gene-environment interaction for case-parent triads with general environmental exposures. *Ann Hum Genet.* 2004 ;68:55-64
- Lange C, Laird NM. On a general class of conditional tests for family-based association studies in genetics: the asymptotic distribution, the conditional power, and optimality considerations. *Genet Epidemiol.* 2002;23:165-80
- Lazzeroni LC, Lange K. A conditional inference framework for extending the transmission/disequilibrium test. *Hum Hered.* 1998;48:67-81
- Maestri NE, Beaty TH, Hetmanski J, Smith EA, McIntosh I, Wyszynski DF, Liang KY, Duffy DL, VanderKolk C. Application of transmission disequilibrium tests to nonsyndromic oral clefts: including candidate genes and environmental exposures in the models. *Am J Med Genet.* 1997;73:337-44
- Morton NE, Collins A. Tests and estimates of allelic association in complex inheritance. *Proc Natl Acad Sci U S A.* 1998 15;95:11389-93
- Ottman R. Gene-environment interaction: definitions and study designs. *Prev Med.* 1996; 25:764-70
- Ottman R. An epidemiologic approach to gene-environment interaction. *Genet Epidemiol.* 1990;7:177-85
- Piegorsch WW, Weinberg CR, Taylor JA. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. *Stat Med.* 1994; 13:153-62
- Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in associations studies. *Am J Hum Genet.* 1999 ;65:220-8
- Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. *Am J Hum Genet.* 2000;67:170-81
- Pritchard JK, Donnelly P. Case-control studies of association in structured or admixed populations. *Theor Popul Biol.* 2001;60:227-37
- Rabinowitz D. A transmission disequilibrium test for quantitative trait loci. *Hum Hered.* 1997 ; 47 :342-50
- Rabinowitz D, Laird N. A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Hum Hered.* 2000;50:211-23
- Rabinowitz D. Unbiased discordant sib-pair tests when parental genotypes are missing. *Am J Med Genet.* 2001;105:57-9
- Reich DE, Goldstein DB. Detecting association in a case-control study while correcting for population stratification. *Genet Epidemiol.* 2001 Jan;20(1):4-16 (abstract)
- Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, Feldman MW. Genetic structure of human populations. *Science.* 2002;298:2381-5
- Satten GA, Flanders WD, Yang Q. R Accounting for unmeasured population substructure in case-control studies of genetic association using a novel latent-class model. *Am J Hum Genet.* 2001;68(2):466-77 (abstract)
- Schaid DJ, Sommer SS. Genotype relative risks: methods for design and analysis of candidate-gene association studies. *Am J Hum Genet.* 1993 ;53:1114-26
- Schaid DJ. General score tests for associations of genetic markers with disease using cases and their parents. *Genet Epidemiol.* 1996;13:423-49
- Schaid DJ, Rowland CM. Use of parents, sibs and unrelated controls for detection of associations between genetic markers and disease. *Am J Hum Genet.* 1998; 63:1492-506
- Schaid DJ. Likelihoods and TDT for the Case-Parents designs. *Genet Epidemiol.* 1999a;16 :250-60
- Schaid DJ. Case-parents design for gene-environment interaction. *Genet Epidemiol.* 1999b;16:261-73
- Schork NJ, Fallin D, Thiel B, Xu X, Broeckel U, Jacob HJ, Cohen D. The future of genetic case-control studies. *Adv Genet.* 2001;42:191-212
- Self SG, Longton G, Kopecky KJ, Liang KY. On estimating HLA/disease association with application to a study of aplastic anemia. *Biometrics* 1991; 47:53-61. (abstract)
- Sham PC, Curtis D. An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Ann Hum Genet.* 1995;59:323-36
- Smith PG, Day NE. The design of case-control studies: the influence of confounding and interaction effects. *Int J Epidemiol.* 1984 ;13:356-65 (abstract)
- Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet.* 1993;52:506-16
- Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am J Hum Genet.* 1998;62:450-8
- Stürmer T, Brenner H. Potential gain in efficiency and power to detect gene-environment interactions by matching in case-control studies. *Genet Epidemiol.* 2000;18:63-80
- Sun F, Flanders WD, Yang Q, Khoury MJ. A new method for estimating the risk ratio in studies using case-parental control design. *Am J Epidemiol.* 1998;148:902-9
- Terwilliger JD, Ott J. A haplotype-based 'haplotype relative risk' approach to detecting allelic associations. *Hum Hered.* 1992;42(6):337-46
- Thomas DC. Case-parents design for gene-environment interaction by Schaid. *Genet Epidemiol.* 2000;19:461-3
- Thomson G. Mapping disease genes: family-based association studies. *Am J Hum Genet.* 1995;57:487-98
- Umbach DM, Weinberg CR. Designing and analysing case-control studies to exploit independence of genotype and exposure. *Stat Med.* 1997; 16:1731-43
- Umbach DM, Weinberg CR. The use of case-parent triads to study joint effects of genotype and exposure. *Am J Hum Genet.* 2000; 66:251-61
- van den Oord EJ. The use of mixture models to perform quantitative tests for linkage disequilibrium, maternal effects, and parent-of-origin effects with incomplete subject-parent triads. *Behav Genet.* 2000;30:335-43
- Wacholder S, Rothman N, Caporaso N. Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. *J Natl Cancer Inst.* 2000;92:1151-8
- Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. *Cancer Epidemiol Biomarkers Prev.* 2002a;11:513-20
- Wacholder S, Chatterjee N, Hartge P. Joint effect of genes and environment distorted by selection biases: implications for hospital-based case-control studies. *Cancer Epidemiol Biomarkers Prev.* 2002b;11:885-9

## BIBLIOGRAPHIE

- Waldman ID, Robinson BF, Rowe DC. A logistic regression based extension of the TDT for continuous and categorical traits. *Ann Hum Genet.* 1999;63:329-40
- Wang Y, Localio R, Rebbeck TR. Evaluating bias due to population stratification in case-control association studies of admixed populations. *Genet Epidemiol.* 2004;27:14-20
- Weinberg CR, Wilcox AJ, Lie RT. A log-linear approach to case-parent-triad data: assessing effects of disease genes that act either directly or through maternal effects and that may be subject to parental imprinting. *Am J Hum Genet.* 1998;62:969-78
- Weinberg CR. Allowing for missing parents in genetic studies of case-parent triads. *Am J Hum Genet.* 1999a;64:1186-93
- Weinberg CR. Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. *Am J Hum Genet.* 1999b;65:229-35
- Weinberg CR, Umbach DM. Choosing a retrospective design to assess joint genetic and environmental contributions to risk. *Am J Epidemiol.* 2000;152:197-203
- Wilcox AJ, Weinberg CR, Lie RT. Distinguishing the effects of maternal and offspring genes through studies of "case-parent triads". *Am J Epidemiol.* 1998;148:893-901
- Witte JS, Gauderman WJ, Thomas DC. Asymptotic bias and efficiency in case-control studies of candidate genes and gene-environment interactions: basic family designs. *Am J Epidemiol.* 1999;149:693-705
- Yang Q, Khoury MJ, Flanders WD. Sample size requirements in case-only designs to detect gene-environment interaction. *Am J Epidemiol.* 1997;146:713-20
- Yang Q, Khoury MJ, Friedman JM, Flanders WD. On the use of population attributable fraction to determine sample size for case-control studies of gene-environment interaction. *Epidemiology.* 2003;14:161-7

## [MATÉRIEL & MÉTHODES]

- Antonarakis SE. Recommendations for a nomenclature system for human gene mutations. Nomenclature Working Group. *Hum Mutat.* 1998;11:1-3
- Arand et al. "A multiplex Polymerase Chain Reaction Protocol for the Simultaneous Analysis of the Glutathione S-Transferase GSTM1 and GSTT1 Polymorphisms", *Analytical Biochemistry* 1996; 236:184-6
- Barbaux S, Poirier O, Cambien F. Use of degenerate oligonucleotide primed PCR (DOP-PCR) for the genotyping of low-concentration DNA samples. *J Mol Med.* 2001;79:329-32
- Bouyer J, Hémon D, Cordier S, Derriennic F, Stücker I, Stengel B, Clavel J. *Epidémiologie Principes et méthodes quantitatives.* Les Editions Inserm, 1993
- Clark JD, Mossey PA, Sharp L, Little J. Socioeconomic status and orofacial clefts in Scotland, 1989 to 1998. *Cleft Palate Craniofac J.* 2003;40:481-5
- Cordier S, Bergeret A, Goujard J, Ha MC, Ayme S, Bianchi F, Calzolari E, De Walle HE, Knill-Jones R, Candela S, Dale I, Dananche B, de Vigan C, Fevotte J, Kiel G, Mandereau L. Congenital malformation and maternal occupational exposure to glycol ethers. Occupational Exposure and Congenital Malformations Working Group. *Epidemiology.* 1997; 8:355-63
- Frelut ML, de Courcy GP, Christides JP et al. Relationship between maternal folate status and foetal hypotrophy in a population with a good socio-economical level. *Int J Vitam Nutr Res* 1995;65:267-71
- Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, Boers GJ, den Heijer M, Kluijtmans LA, van den Heuvel LP, et al. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet.* 1995;10:111-3
- Garte S, Crosti F. A nomenclature system for metabolic gene polymorphisms. *IARC Sci Publ.* 1999;148:5-12
- Ingelman-Sundberg M, Oscarson M, Daly AK, Garte S, Nebert DW. Human cytochrome P-450 (CYP) genes: a web page for the nomenclature of alleles. *Cancer Epidemiol Biomarkers Prev.* 2001;10:1307-8
- Little J, Bradley L, Bray MS, Clyne M, Dorman J, Ellsworth DL, Hanson J, Khoury M, Lau J, O'Brien TR, Rothman N, Stroup D, Taioli E, Thomas D, Vainio H, Wacholder S, Weinberg C. Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. *Am J Epidemiol.* 2002;156:300-10
- Lorente C, Cordier S, Goujard J, Ayme S, Bianchi F, Calzolari E, De Walle HE, Knill-Jones R. Tobacco and alcohol use during pregnancy and risk of oral clefts. Occupational Exposure and Congenital Malformation Working Group. *Am J Public Health.* 2000; 90:415-9
- Nebert DW. Recommendations for the nomenclature of human alleles: relevance to ecogenetics, pharmacogenetics and molecular epidemiology. *Pharmacogenetics.* 2000, 10: 279-90
- Saal HM. Classification and Description of nonsyndromic Clefts. Chapter 5 in Wyszynski DF, ed. *Cleft lip and palate. From Origin to Treatment.* Oxford, Oxford University Press, 2002, p47-65
- Saiki RK, Bugawan TL, Horn GT, Mullis KB, Erlich HA. Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. *Nature.* 1986;324:163-6
- Sarmanova J, Benesova K, Gut I, Nedelcheva-Kristensen V, Tynkova L, Soucek P. Genetic polymorphisms of biotransformation enzymes in patients with Hodgkin's and non-Hodgkin's lymphomas. *Hum Mol Genet.* 2001;10:1265-73

## [RESULTATS] Rôle d'un polymorphisme du gène *ADH1C* en interaction avec la consommation maternelle d'alcool sur le risque de fentes orales non syndromiques

- Andersson S, Halmesmaki E, Koivusalo M, Lapatto R, Ylikorkala O. Placental alcohol metabolism in chronic alcohol abuse. *Biol Neonate* 1989; 56:90-3
- Beiraghi S, Foroud T, Diouhy S, Bixler D, Conneally PM, Delozier-Blanchet D, Hodes ME. Possible localization of a major gene for cleft lip and palate to 4q. *Clin Genet* 1994; 46(3):255-6
- Borras E, Coutelle C, Rosell A, Fernandez-Muixi F, Broch M, Crosas B, Hjelmqvist L, Lorenzo A, Gutierrez C, Santos M, Szczepanek M, Heilig M, Quattrocchi P, Farres J, Vidal F, Richart C, Mach T, Bogdal J, Jornvall H, Seitz HK, Couzigou P, Pares X. Genetic polymorphism of alcohol dehydrogenase in europeans: the ADH2\*2 allele decreases the risk for alcoholism and is associated with ADH3\*1. *Hepatology*. 2000;31:984-9
- Bosron WF, Li TK. Genetic polymorphism of human liver alcohol and aldehyde dehydrogenases, and their relationship to alcohol metabolism and alcoholism. *Hepatology* 1986; 6(3): 502-10
- Brennan P, Lewis S, Hashibe M, Bell DA, Boffetta P, Bouchardy C, Caporaso N, Chen C, Coutelle C, Diehl SR, Hayes RB, Olshan AF, Schwartz SM, Sturgis EM, Wei Q, Zavras AI, Benhamou S. Pooled analysis of alcohol dehydrogenase genotypes and head and neck cancer: a HuGE review. *Am J Epidemiol*. 2004 ;159:1-16
- Chen CC, Lu RB, Chen YC, Wang MF, Chang YC, Li TK, Yin SJ. Interaction between the functional polymorphisms of the alcohol-metabolism genes in protection against alcoholism. *Am J Hum Genet*. 1999;65:795-807
- Clayton D, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet*. 2001;358:1356-60
- Coutelle C, Ward PJ, Fleury B, Quattrocchi P, Chambrin H, Iron A, Couzigou P, Cassaigne A. Laryngeal and oropharyngeal cancer, and alcohol dehydrogenase 3 and glutathione S-transferase M1 polymorphisms. *Hum Genet*. 1997;99:319-25
- Coutelle C, Ward PJ, Quattrocchi P, Fleury B. Population distribution of alcohol dehydrogenase class I in France: comparison with other populations, and distribution with respect to gender and age. The French Group for Research on Alcohol and Liver. *Alcohol Alcohol*. 1998;33:173-83
- Crabb DW, Matsumoto M, Chang D, You M. Overview of the role of alcohol dehydrogenase and aldehyde dehydrogenase and their variants in the genesis of alcohol-related pathology. *Proc Nutr Soc*. 2004;63:49-63
- Duester G, Shean ML, McBride MS, Stewart MJ. Retinoic acid response element in the human alcohol dehydrogenase gene ADH3: implications for regulation of retinoic acid synthesis. *Mol Cell Biol*. 1991;11:1638-46
- Expertise Collective INSERM. Métabolisme et toxicocinétique chez l'animal. In : Ethers de glycol: quels risques pour la santé? Les Editions INSERM, 1999:21-38
- Freudenheim JL, Ambrosone CB, Moysich KB, Vena JE, Graham S, Marshall JR, Muti P, Laughlin R, Nemoto T, Harty LC, Crits GA, Chan AW, Shields PG. Alcohol dehydrogenase 3 genotype modification of the association of alcohol consumption with breast cancer risk. *Cancer Causes Control*. 1999;10:369-77
- Hard ML, Einarson TR, Koren G. The role of acetaldehyde in pregnancy outcome after prenatal alcohol exposure. *Ther Drug Monit*. 2001;23:427-34
- Harty LC, Caporaso NE, Hayes RB, Winn DM, Bravo-Otero E, Blot WJ, Kleinman DV, Brown LM, Armenian HK, Fraumeni JF Jr, Shields PG. Alcohol dehydrogenase 3 genotype and risk of oral cavity and pharyngeal cancers. *J Natl Cancer Inst*. 1997;89:1698-705
- Hoog JO, Heden LO, Larsson K, Jornvall H, von Bahr-Lindstrom H. The gamma 1 and gamma 2 subunits of human liver alcohol dehydrogenase. cDNA structures, two amino acid replacements, and compatibility with changes in the enzymatic properties. *Eur J Biochem*. 1986;159:215-8
- Jugessur A, Lie RT, Wilcox AJ, Murray JC, Taylor JA, Saugstad OD, Vindenes HA, Abyholm FE. Cleft palate, transforming growth factor alpha gene variants, and maternal exposures: assessing gene-environment interactions in case-parent triads. *Genet Epidemiol*. 2003;25:367-74
- McCarver DG, Thomasson HR, Martier SS, Sokol RJ, Li T. Alcohol dehydrogenase-2\*3 allele protects against alcohol-related birth defects among African Americans. *J Pharmacol Exp Ther*. 1997;283:1095-101
- Mitchell LE, Healey SC, Chenevix-Trench G. Evidence for an association between nonsyndromic cleft lip with or without cleft palate and a gene located on the long arm of chromosome 4. *Am J Hum Genet*. 1995;57:1130-6
- Mitchell LE, Murray JC, O'Brien S, Christensen K. Evaluation of two putative susceptibility loci for oral clefts in the Danish population. *Am J Epidemiol*. 2001;153:1007-15
- OMIM - Online Mendelian Inheritance in Man (<http://www.ncbi.nlm.nih.gov/omim/>).
- Osier M, Pakstis AJ, Kidd JR, Lee JF, Yin SJ, Ko HC, Edenberg HJ, Lu RB, Kidd KK. Linkage disequilibrium at the ADH2 and ADH3 loci and risk of alcoholism. *Am J Hum Genet*. 1999;64:1147-57
- Osier MV, Pakstis AJ, Goldman D, Edenberg HJ, Kidd JR, Kidd KK. A proline-threonine substitution in codon 351 of ADH1C is common in Native Americans. *Alcohol Clin Exp Res*. 2002;26:1759-63
- Raimondi S, Benhamou S, Coutelle C, Garte S, Hayes R, Kiemeny L, Lazarus P, Marchand LL, Morita S, Povey A, Romkes M, Zijno A, Taioli E. Association of metabolic gene polymorphisms with alcohol consumption in controls. *Biomarkers*. 2004;9:180-9
- Romitti PA, Lidral AC, Munger RG, Daack-Hirsch S, Burns TL, Murray JC. Candidate genes for nonsyndromic cleft lip and palate and maternal cigarette smoking and alcohol consumption: evaluation of genotype-environment interactions from a population-based case-control study of orofacial clefts. *Teratology*. 1999;59:39-50

- Stoler JM, Ryan LM, Holmes LB. Alcohol dehydrogenase 2 genotypes, maternal alcohol use, and infant outcome. *J Pediatr*. 2002;141:780-5
- Umbach DM, Weinberg CR. The use of case-parent triads to study joint effects of genotype and exposure. *Am J Hum Genet*. 2000;66:251-61
- van Ooij C, Snyder RC, Paepker BW, Duester G. Temporal expression of the human alcohol dehydrogenase gene family during liver development correlates with differential promoter activation by hepatocyte nuclear factor 1, CCAAT/enhancer-binding protein alpha, liver activator protein, and D-element-binding protein. *Mol Cell Biol*. 1992;12:3023-31
- Viljoen DL, Carr LG, Foroud TM, Brooke L, Ramsay M, Li TK. Alcohol dehydrogenase-2\*2 allele is associated with decreased prevalence of fetal alcohol syndrome in the mixed-ancestry population of the Western Cape Province, South Africa. *Alcohol Clin Exp Res*. 2001;25:1719-22
- Weinberg CR, Wilcox AJ, Lie RT. A log-linear approach to case-parent-triad data: assessing effects of disease genes that act either directly or through maternal effects and that may be subject to parental imprinting. *Am J Hum Genet*. 1998;62:969-78
- Weinberg CR. Allowing for missing parents in genetic studies of case-parent triads. *Am J Hum Genet*. 1999a;64:1186-93
- Weinberg CR. Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. *Am J Hum Genet*. 1999b;65:229-35
- Whitfield JB. Alcohol dehydrogenase and alcohol dependence: variation in genotype-associated risk between populations. *Am J Hum Genet*. 2002;71:1247-50
- Yang ZN, Davis GJ, Hurley TD, Stone CL, Li TK, Bosron WF. Catalytic efficiency of human alcohol dehydrogenases for retinol oxidation and retinal reduction. *Alcohol Clin Exp Res*. 1994;18:587-91
- Zavras AI, Wu T, Laskaris G, Wang YF, Cartos V, Segas J, Lefantzis D, Joshipura K, Douglass CW, Diehl SR. Interaction between a single nucleotide polymorphism in the alcohol dehydrogenase 3 gene, alcohol consumption and oral cancer risk. *Int J Cancer*. 2002;97:526-30
- Zgombic-Knight M, Satre MA, Duester G. Differential activity of the promoter for the human alcohol dehydrogenase (retinol dehydrogenase) gene ADH3 in neural tube of transgenic mouse embryos. *J Biol Chem*. 1994;269:6790-5

## [RESULTATS] Les génotypes *MTHFR C677T* de l'enfant et de la mère, l'apport maternel en folates et le risque de fentes orales non syndromiques

- Bailey LB, Gregory JF 3rd. Polymorphisms of methylenetetrahydrofolate reductase and other enzymes: metabolic significance, risks and impact on folate requirement. *J Nutr* 1999;129:919-22
- Berry RJ, Li Z, Erickson JD, Li S, Moore CA, Wang H, Mulinare J, Zhao P, Wong LY, Gindler J, Hong SX, Correa A. Prevention of neural-tube defects with folic acid in China. China-U.S. Collaborative Project for Neural Tube Defect Prevention. *N Engl J Med*. 1999;341:1485-90
- Botto LD, Yang Q. 5,10-Methylenetetrahydrofolate reductase gene variants and congenital anomalies: a HuGE review. *Am J Epidemiol*. 2000;151:862-77
- Chango A, Potier De Courcy G, Boisson F, Guillaud JC, Barbe F, Perrin MO, Christides JP, Rabhi K, Pfister M, Galan P, Hercberg S, Nicolas JP. 5,10-methylenetetrahydrofolate reductase common mutations, folate status and plasma homocysteine in healthy French adults of the Supplémentation en Vitamines et Minéraux Antioxydants (SU.VI.MAX) cohort. *Br J Nutr*. 2000;84:891-6
- Christensen B, Arbour L, Tran P, Leclerc D, Sabbaghian N, Platt R, Gilfix BM, Rosenblatt DS, Gravel RA, Forbes P, Rozen R. Genetic polymorphisms in methylenetetrahydrofolate reductase and methionine synthase, folate levels in red blood cells, and risk of neural tube defects. *Am J Med Genet*. 1999;84:151-7
- Czeizel AE, Dudas I. Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. *N Engl J Med* 1992; 327:1832-35
- de Bree A, van Dusseldorp M, Brouwer IA, van het Hof KH, Steegers-Theunissen RP. Folate intake in Europe: recommended, actual and desired intake. *Eur J Clin Nutr*. 1997;51:643-60
- Faure-Delaneuf L, Quere I, Chasse JF, Guerassimenko O, Lesaulnier M, Bellet H, Zittoun J, Kamoun P, Cohen D. Methylenetetrahydrofolate reductase thermolabile variant and human longevity. *Am J Hum Genet*. 1997;60:999-1001
- Favier JC, Ireland-Ripert J, Toque C, Feinberg M. Répertoire général des aliments. Table de composition. (General repertory of food. Table of composition). Technique & Documentation, Paris 1995
- Finnell RH, Shaw GM, Lammer EJ, Brandl KL, Carmichael SL, Rosenquist TH. Gene-nutrient interactions: importance of folates and retinoids during early embryogenesis. *Toxicol Appl Pharmacol*. 2004;198:75-85
- Frelut ML, de Courcy GP, Christides JP, Blot P, Navarro J. Relationship between maternal folate status and foetal hypotrophy in a population with a good socio-economical level. *Int J Vitam Nutr Res*. 1995;65:267-71
- Gaspar DA, Pavanello RC, Zatz M, Passos-Bueno MR, Andre M, Steman S, Wyszynski DF, Mاتيولli SR. Role of the C677T polymorphism at the MTHFR gene on risk to nonsyndromic cleft lip with/without cleft palate: results from a case-control study in Brazil. *Am J Med Genet*. 1999;87:197-9
- Gueant JL, Gueant-Rodriguez RM, Anello G, Bosco P, Brunaud L, Romano C, Ferri R, Romano A, Candito M, Namour B. Genetic determinants of folate and vitamin B12 metabolism: a common pathway in neural tube defect and Down syndrome? *Clin Chem Lab Med*. 2003;41:1473-7
- Jugessur A, Wilcox AJ, Lie RT, Murray JC, Taylor JA, Ulvik A, Drevon CA, Vindenes HA, Abyholm FE. Exploring the effects of methylenetetrahydrofolate reductase gene variants C677T and A1298C on the risk of orofacial clefts in 261 Norwegian case-parent triads. *Am J Epidemiol*. 2003;157:1083-91
- Mennen LI, de Courcy GP, Guillaud JC, Ducros V, Bertrais S, Nicolas JP, Maurel M, Zarebska M, Favier A, Franchisseur C, Hercberg S, Galan P. Homocysteine, cardiovascular disease risk factors, and habitual diet in the French Supplémentation with Antioxydant Vitamins and Minerals Study. *Am J Clin Nutr*. 2002;76:1279-89
- Mornet E, Muller F, Lenvoise-Furet A, Delezoide AL, Col JY, Simon-Bouy B, Serre JL. Screening of the C677T mutation on the methylenetetrahydrofolate reductase gene in French patients with neural tube defects. *Hum Genet*. 1997;100:512-4
- MRC Vitamin Research Group. Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. *Lancet* 1991; 338:131-7.
- Rosenquist TH, Finnell RH. Genes, folate and homocysteine in embryonic development. *Proc Nutr Soc*. 2001;60:53-61
- Rouget F, Monfort C, Bahuaud M, Nelva A, Herman C, Francannet C, Robert-Gnansia E, Cordier S. Folates during the periconceptional period and the prevention of oral clefts: role of dietary intake and vitamin supplements in France. 2005 (sous presse)
- Shaw GM, Rozen R, Finnell RH, Todoroff K, Lammer EJ. Infant C677T mutation in MTHFR, maternal periconceptional vitamin use, and cleft lip. *Am J Med Genet*. 1998;80:196-8
- Shaw GM, Todoroff K, Finnell RH, Rozen R, Lammer EJ. Maternal vitamin use, infant C677T mutation in MTHFR, and isolated cleft palate risk. *Am J Med Genet*. 1999;85:84-5
- Shaw GM, Zhu H, Lammer EJ, Yang W, Finnell RH. Genetic variation of infant reduced folate carrier (A80G) and risk of orofacial and conotruncal heart defects. *Am J Epidemiol*. 2003;158:747-52
- Shotelersuk V, Ittiwut C, Siriwan P, Angspatt A. Maternal 677CT/1298AC genotype of the MTHFR gene as a risk factor for cleft lip. *J Med Genet*. 2003;40:e64 available.
- Thuillier L, Chadefaux-Vekemans B, Bonnefont JP, Kara A, Aupetit J, Rochette C, Montalescot G, Couty MC, Kamoun P, Ankri A. Does the polymorphism 677C-T of the 5,10-methylenetetrahydrofolate reductase gene contribute to homocysteine-related vascular disease? *J Inherit Metab Dis*. 1998;21:812-22
- Umbach DM, Weinberg CR. The use of case-parent triads to study joint effects of genotype and exposure. *Am J Hum Genet*. 2000;66:251-61
- van Rooij IA, Vermeij-Keers C, Kluijtmans LA, Ocke MC, Zielhuis GA, Goorhuis-Brouwer SM, van der Biezen JJ, Kuijpers-Jagtman AM, Steegers-Theunissen RP. Does the interaction between maternal folate intake and the methylenetetrahydrofolate reductase polymorphisms affect the risk of cleft lip with or without cleft palate? *Am J Epidemiol*. 2003;157:583-91

- Vieira AR, Murray JC, Trembath D, Orioli IM, Castilla EE, Cooper ME, Marazita ML, Lennon-Graham F, Speer M. Studies of reduced folate carrier 1 (RFC1) A80G and 5,10-methylenetetrahydrofolate reductase (MTHFR) C677T polymorphisms with neural tube and orofacial cleft defects. *Am J Med Genet A*. 2005 (sous presse)
- Weinberg CR, Wilcox AJ, Lie RT. A log-linear approach to case-parent-triad data: assessing effects of disease genes that act either directly or through maternal effects and that may be subject to parental imprinting. *Am J Hum Genet*. 1998;62:969-78
- Weinberg CR. Allowing for missing parents in genetic studies of case-parent triads. *Am J Hum Genet*. 1999a;64:1186-93
- Weinberg CR. Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. *Am J Hum Genet*. 1999b;65:229-35
- Wilcken B, Bamforth F, Li Z, Zhu H, Ritvanen A, Renlund M, Stoll C, Alembik Y, Dott B, Czeizel AE, Gelman-Kohan Z, Scarano G, Bianca S, Ettore G, Tenconi R, Bellato S, Scala I, Mutchinick OM, Lopez MA, de Walle H, Hofstra R, Joutchenko L, Kavteladze L, Bermejo E, Martinez-Frias ML, Gallagher M, Erickson JD, Vollset SE, Mastroiacovo P, Andria G, Botto LD. Geographical and ethnic variation of the 677C>T allele of 5,10 methylenetetrahydrofolate reductase (MTHFR): findings from over 7000 newborns from 16 areas world wide. *J Med Genet*. 2003;40:619-25
- Zetterberg H. Methylenetetrahydrofolate reductase and transcobalamin genetic polymorphisms in human spontaneous abortion: biological and clinical implications. *Reprod Biol Endocrinol*. 2004;2:7

## [RESULTATS] L'exposition professionnelle de la mère aux solvants organiques en interaction avec des gènes de métabolisme/détoxication, sur le risque de fentes orales non syndromiques

- Bianchi F, Cianciulli D, Pierini A, Seniori Costantini A. Congenital malformations and maternal occupation: a registry based case-control study. *Occup Environ Med*. 1997;54:223-8
- Bove FJ, Fulcomer MC, Klotz JB, Esmart J, Dufficy EM, Savrin JE. Public drinking water contamination and birth outcomes. *Am J Epidemiol*. 1995;141:850-62
- Carpenter SP, Lasker JM, Raucy JL. Expression, induction, and catalytic activity of the ethanol-inducible cytochrome P450 (CYP2E1) in human fetal liver and hepatocytes. *Mol Pharmacol*. 1996;49:260-8
- Cordier S, Bergeret A, Goujard J, Ha MC, Ayme S, Bianchi F, Calzolari E, De Walle HE, Knill-Jones R, Candela S, Dale I, Dananche B, de Vigan C, Fevotte J, Kiel G, Mandereau L. Congenital malformation and maternal occupational exposure to glycol ethers. Occupational Exposure and Congenital Malformations Working Group. *Epidemiology*. 1997;8:355-63
- Cordier S, Szabova E, Fevotte J, Bergeret A, Plackova S, Mandereau L. Congenital malformations and maternal exposure to glycol ethers in the Slovak Republic. *Epidemiology*. 2001;12:592-3
- Expertise Collective INSERM. Etudes épidémiologiques. In Ethers de glycol: quels risques pour la santé? Les Editions INSERM, 1999,p171-87
- Garcia AM, Fletcher T. Maternal occupation in the leather industry and selected congenital malformations. *Occup Environ Med*. 1998;55:284-6
- Garte S, Crosti F. A nomenclature system for metabolic gene polymorphisms. *IARC Sci Publ*. 1999;148:5-12
- Garte S, Gaspari L, Alexandrie AK, Ambrosone C, Autrup H, Autrup JL, Baranova H, Bathum L, Benhamou S, Boffetta P, Bouchardy C, Breskvar K, Brockmoller J, Cascorbi I, Clapper ML, Coutelle C, Daly A, Dell'Omo M, Dolzan V, Dresler CM, Fryer A, Haugen A, Hein DW, Hildesheim A, Hirvonen A, Hsieh LL, Ingelman-Sundberg M, Kalina I, Kang D, Kihara M, Kiyohara C, Kremers P, Lazarus P, Le Marchand L, Lechner MC, van Lieshout EM, London S, Manni JJ, Maugard CM, Morita S, Nazar-Stewart V, Noda K, Oda Y, Parl FF, Pastorelli R, Persson I, Peters WH, Rannug A, Rebbeck T, Risch A, Roelandt L, Romkes M, Ryberg D, Salagovic J, Schoket B, Seidegard J, Shields PG, Sim E, Sinnet D, Strange RC, Stucker I, Sugimura H, To-Figueras J, Vineis P, Yu MC, Taioli E. Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev*. 2001;10:1239-48
- Hakkola J, Pelkonen O, Pasanen M, Raunio H. Xenobiotic-metabolizing cytochrome P450 enzymes in the human fetoplacental unit: role in intrauterine toxicity. *Crit Rev Toxicol*. 1998;28:35-72
- Johnson PD, Goldberg SJ, Mays MZ, Dawson BV. Threshold of trichloroethylene contamination in maternal drinking waters affecting fetal heart development in the rat. *Environ Health Perspect*. 2003;111:289-92
- Kersemakers WM, Roeleveld N, Zielhuis GA. Reproductive disorders among hairdressers. *Epidemiology*. 1997;8:396-401
- Laumon B, Martin JL, Collet P, Bertucat I, Verney MP, Robert E. Exposure to organic solvents during pregnancy and oral clefts: a case-control study. *Reprod Toxicol*. 1996;10:15-9
- Lorente C, Cordier S, Bergeret A, De Walle HE, Goujard J, Ayme S, Knill-Jones R, Calzolari E, Bianchi F. Maternal occupational risk factors for oral clefts. Occupational Exposure and Congenital Malformation Working Group. *Scand J Work Environ Health*. 2000;26:137-45
- Multigner L, Catala M, Cordier S, Delaforge M, Fenaux P, Garnier R, Rico-Lattes I, Vasseur P; INSERM (French Institute of Health and Medical Research). The INSERM expert review on glycol ethers: findings and recommendations. *Toxicol Lett*. 2005;156:29-37
- Raijmakers MT, Steegers EA, Peters WH. Glutathione S-transferases and thiol concentrations in embryonic and early fetal tissues. *Hum Reprod*. 2001;16:2445-50
- Rylander L, Axmon A, Toren K, Albin M. Reproductive outcome among female hairdressers. *Occup Environ Med*. 2002;59:517-22
- Schardein JL. Industrial solvents. In: Chemically induced birth defects 2nd edition, revised and expanded. New York, Marcel Dekker, 1993, p751-75
- Shaw GM, Nelson V, Iovannisci DM, Finnell RH, Lammer EJ. Maternal occupational chemical exposures and biotransformation genotypes as risk factors for selected congenital anomalies. *Am J Epidemiol*. 2003;157:475-84
- Smith CA, Smith G, Wolf CR. Genetic polymorphisms in xenobiotic metabolism. *Eur J Cancer*. 1994;30A:1921-35
- Smith G, Stanley LA, Sim E, Strange RC, Wolf CR. Metabolic polymorphisms and cancer susceptibility. *Cancer Surv*. 1995;25:27-65
- Stubbins MJ, Wolf CR. Chapter 21. Additional polymorphisms and cancer. *IARC Sci Publ*. 1999:271-302
- Wennborg H, Magnusson LL, Bonde JP, Olsen J. Congenital malformations related to maternal exposure to specific agents in biomedical research laboratories. *J Occup Environ Med*. 2005;47:11-9
- Wormhoudt LW. Inter-individual Variability in Drug Metabolism. Thesis, Amsterdam, 1998

## [CONCLUSION GÉNÉRALE]

- Andrieu N, Goldstein AM. The case-combined-control design was efficient in detecting gene-environment interactions. *J Clin Epidemiol*. 2004;57:662-71
- Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet*. 2004;74:106-20
- Clayton D, Jones H. Transmission/disequilibrium tests for extended marker haplotypes. *Am J Hum Genet*. 1999;65:1161-9
- Clayton D, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet*. 2001;358:1356-60
- Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. *Lancet*. 2003;361:865-72
- Cordell HJ, Barratt BJ, Clayton DG. Case/pseudocontrol analysis in genetic association studies: A unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. *Genet Epidemiol*. 2004;26:167-85
- Curtis D, North BV, Sham PC. Use of an artificial neural network to detect association between a disease and multiple marker genotypes. *Ann Hum Genet*. 2001;65:95-107
- Editorial de *Cancer Epidemiology, Biomarkers & Prevention* 2004: Rebbeck TR, Martinez ME, Sellers TA, Shields PG, Wild CP, Potter JD. Genetic variation and cancer: improving the environment for publication of association studies. *Cancer Epidemiol Biomarkers Prev*. 2004;13:1985-6
- Editorial, *Freely Associating*. *Nature Genetics* 1999 ; 22 :1-2
- Epstein MP, Veal CD, Trembath RC, Barker JN, Li C, Satten GA. Genetic association analysis using data from triads and unrelated subjects. *Am J Hum Genet*. 2005;76:592-608
- Foulkes AS, Reilly M, Zhou L, Wolfe M, Rader DJ. Mixed modelling to characterize genotype-phenotype associations. *Stat Med*. 2005;24:775-89
- Gauderman WJ. Sample size requirements for matched case-control studies of gene-environment interaction. *Stat Med*. 2002;21:35-50
- Goodman SN, Berlin JA. The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Ann Intern Med*. 1994;121:200-6
- Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet*. 2001;29:306-9
- Ioannidis JP, Rosenberg PS, Goedert JJ, O'Brien TR; International Meta-analysis of HIV Host Genetics. Commentary: meta-analysis of individual participants' data in genetic epidemiology. *Am J Epidemiol*. 2002;156:204-10
- Lake SL, Laird NM. Tests of gene-environment interaction for case-parent triads with general environmental exposures. *Ann Hum Genet*. 2004;68:55-64
- Little J, Bradley L, Bray MS, Clyne M, Dorman J, Ellsworth DL, Hanson J, Khoury M, Lau J, O'Brien TR, Rothman N, Stroup D, Taioli E, Thomas D, Vainio H, Wacholder S, Weinberg C. Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. *Am J Epidemiol*. 2002;156:300-10
- Mitchell LE, Beaty TH, Lidral AC, Munger RG, Murray JC, Saal HM, Wyszynski DF; International Consortium for Oral Clefts Genetics. Guidelines for the design and analysis of studies on nonsyndromic cleft lip and cleft palate in humans: summary report from a Workshop of the International Consortium for Oral Clefts Genetics. *Cleft Palate Craniofac J*. 2002;39:93-100
- Munafò M. R, Flint J. Meta-analysis of genetic association studies. *Trends Genet*. 2004;20:439-44
- Nagelkerke NJ, Hoebee B, Teunis P, Kimman TG. Combining the transmission disequilibrium test and case-control methodology using generalized logistic regression. *Eur J Hum Genet*. 2004;12:964-70
- Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. *Am J Hum Genet*. 2004;75:353-62
- Rebbeck TR, Spitz M, Wu X. Assessing the function of genetic variants in candidate gene association studies. *Nat Rev Genet*. 2004;5:589-97
- Ritchie MD, White BC, Parker JS, Hahn LW, Moore JH. Optimization of neural network architecture using genetic programming improves detection and modeling of gene-gene interactions in studies of human diseases. *BMC Bioinformatics*. 2003;4:28
- Schaid DJ. Evaluating associations of haplotypes with traits. *Genet Epidemiol*. 2004;27:348-64
- Slager SL, Schaid DJ, Wang L, Thibodeau SN. Candidate-gene association studies with pedigree data: controlling for environmental covariates. *Genet Epidemiol*. 2003;24:273-83
- Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst*. 2004;96:434-42
- Zeiger JS, Beaty TH, Liang KY. Oral Clefts, Maternal Smoking, and TGFA: A Meta-Analysis of Gene-Environment Interaction. *Cleft Palate Craniofac J*. 2005;42:58-63
- Zhang H, Bonney G. Use of classification trees for association studies. *Genet Epidemiol*. 2000;19:323-32
- Zhang W, Collins A, Morton NE. Does haplotype diversity predict power for association mapping of disease susceptibility? *Hum Genet*. 2004;115:157-64
- Zhao H, Pfeiffer R, Gail MH. Haplotype analysis in population genetics and association studies. *Pharmacogenomics*. 2003;4:171-8



# ANNEXES



# Index des tableaux

Tableau 1: Prévalence des fentes orales en France .....	16
Tableau 2 : Etudes cas-témoins évaluant l'association entre la prise de multivitamines par la mère pendant la grossesse et le risque de fentes orales .....	21
Tableau 3: Résultats des deux méta-analyses s'intéressant à l'association entre la consommation maternelle de tabac pendant la grossesse et le risque de fentes orales non syndromiques .....	23
Tableau 4: Risques de FL/P et de FP associés aux plus fortes doses d'alcool ingérées par la mère pendant la grossesse, des études épidémiologiques mesurant l'association entre alcool et risque de fentes orales .....	24
Tableau 5: Etudes d'association du gène <i>TGF<math>\alpha</math></i> avec le risque de fentes orales (populations caucasiennes) .....	29
Tableau 6: Etudes d'association du gène <i>TGF<math>\beta</math>3</i> avec le risque de fentes orales (populations caucasiennes) .....	30
Tableau 7: Etudes d'association du gène <i>MSX1</i> avec le risque de fentes orales (populations caucasiennes) .....	31
Tableau 8: Etudes d'association du gène <i>MTHFR</i> avec le risque de fentes orales (populations caucasiennes) .....	33
Tableau 9: Résumé des résultats des études s'intéressant à la consommation maternelle de tabac en interaction avec un gène candidat .....	39
Tableau 10 : Etudes épidémiologiques d'interaction gène-environnement dans l'étiologie des fentes orales .....	40
Tableau 11: Distribution dans la population des génotypes selon le statut de la maladie .....	48
Tableau 12: Analyse non appariée des génotypes transmis et non transmis, pour la construction du test HRR de Falk et Rubinstein (1987 - Cas d'un gène candidat à 2 allèles a et A, où A est l'allèle d'intérêt) .....	53
Tableau 13: Analyse non appariée des allèles transmis et non transmis, pour la construction des tests étendus du HRR ....	54
Tableau 14: Analyse appariée des allèles transmis et non transmis, pour la construction du TDT .....	55
Tableau 15: Tableau 2x4 des effets propres et conjoint des facteurs génétique G et environnemental E, par l'approche cas-témoins .....	66
Tableau 16: Effet conjoint des facteurs génétique G et environnemental E, par l'approche cas-seuls .....	68
Tableau 17: Caractéristiques des approches disponibles dans l'étude évaluant une interaction gène-environnement (notée GxE, avec le gène G et le facteur environnemental E) .....	73
Tableau 18: Support de prélèvement sanguin par catégorie de sujets (France, 1998-2001) .....	82
Tableau 19 : Effectifs des informations génétiques disponibles .....	85
Tableau 20: Description des populations cas non syndromiques et témoins (France, 1998-2001) .....	89
Tableau 21: Propriétés des formes polymorphiques des gènes humains <i>ADH1B</i> et <i>ADH1C</i> .....	98
Tableau 22: Description des populations de cas non syndromiques et de témoins (France, 1998-2001) .....	103
Tableau 23: Distribution des génotypes de l' <i>ADH1C</i> parmi les enfants et parents des populations cas et témoins (France, 1998-2001) .....	104
Tableau 24: Distribution des génotypes de l' <i>ADH1C</i> des familles d'enfants atteints de fentes orales non syndromiques, selon le profil parental (France, 1998-2001) .....	104
Tableau 25: Estimations des effets des génotypes de l' <i>ADH1C</i> de l'enfant et de la mère sur le risque de fentes orales nonsyndromiques, par l'approche cas-parents en présence de triades incomplètes (France, 1998-2001) .....	105
Tableau 26: Estimations des effets des génotypes de l' <i>ADH1C</i> de l'enfant et de la mère sur le risque de fentes orales nonsyndromiques, par l'approche cas-parents en présence de triades incomplètes, pour les familles d'origine européenne (France, 1998-2001) .....	105
Tableau 27: Estimations des effets des génotypes de l' <i>ADH1C</i> de l'enfant et de la mère sur le risque de fentes orales nonsyndromiques, par l'approche cas-parents en présence de triades incomplètes, incluant ou excluant diverses sous-populations (France, 1998-2001) .....	106
Tableau 28: Description des populations cas non syndromiques et témoins (France, 1998-2001) .....	116
Tableau 29: Estimations des effets du génotype <i>MTHFR C677T</i> de l'enfant sur le risque de fentes orales nonsyndromiques, par l'approche cas-témoins (France, 1998-2001) .....	117
Tableau 30: Estimations des effets des génotypes <i>MTHFR C677T</i> de l'enfant et de la mère sur les risques de fentes orales non syndromiques, par l'approche cas-parents (France, 1998-2001) .....	117

Tableau 31: Estimations des effets des génotypes <i>MTHFR C677T</i> de l'enfant et de la mère sur le risque de fentes orales nonsyndromiques, pour diverses conditions de population, par l'approche cas-parents (France, 1998-2001)..	118
Tableau 32: Estimations des effets du génotype <i>MTHFR C677T</i> de l'enfant sur le risque de FL/P pour chaque strate d'apport en folates de la mère, par l'approche cas-témoins (France, 1998-2001) .....	119
Tableau 33: Fréquence du génotype <i>MTHFR TT</i> rapportées dans la population française .....	122
Tableau 34: Résumé des risques de FL/P et de FP portés par les génotypes <i>MTHFR</i> sur chaque strate d'apport en folates, parmi les études épidémiologiques existantes incluant l'étude présente .....	124
Tableau 35: Etudes épidémiologiques s'intéressant au risque de fentes orales associé à l'exposition professionnelle des mères aux solvants.....	126
Tableau 36: Risques de FL/P et de FP associés à l'exposition professionnelle des mères aux solvants organiques ou à des activités extra-professionnelles impliquant possiblement la présence de solvants (France, 1998-2001) .....	131
Tableau 37: Risques de FL/P et de FP associés à l'exposition professionnelle des mères aux solvants oxygénés parmi les femmes ayant une activité professionnelle au premier trimestre de la grossesse, sachant que les femmes exposées aux solvants chlorés ou pétroliers sont exclues de l'analyse (France, 1998-2001) .....	132
Tableau 38: Interaction entre le polymorphisme du <i>CYP2E1 *RsaI</i> de l'enfant et l'exposition professionnelle et extra-professionnelles de la mère aux solvants sur le risque de fentes orales non syndromiques (France, 1998-2001) .....	133
Tableau 39: Interaction entre le polymorphisme <i>Ile349Val ADH1C</i> de l'enfant et l'exposition professionnelle et extra-professionnelles de la mère aux solvants sur le risque de fentes orales non syndromiques (France, 1998-2001) .....	134
Tableau 40: Risques de fentes orales non syndromiques associés à l'exposition professionnelle des mères aux solvants organiques ou à des activités extra-professionnelles impliquant possiblement la présence de solvants, stratifiés sur le génotype <i>GSTM1</i> de l'enfant (France, 1998-2001) .....	135
Tableau 41: Risques de fentes orales non syndromiques associés à l'exposition professionnelle des mères aux solvants organiques ou à des activités extra-professionnelles impliquant possiblement la présence de solvants, stratifiés sur le génotype <i>GSTT1</i> de l'enfant (France, 1998-2001) .....	136
Tableau 42: Professions des mères exposées aux solvants organiques (France, 1998-2001).....	139
Tableau 43: Puissance statistique calculée a posteriori d'un test d'interaction gène-environnement, pour les trois applications du projet.....	150
Tableau 44: Simulations par QUANTO de la taille de l'échantillon requise pour un test d'interaction gène-environnement par les schémas d'étude cas-témoins et cas-parents.....	151

# Glossaire

Allèle : forme d'un gène détectable à un locus

Autosomique : indépendant du chromosome sexuel.

Catabolisme : phase du métabolisme au cours de laquelle des molécules relativement grosses et complexes sont dégradées en molécules plus petites et plus simples. De l'énergie est libérée au cours de cette phase. (Autre phase du Métabolisme : Anabolisme)

Déséquilibre de liaison (*Linkage Disequilibrium*) : association entre des allèles à différents loci dans une population. Un déséquilibre de liaison existe car des allèles peuvent être physiquement proches, et ont ainsi tendance à être hérités ensemble.

Désordre mendélien (traduit littéralement de *Mendelian disorder*) : maladie génétique à hérédité mendélienne et causée par une unique mutation dans la structure de l'ADN, appelée aussi « maladie monogénique ».

Dominance (respectivement Récessivité) d'un allèle : état qui présente un phénotype semblable entre l'hétérozygotie et l'homozygotie de l'allèle (respectivement, le génotype ne contenant pas l'allèle).

Empreinte parentale : différence d'effet sur le phénotype de l'enfant d'un allèle hérité selon son origine paternelle ou maternelle.

Équilibre de Hardy-Weinberg, encore appelé équilibre panmictique : modèle théorique central de la génétique des populations. La notion d'équilibre dans le modèle de Hardy-Weinberg est soumise aux hypothèses/conditions suivantes : la population est panmictique (les couples se forment au hasard, et leurs gamètes se rencontrent au hasard. La population est très grande. Il ne doit y avoir ni sélection, ni mutation, ni migration (pas de perte/gain d'allèle). Il n'existe pas de croisement entre générations différentes. Dans ces conditions, la diversité génétique de la population se maintient et doit tendre vers un équilibre stable de la distribution génotypique. La loi de Hardy-Weinberg calcule des fréquences génotypiques à partir des fréquences alléliques, utilisant la formule  $p^2 + 2pq + q^2 = 1$  où  $p$  est la fréquence du gène normal et  $q$  celle du gène mutant.

Gène : séquence d'ADN (incluant des introns, des exons, des régions « régulatrices ») localisé par le locus. Les termes gène et locus sont couramment employés de manière équivalente.

Génotype : combinaison formée par les deux allèles hérités à un locus spécifique. Si les deux allèles sont les mêmes, le génotype est dit homozygote, s'ils sont différents le génotype est qualifié d'hétérozygote.

Haplotype : combinaison formée par les allèles d'une série de loci liés sur un même chromosome.

Locus (pluriel, loci) : position dans la séquence de l'ADN. Un locus peut être utilisé pour localiser un polymorphisme sur un seul nucléotide ou sur une région plus large de la séquence d'ADN

Marqueur : segment de l'ADN avec une localisation physique sur le chromosome identifiable. Un marqueur peut être un gène, un SNP ou une section de l'ADN.

Mutation : Modification soudaine, transmissible ou non, du matériel héréditaire. Conséquence d'une erreur d'écriture dans l'ADN, une mutation peut arriver spontanément, de façon naturelle, ou être induite par des agents dits "mutagènes", tels que des radiations ou des produits toxiques. La modification peut concerner toute une région du chromosome ou ne porter que sur une seule lettre de l'ADN, on parle alors de mutation ponctuelle. Il existe des mutations utiles, qui permettent par exemple de répondre efficacement aux agressions d'organismes pathogènes tels que des virus. Elles participent à l'évolution des espèces. D'autres mutations, au contraire, conduisent à un dysfonctionnement d'une protéine, et peuvent être responsables d'une maladie héréditaire. Les mutations somatiques touchent n'importe quelle cellule à l'exception des cellules germinales et ne sont alors pas héréditaires. En revanche, les mutations constitutionnelles concernent toutes les cellules, y compris les cellules reproductrices, et sont transmissibles à la descendance.

Panmixie : mariages aléatoires (indépendants des gènes) dans une population.

Pénétrance d'un génotype : probabilité d'être malade parmi les porteurs du génotype.

Phénotype : caractéristique observable de l'individu. Il peut s'agir de caractères morphologiques, biochimiques, physiologiques...

Pléiotropie : possibilité pour les polymorphismes d'avoir plus d'un effet spécifique phénotypique

Polymorphisme : existence d'au moins deux formes différentes du gène (ou allèles). Conventionnellement, nous parlons de polymorphisme dans une population si au moins deux des allèles sont prévalents à plus de 1%, sinon nous employons le terme de mutation.

Séquence : ensemble de plusieurs anomalies dérivées d'une anomalie ou d'un facteur mécanique unique, qui sont connus ou supposés. On parle aussi de cascade d'événements secondaires liés à l'événement initial. L'origine d'une séquence est généralement déterminée par plusieurs causes.

SNP (*Single Nucleotide Polymorphism*) : position sur le chromosome qui identifie une variation du code génétique sur une seule paire de nucléotide.

Syndrome : ensemble de plusieurs anomalies patho-génétiquement liées et qui ne représente pas une séquence. Un syndrome implique généralement une étiologie unique (exemple : la délétion d'un gène).

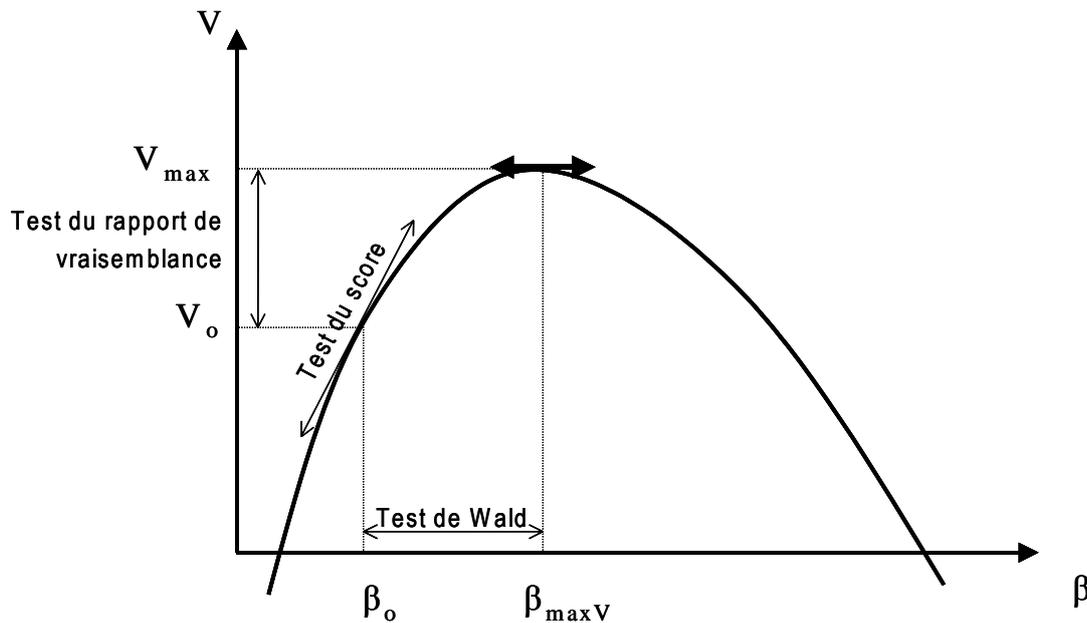
Recombinaison : Désigne généralement l'échange de matériel entre deux chromosomes homologues lors de la méiose. La recombinaison génétique assure la variabilité des espèces, c'est elle qui fait que les enfants d'un même couple ne seront jamais identiques (à l'exception des jumeaux).

# Annexe 1 : Tests associés à l'estimation du maximum de vraisemblance

Il existe trois façons de tester une hypothèse dans les modèles pour lesquels la vraisemblance est spécifiée :

- le test du rapport de vraisemblance (ou test de déviance)
- le test du score (ou du multiplicateur de Lagrange)
- le test de Wald

Ces trois tests s'illustrent sur un graphique représentant la fonction de la vraisemblance et son maximum.



La fonction atteint son maximum au point  $(\beta_{\max V}, V_{\max})$ . Le paramètre  $\beta_0$  est la valeur prise par  $\beta$  sous  $H_0$ . Les tests s'écrivent tels que :

$$H_0 : \beta = \beta_0$$

$$H_1 : \beta \neq \beta_0$$

Le principe des tests est de mesurer l'écart entre la situation sous  $H_0$  et la situation du maximum de vraisemblance. Cet écart peut s'exprimer de trois façons :

- l'écart entre les vraisemblances  $V_0$  et  $V_{\max}$
- la tangente à la courbe en  $\beta_0$ , comparée à l'horizontale de la situation du maximum de vraisemblance
- l'écart entre les paramètres  $\beta_{\max V}$  et  $\beta_0$

Les trois tests sont asymptotiquement équivalents. Le calcul du test de Wald s'appuie sur une hypothèse supplémentaire d'approximation (normalité de l'estimateur du maximum de vraisemblance). Elle permet de calculer facilement les intervalles de confiance, classiquement utilisé en épidémiologie, cependant ce test est le moins robuste des trois lorsque la condition asymptotique n'est pas optimale.

## Annexe 2 : Causes d'hospitalisation des témoins

Causes	Effectif
Gastroentérite	14
Diarrhée	3
Coqueluche	1
Varicelle	2
Naevus (jambe, cuisse, cuir chevelu, bras gauche, non précisé)	28
Angiome (scrotum, hanche)	4
Chirurgie sur résidu embryonnaire	2
Tumeur cutanée (nez)	2
Méningite (ou suspicion de méningite)	6
Otite	4
Bronchiolite	16
Bronchite	1
Asthme	1
Pneumopathie	6
Hernie inguinale ou ombilicale	6
Constipation	2
Pyélonéphrite (ou suspicion)	11
Infection urinaire	16
Kyste de la queue de sourcil	5
Ostéoarthrite	1
Kyste synovial du bras	1
Fibrochondrome	2
Frein de langue	2
Sténose du pylore	1
Aplasie du cuir chevelu	1
Infection streptocoque B, septicémie	1
Fièvre, hyperthermie, convulsions hyperthermiques	10
Malaise, collapse	2
Stagnation pondérale	1
Vomissement	1
Fracture du fémur	1
Trauma crânien	1
Allergie alimentaire	1
	<hr/>
	156

# Annexe 3 : Complément des résultats du gène *ADH1C* sur le risque de fentes orales, en interaction avec l'exposition maternelle à l'alcool

- 
1. Calcul du TDT à partir du Tableau 24, pour le polymorphisme *ADH1C Ile349Val*
  2. Interactions entre *ADH1C* et exposition maternelle à l'alcool
- 

## 1. Calcul du TDT à partir du Tableau 24, pour le polymorphisme *ADH1C Ile349Val*

Parmi les fentes orales non syndromiques (174 triades)

TDT	Allèle non transmis			
	*C	*T		
	<b>Allèle *C</b>	<b>105</b>	<b>60</b>	165
	<b>transmis *T</b>	<b>54</b>	<b>33</b>	87
		159	93	252
Méthode classique	Odds= 0,90		[ 0,62 - 1,30 ]	
Méthode Fleiss	std (Odds)= 0,17			
$(b-c)^2/(b+c)=$	$X^2(1)= 0,32$		p=0,57	

## 2. Interactions entre *ADH1C* et exposition maternelle à l'alcool dans l'étiologie des fentes orales

Le polymorphisme de l'*ADH1C* est étudié en interaction avec la consommation maternelle d'alcool (le mois précédant la conception et le 1<sup>er</sup> trimestre de grossesse) dans la 1<sup>ère</sup> application du projet, puis en interaction avec l'exposition professionnelle des mères aux solvants organiques, en particulier aux éthers de glycol, dans la 3<sup>ème</sup> application du projet.

Quelques expositions supplémentaires peuvent être étudiées en interaction avec le gène *ADH1C* :

- Consommation maternelle d'alcool au premier trimestre de grossesse
- Suppléments vitaminiques contenant de la vitamine A, ou retinol, au premier trimestre de la grossesse
- Consommation régulière de viande de foie par la mère

Les résultats d'interaction sont présentés sous la forme exponentielle des coefficients estimés (et non sous la forme d'analyses stratifiées).

<b>Suppléments vitaminiques contenant de la vitamine A, ou retinol (au premier trimestre de la grossesse)</b>									
Familles exposées :		n=17 dont 9 avec un enfant hétérozygote et 0 avec un enfant homozygote pour le variant							
Familles non exposées :		n=157 dont 68 avec un enfant hétérozygote et 17 avec un enfant homozygote pour le variant							
	<i>ADH1C</i>	Pseudo-sibs		Modèle log-linéaire Simple		Modèle log-linéaire Avec le génotype de la mère		Modèle log-linéaire Interaction avec le génotype de la mère	
		RR	IC95% Wald	RR	IC95% Wald	RR	IC95% Wald	RR	IC95% Wald
<b>Génotype de l'enfant</b>	<i>*1/*1</i>	1.00		1.00		1.00		1.00	
	<i>*1/*2</i>	0.77	0.51-1.15	0.77	xx	0.77	xx	0.83	xx
	<i>*2/*2</i>	0.75	0.36-1.55	0.75	xx	0.75	xx	0.74	xx
	Interaction: Exposition <i>*1/*2</i>	2.52	0.61-10.5	2.52	xx	2.52	xx		
	Interaction: Exposition <i>*2/*2</i>	xx	x	0.83	xx	0.83	xx		
	<i>p interaction</i>		0.45		x		x		
<b>Génotype de la mère</b>	<i>*1/*1</i>					1.00		1.00	
	<i>*1/*2</i>					1.33	xx	1.35	xx
	<i>*2/*2</i>					0.51	xx	0.57	xx
	Interaction: Exposition <i>*1/*2</i>							0.99	xx
	Interaction: Exposition <i>*2/*2</i>							0.00	xx
	<i>p interaction</i>								x

n : effectif ; RR : Risque Relatif ; IC95%Wald: Intervalle de Confiance à 95% (de Wald) ; *p* : degré de signification du test du rapport de vraisemblance

### Consommation régulière de viande de foie par la mère

Familles exposées : n=45 dont 21 avec un enfant hétérozygote et 2 avec un enfant homozygote pour le variant

Familles non exposées : n=126 dont 54 avec un enfant hétérozygote et 14 avec un enfant homozygote pour le variant

	ADH1C	Pseudo-sibs		Modèle log-linéaire Simple		Modèle log-linéaire Avec le génotype de la mère		Modèle log-linéaire Interaction avec le génotype de la mère	
		RR	IC95% Wald	RR	IC95% Wald	RR	IC95% Wald	RR	IC95% Wald
<b>Génotype de l'enfant</b>	<i>*1/*1</i>	1.00		1.00		1.00		1.00	
	<i>*1/*2</i>	0.78	0.50 -1.22	0.78	0.50 -1.22	0.78	0.50 -1.22	0.82	0.55 -1.20
	<i>*2/*2</i>	0.86	0.38 -1.92	0.86	0.38 -1.92	0.86	0.38 -1.92	0.73	0.35 -1.50
Interaction: Exposition	<i>*1/*2</i>	1.20	0.48 -2.97	1.20	0.48 -2.97	1.20	0.48 -2.97		
Interaction: Exposition	<i>*2/*2</i>	0.45	0.07 -2.94	0.45	0.07 -2.94	0.45	0.07 -2.94		
<i>p interaction</i>				0.40		0.40			
<b>Génotype de la mère</b>	<i>*1/*1</i>					1.00		1.00	
	<i>*1/*2</i>					1.32	0.85 -2.04	1.28	0.77 -2.13
	<i>*2/*2</i>					0.48	0.21 -1.06	0.44	0.17 -1.15
Interaction: Exposition	<i>*1/*2</i>							1.11	0.41 -3.03
Interaction: Exposition	<i>*2/*2</i>							1.33	0.24 -7.49
<i>p interaction</i>								0.94	

n : effectif ; RR : Risque Relatif ; IC95%Wald: Intervalle de Confiance à 95% (de Wald) ; p : degré de signification du test du rapport de vraisemblance

### Consommation maternelle d'alcool (au premier trimestre)

Familles exposées : n=32 dont 11 avec un enfant hétérozygote et 6 avec un enfant homozygote pour le variant

Familles non exposées : n=142 dont 66 avec un enfant hétérozygote et 11 avec un enfant homozygote pour le variant

	ADH1C	Pseudo-sibs		Modèle log-linéaire Simple		Modèle log-linéaire Avec le génotype de la mère		Modèle log-linéaire Interaction avec le génotype de la mère	
		RR	IC95% Wald	RR	IC95% Wald	RR	IC95% Wald	RR	IC95% Wald
<b>Génotype de l'enfant</b>	<i>*1/*1</i>	1.00		1.00		1.00		1.00	
	<i>*1/*2</i>	0.87	0.57 -1.33	0.87	0.57 -1.33	0.87	0.57 -1.33	0.83	0.57 -1.23
	<i>*2/*2</i>	0.58	0.25 -1.36	0.58	0.25 -1.36	0.58	0.25 -1.36	0.74	0.37 -1.51
Interaction: Exposition	<i>*1/*2</i>	0.75	0.27 -2.08	0.75	0.27 -2.08	0.75	0.27 -2.08		
Interaction: Exposition	<i>*2/*2</i>	2.29	0.48 -11.01	2.30	0.48 -11.01	2.30	0.48 -11.01		
<i>p interaction</i>				0.21		0.21			
<b>Génotype de la mère</b>	<i>*1/*1</i>					1.00		1.00	
	<i>*1/*2</i>					1.33	0.86 -2.04	1.31	0.82 -2.09
	<i>*2/*2</i>					0.51	0.24 -1.11	0.62	0.28 -1.38
Interaction: Exposition	<i>*1/*2</i>							1.07	0.31 -3.69
Interaction: Exposition	<i>*2/*2</i>							0.00	0.00 -0.00
<i>p interaction</i>								0.20	

n : effectif ; RR : Risque Relatif ; IC95%Wald: Intervalle de Confiance à 95% (de Wald) ; p : degré de signification du test du rapport de vraisemblance

## Annexe 4 : Complément des résultats sur le gène *MTHFR* et l'apport alimentaire en folates

1. Distribution des génotypes *MTHFR* C677T parmi les groupes cas et témoins chez les enfants et les parents
2. Tableau des triades et calcul du TDT, pour le polymorphisme *MTHFR* C677T
3. Schéma décrivant les cycles de métabolisme des folates et de l'homocystéine
4. Interaction entre le polymorphisme *MTHFR* C677T et l'apport en folates de la mère sur le risque de fentes orales non syndromiques par l'approche cas-témoins (présentation sous forme d'analyses stratifiées)
5. Interaction entre le polymorphisme *MTHFR* C677T et l'apport en folates de la mère sur le risque de fentes orales non syndromiques par l'approche cas-parents (présentation sous forme d'analyses stratifiées)

### 1. Distribution des génotypes *MTHFR* C677T parmi les groupes cas et témoins chez les enfants et les parents

Génotype	CC		CT		TT		Total n	Fréquence de l'allèle T	pHWE
	n	%	n	%	n	%			
<b>FL/P</b>									
Enfant	66	44.6	60	40.5	22	14.9	148	0.35	0.18
Mère	60	40.8	70	47.6	17	11.6	147	0.35	0.61
Père	55	40.4	61	44.9	20	14.7	136	0.37	0.65
<b>FP</b>									
Enfant	25	42.4	29	49.2	5	8.5	59	0.33	0.40
Mère	22	35.5	32	51.6	8	12.9	62	0.39	0.49
Père	26	44.1	27	45.8	6	10.2	59	0.33	0.79
<b>Fentes orales non syndromiques</b>									
Enfant	91	44.0	89	43.0	27	13.0	207	0.35	0.48
Mère	82	39.2	102	48.8	25	12.0	209	0.36	0.43
Père	81	41.5	88	45.1	26	13.3	195	0.36	0.79
<b>Témoins</b>									
Enfant	54	32.1	81	48.2	33	19.6	168	0.44	0.79
Mère	23	39.0	30	50.8	6	10.2	59	0.36	0.40

n: effectif ; %: pourcentage ; p HWE : degré de signification du test de comparaison à la distribution attendue sous l'hypothèse de l'équilibre d'Hardy-Weinberg

## 2. Tableau des triades et calcul du TDT, pour le polymorphisme *MTHFR C677T*

FL/P: 126 triades complètes et 17 diades (triades avec un génotype parental manquant)

		Enfant			
Mère	Père	CC	CT	TT	Total
CC	CC	22	x	x	
CC	CT	11	13	x	
CC	TT	x	7	x	
CT	CC	14	10	x	
CT	CT	10	9	7	
CT	TT	x	3	5	
TT	CC	x	6	x	
TT	CT	x	3	3	
TT	TT	x	x	3	126
Quand le génotype de la mère est manquant					
Père	CC	2	0	x	
	CT	0	0	1	
	TT	x	0	0	3
Quand le génotype du père est manquant					
Mère	CC	3	0	x	
	CT	2	6	2	
	TT	x	1	0	14

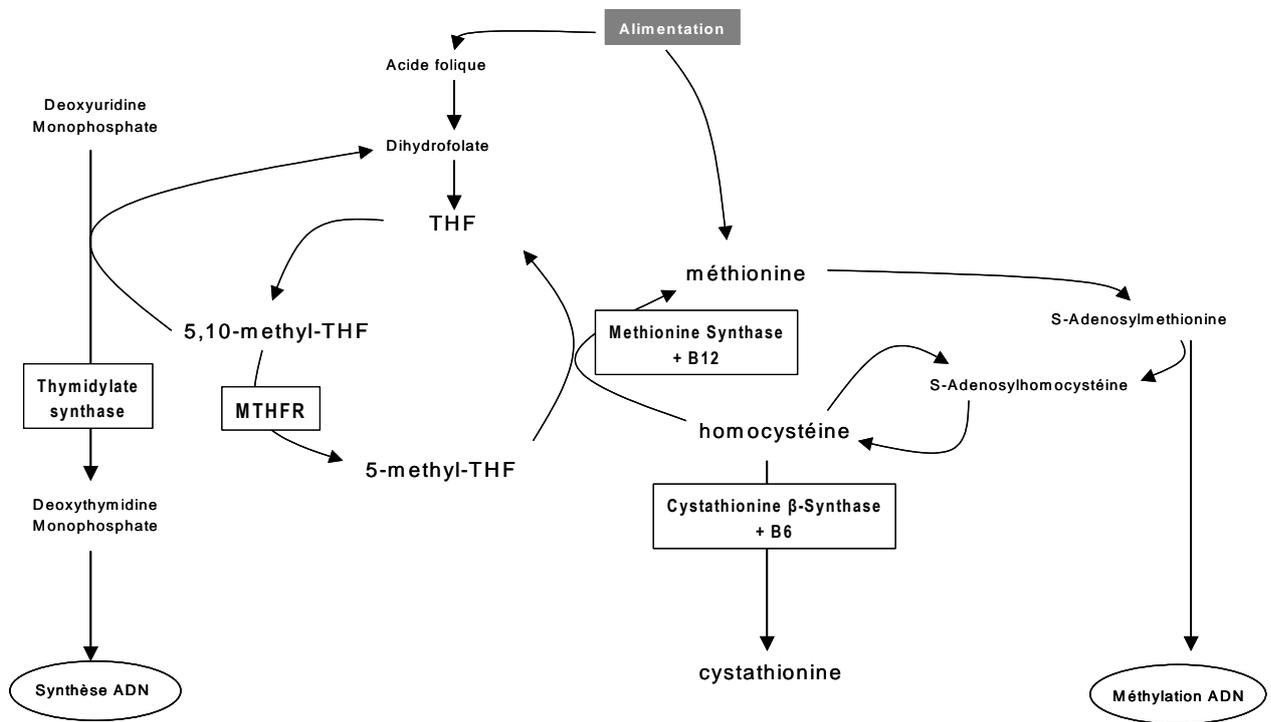
TDT	Allèle non transmis		
	*C	*T	
Allèle transmis *C	105	60	165
Allèle transmis *T	54	33	87
	159	93	252
Méthode classique	Odds= 0,90		[ 0,62 - 1,30 ]
Méthode Fleiss	std (Odds)= 0,17		
(b-c) <sup>2</sup> /(b+c)=	X <sup>2</sup> (1)= 0,32		p=0,57

FP: 50 triades complètes et 6 diades (triades avec un génotype parental manquant)

		Enfant			
Mère	Père	CC	CT	TT	Total
CC	CC	11	x	x	
CC	CT	2	5	x	
CC	TT	x	3	x	
CT	CC	6	3	x	
CT	CT	4	5	3	
CT	TT	x	2	1	
TT	CC	x	2	x	
TT	CT	x	2	1	
TT	TT	x	x	0	50
Quand le génotype de la mère est manquant					
Père	CC	1	1	x	
	CT	0	1	0	
	TT	x	0	0	3
Quand le génotype du père est manquant					
Mère	CC	0	1	x	
	CT	0	2	0	
	TT	x	0	0	3

TDT	Allèle non transmis		
	*C	*T	
Allèle transmis *C	43	25	68
Allèle transmis *T	21	11	32
	64	36	100
Méthode classique	Odds= 0,84		[ 0,47 - 1,50 ]
Méthode Fleiss	std (Odds)= 0,25		
(b-c) <sup>2</sup> /(b+c)=	X <sup>2</sup> (1)= 0,35		p=0,56

### 3. Schéma décrivant les cycles de métabolisme des folates et de l'homocystéine



Source : Sharp L, Little J. Polymorphisms in genes involved in folate metabolism and colorectal neoplasia : a HuGE review. Human Genome Epidemiology Review. Am J Epidemiol. 2004 ; 159 :423-43

**4. Interaction entre le polymorphisme *MTHFR* C677T et l'apport en folates de la mère sur le risque de fentes orales non syndromiques par l'approche cas-témoins (présentation sous forme d'analyses stratifiées)**

		Témoins	Fentes orales non syndromiques	
			OR*	IC95% Wald
		n=168	n=205	
<b>Génotype <i>MTHFR</i> de l'enfant et Apport alimentaire en folates de la mère</b>				
<b>CC†</b>	Niveau moyen-fort†		<b>1.0</b>	
	Niveau faible (<217 µg/jour)		<b>2.5</b>	1.2,5.2
<b>CT</b>	Niveau moyen-fort		<b>0.8</b>	0.4,1.4
	Niveau faible (<217 µg/jour)		<b>1.2</b>	0.6,2.3
<b>TT</b>	Niveau moyen-fort		<b>0.5</b>	0.2,1.0
	Niveau faible (<217 µg/jour)		<b>1.2</b>	0.5,3.0
<i>p</i> interaction				0.49
<b>Génotype <i>MTHFR</i> de l'enfant et Mère utilisatrice de suppléments en multivitamines contenant de l'acide folique (le mois précédant la grossesse OU le premier trimestre de la grossesse)</b>				
		n=168	n=206	
<b>CC†</b>	Oui		<b>2.1</b>	0.8,5.3
	Non†		<b>1.0</b>	
<b>CT</b>	Oui		<b>0.8</b>	0.4,1.6
	Non		<b>0.8</b>	0.5,1.3
<b>TT</b>	Oui		<b>0.6</b>	0.2,1.5
	Non		<b>0.6</b>	0.3,1.1
<i>p</i> interaction				0.42
<b>Génotype <i>MTHFR</i> de l'enfant et Apport alimentaire en folates de la mère (exclusion des femmes supplémentées en multivitamines qui ont un faible apport alimentaire en folates)</b>				
		n=155	n=179	
<b>CC†</b>	Niveau moyen-fort†		<b>1.0</b>	
	Niveau faible (<217 µg/jour)		<b>2.5</b>	1.1,5.5
<b>CT</b>	Niveau moyen-fort		<b>0.8</b>	0.4,1.4
	Niveau faible (<217 µg/jour)		<b>1.1</b>	0.5,2.3
<b>TT</b>	Niveau moyen-fort		<b>0.5</b>	0.2,1.0
	Niveau faible (<217 µg/jour)		<b>1.0</b>	0.3,3.2
<i>p</i> interaction				0.55

n: effectif ; %: pourcentage ; OR: Odds-Ratio ; IC: Intervalle de Confiance ; *p* : degré de signification du test du rapport de vraisemblance

\* Ajustement sur le centre (Lyon-Grenoble, Paris, Clermont-Ferrand), l'âge maternel en classes ( $\leq 30$ , 30-35,  $>35$  ans), l'origine ethnique de la mère (Afrique subsaharienne, Afrique du Nord, Europe, Autres)

† Catégories de référence

### 5. Interaction entre le polymorphisme *MTHFR C677T* et l'apport en folates de la mère sur le risque de fentes orales non syndromiques par l'approche cas-parents (présentation sous forme d'analyses stratifiées)

#### (1) Apport alimentaire de la mère en folates

Génotype de l'enfant	CT	TT	
<b>RR [IC95%Wald]</b>			
Niveau moyen-fort	0.92 [0.54, 1.54]	0.80 [0.33, 1.92]	
Niveau faible (<217 µg/jour)	0.60 [0.31, 1.19]	0.96 [0.35, 2.60]	
<i>p</i> interaction			0.44

RR :Risque Relatif ; IC : Intervalle de Confiance ; *p* : degré de signification du test du rapport de vraisemblance

Génotype de la mère	CT	TT	
<b>RR [IC95%Wald]</b>			
Niveau moyen-fort	0.92 [0.50, 1.70]	0.44 [0.15, 1.31]	
Niveau faible (<217 µg/jour)	1.41 [0.69, 2.88]	1.50 [0.56, 4.03]	
<i>p</i> interaction			0.24

RR :Risque Relatif ; IC : Intervalle de Confiance ; *p* : degré de signification du test du rapport de vraisemblance

#### (2) Supplémentation en multivitamines contenant de l'acide folique le mois précédant la conception OU le premier trimestre de grossesse

Génotype de l'enfant	CT	TT	
<b>RR [IC95%Wald]</b>			
Supplémentation en multivitamines	0.57 [0.24, 1.37]	0.74 [0.22, 2.51]	
Pas de supplémentation	0.83 [0.52, 1.33]	0.85 [0.39, 1.84]	
<i>p</i> interaction			0.72

RR :Risque Relatif ; IC : Intervalle de Confiance ; *p* : degré de signification du test du rapport de vraisemblance

Génotype de la mère	CT	TT	
<b>RR [IC95%Wald]</b>			
Supplémentation en multivitamines	1.45 [0.57, 3.67]	1.66 [0.51, 5.45]	
Pas de supplémentation	0.95 [0.55, 1.62]	0.56 [0.32, 1.37]	
<i>p</i> interaction			0.34

RR :Risque Relatif ; IC : Intervalle de Confiance ; *p* : degré de signification du test du rapport de vraisemblance

(3) Apport alimentaire de la mère en folates (exclusion des femmes supplémentées en multivitamines qui ont un faible apport alimentaire en folates)

<b>Génotype de l'enfant</b>	<b>CT</b>	<b>TT</b>	
<b>RR [IC95%Wald]</b>			
Niveau moyen-fort	0.92 [0.54, 1.54]	0.80 [0.33, 1.92]	
Niveau faible (<217 µg/jour)	0.59 [0.27, 1.30]	0.71 [0.19, 2.69]	
<i>p</i> interaction			0.63

RR :Risque Relatif ; IC : Intervalle de Confiance ; *p* : degré de signification du test du rapport de vraisemblance

<b>Génotype de la mère</b>	<b>CT</b>	<b>TT</b>	
<b>RR [IC95%Wald]</b>			
Niveau moyen-fort	0.92 [0.50, 1.70]	0.44 [0.15, 1.31]	
Niveau faible (<217 µg/jour)	1.50 [0.65, 3.47]	1.50 [0.42, 5.40]	
<i>p</i> interaction			0.31

RR :Risque Relatif ; IC : Intervalle de Confiance ; *p* : degré de signification du test du rapport de vraisemblance

## Annexe 5 : Complément des résultats sur les gènes *CYP2E1*, *GSTM1* et *GSTT1* et le risque de fentes orales

1. Distribution des génotypes *Cyp2E1\* RsaI* parmi les groupes cas et témoins chez les enfants et les parents
2. Tableau des triades et calcul du TDT, pour le polymorphisme du *Cyp2E1\* RsaI*
3. Estimation des effets des génotypes *CYP2E1\*RsaI* de l'enfant et de la mère sur le risque de fentes orales non syndromiques par l'approche log-linéaire
4. Estimation des effets des génotypes *CYP2E1\*RsaI* de l'enfant et de la mère sur le risque de fentes orales non syndromiques par l'approche log-linéaire, parmi les familles dont les parents sont d'origine européenne
5. Interaction entre le polymorphisme de l'enfant du *CYP2E\*RsaI* et la consommation maternelle d'alcool sur le risque de fentes orales
6. Distribution des génotypes *GSTM1* parmi les groupes cas et témoins chez les enfants et les parents
7. Distribution des génotypes *GSTT1* parmi les groupes cas et témoins chez les enfants et les parents

### 1. Distribution des génotypes *Cyp2E1\* RsaI* parmi les groupes cas et témoins chez les enfants et les parents

Génotype	*1*1		*1*5		*5*5		Total n	Fréquence de l'allèle variant <i>CYP2E1*5</i>	pHWE
	n	%	n	%	n	%			
<b>FL/P</b>									
Enfant	136	93.8	9	6.2	0	0	145	0.031	0.70
Mère	133	90.5	14	9.5	0	0	147	0.048	0.54
Père	124	91.2	11	8.1	1	0.7	136	0.048	0.19
<b>FP</b>									
Enfant	56	94.9	3	5.1	0	0	59	0.025	0.84
Mère	55	90.2	6	9.8	0	0	61	0.049	0.69
Père	55	93.2	4	6.8	0	0	59	0.034	0.79
<b>Fentes orales non syndromiques</b>									
Enfant	192	94.1	12	5.9	0	0	204	0.029	0.67
Mère	188	90.4	20	9.6	0	0	208	0.048	0.47
Père	179	91.8	15	7.7	1	0.5	195	0.044	0.28
<b>Témoins</b>									
Enfant	34	89.5	4	10.5	0	0	38	0.053	0.73
Mère	44	88.0	6	12.0	0	0	50	0.060	0.65

n: effectif ; %: pourcentage ; p HWE : degré de signification du test de comparaison à la distribution attendue sous l'hypothèse de l'équilibre d'Hardy-Weinberg

## 2. Tableau des triades et calcul du TDT, pour le polymorphisme *Cyp2E1\* RsaI*

Fentes orales non syndromiques: 175 triades complètes et 21 diades (triades avec un génotype parental manquant)

Mère	Père	Enfant			Total
		*1*1	*1*5	*5*5	
*1*1	*1*1	151	x	x	
*1*1	*1*5	3	5	x	
*1*1	*5*5	x	1	x	
*1*5	*1*1	10	2	x	
*1*5	*1*5	2	1	0	
*1*5	*5*5	x	0	0	
*5*5	*1*5	x	0	x	
*5*5	*1*5	x	0	0	
*5*5	*5*5	x	x	0	175
Quand le génotype de la mère est manquant					
Père	*1*1	5	0	x	
	*1*5	0	1	0	
	*5*5	x	0	0	6
Quand le génotype du père est manquant					
Mère	*1*1	12	1	x	
	*1*5	2	0	0	
	*5*5	x	0	0	15

TDT	Allèle non transmis			
	*1	*5		
	<b>Allèle transmis *1</b>	<b>323</b>	<b>18</b>	341
	<b>Allèle transmis *5</b>	<b>8</b>	<b>1</b>	9
		331	19	350
Méthode classique	Odds= 0,44 [ 0,19 - 1,01 ]			
Méthode Fleiss	std (Odds)= 0,19			
(b-c) <sup>2</sup> /(b+c)=	X <sup>2</sup> (1)= 3.85 p=0,049			

### FL/P: 125 triades

TDT	Allèle non transmis			
	*1	*5		
	<b>Allèle transmis *1</b>	<b>232</b>	<b>12</b>	244
	<b>Allèle transmis *5</b>	<b>5</b>	<b>1</b>	6
		237	13	250
Méthode classique	Odds= 0,42 [ 0,16 - 1,13 ]			
Méthode Fleiss	std (Odds)= 0,22			
(b-c) <sup>2</sup> /(b+c)=	X <sup>2</sup> (1)= 2.88 p=0,09			

### FP: 50 triades

TDT	Allèle non transmis			
	*1	*5		
	<b>Allèle transmis *1</b>	<b>91</b>	<b>6</b>	97
	<b>Allèle transmis *5</b>	<b>3</b>	<b>0</b>	3
		94	6	100
Méthode classique	Odds= 0,50 [ 0,13 - 2,00 ]			
Méthode Fleiss	std (Odds)= 0,35			
(b-c) <sup>2</sup> /(b+c)=	X <sup>2</sup> (1)= 1 p=0,32			

### 3. Estimation des effets des génotypes *CYP2E1\*Rsai* de l'enfant et de la mère sur le risque de fentes orales non syndromiques par l'approche log-linéaire

RR : Risque Relatif

IC : Intervalle de Confiance

		Modèle sur les 175 triades complètes		Modèle avec les 21 diades supplémentaires	
Nombre de copies de l'allèle variant <i>CYP2E1*5</i> (mode dominant) porté par...		RR	95%IC Wald	RR	95%IC Wald
... l'enfant :	1,2	<b>0.46</b>	<b>[ 0.19 – 1.09 ]</b>	<b>0.26</b>	<b>[ 0.13 – 0.53 ]</b>
	0	<b>1.00</b>		<b>1.00</b>	
... la mère :	1,2	<b>1.33</b>	<b>[ 0.56 – 3.16 ]</b>	<b>0.82</b>	<b>[ 0.45 – 1.51 ]</b>
	0	<b>1.00</b>		<b>1.00</b>	

#### Tests du rapport de vraisemblance

Effets testés (mode dominant)	Modèle sur les 175 triades complètes	Modèle avec les 21 diades supplémentaires
Génotype de l'enfant :	<b><math>p=0.14</math></b>	<b><math>p&lt;0.001</math></b>
Génotype de la mère :	<b><math>p=0.81</math></b>	<b><math>p=0.83</math></b>

**4. Estimation des effets des génotypes *CYP2E1\*Rsai* de l'enfant et de la mère sur le risque de fentes orales non syndromiques par l'approche log-linéaire, parmi les familles dont les parents sont d'origine européenne**

RR : Risque Relatif

IC : Intervalle de Confiance

<b>Familles d'origine européenne</b>		Modèle réalisé parmi 144 triades complètes et 14 diades	
Nombre de copies de l'allèle variant <i>CYP2E1*5</i> (mode dominant) porté par...		RR	95%IC Wald
... l'enfant :	1,2	<b>0.33</b>	<b>[ 0.15 – 0.71 ]</b>
	0	<b>1.00</b>	
... la mère :	1,2	<b>0.76</b>	<b>[ 0.36 – 1.57 ]</b>
	0	<b>1.00</b>	

**Tests du rapport de vraisemblance**

<b>Effets testés</b> (mode dominant)	Modèle sur les 175 triades complètes
Génotype de l'enfant :	<b><math>p=0.001</math></b>
Génotype de la mère :	<b><math>p=0.47</math></b>

### 5. Interaction entre le polymorphisme de l'enfant du CYP2E\*RsaI et la consommation maternelle d'alcool sur le risque de fentes orales

Polymorphisme de l'enfant du CYP2E1 *RsaI	n triades* (impliqué dans le calcul de l'interaction)	Interaction entre l'exposition et la présence du variant chez l'enfant †		p interaction
		RR	IC 95%	
<b>Consommation maternelle d'alcool</b>				
Sur la période allant du mois précédant la conception et le 1 <sup>er</sup> trimestre de grossesse	4	6.6	0.8, 51.6	0.07
Pendant le 1 <sup>er</sup> trimestre de grossesse	2	5.6	0.4, 75.2	0.19

\*n triades: nombre de familles exposées et dont l'enfant est porteur du variant *RsaI* ;

RR: Risque Relatif ; IC: Intervalle de Confiance

† Interaction sous l'échelle multiplicative entre le polymorphisme de l'enfant et l'exposition d'intérêt (=1 si l'hypothèse multiplicative est vérifiée (absence d'interaction), ≠ 1 si déviation de l'hypothèse multiplicative (présence d'interaction))

p interaction : test du rapport de vraisemblance de l'interaction évalué par l'approche des pseudo-sibs

## 6. Distribution des génotypes *GSTM1* parmi les groupes cas et témoins chez les enfants et les parents

<i>GSTM1</i>	Null		Wild		Total n
	n	%	n	%	
<b>FL/P</b>					
Enfant	78	52.7	70	47.3	148
Mère	79	53.4	69	46.6	148
Père	71	52.2	65	47.8	136
<b>FP</b>					
Enfant	36	59.0	25	41.0	61
Mère	39	61.9	24	38.1	63
Père	34	56.7	26	43.3	60
<b>Fentes orales non syndromiques</b>					
Enfant	114	54.5	95	45.5	209
Mère	118	55.9	93	44.1	211
Père	105	53.6	91	46.4	196
<b>Témoins</b>					
Enfant	93	56.4	72	43.6	165
Mère	26	47.3	29	52.7	55
<hr/>					
<b>FL/P</b>					
Enfant : OR*; IC 95%	0.87	0.55-1.36	1.00		
Mère : OR†; IC 95%	1.35	0.72-2.54	1.00		
<b>FP</b>					
Enfant : OR*; IC 95%	1.10	0.60-2.02	1.00		
Mère : OR†; IC 95%	1.73	0.83-3.64	1.00		
<b>Fentes orales non syndromiques</b>					
Enfant : OR*; IC 95%	0.93	0.61-1.41	1.00		
Mère : OR†; IC 95%	1.44	0.79-2.63	1.00		
<b>Fentes orales non syndromiques : analyses restreintes aux familles d'origine européenne</b>					
Enfant : OR*; IC 95%	0.93	0.57-1.44	1.00		
Mère : OR†; IC 95%	1.49	0.78-2.83	1.00		

OR : Odds-ratio ; IC : Intervalle de Confiance

\* ajusté sur l'âge de la mère (3 classes), le centre (3 catégories) (et l'origine ethnique de la mère (4 catégories))

† ajusté sur l'âge de la mère (3 classes), le centre (2 catégories) (et l'origine ethnique de la mère (2 catégories))

7. Distribution des génotypes *GSTT1* parmi les groupes cas et témoins chez les enfants et les parents

<i>GST T1</i>	Null		Wild		Total
	n	%	n	%	n
<b>FL/P</b>					
Enfant	30	20.3	118	79.7	148
Mère	25	16.9	123	83.1	148
Père	28	20.6	108	79.4	136
<b>FP</b>					
Enfant	14	23.3	46	76.7	60
Mère	12	19.0	51	81.0	63
Père	11	18.3	49	81.7	60
<b>Fentes orales non syndromiques</b>					
Enfant	44	21.2	164	78.8	208
Mère	37	17.5	174	82.5	211
Père	39	19.9	157	80.1	196
<b>Témoins</b>					
Enfant	30	18.2	135	81.8	165
Mère	6	10.9	49	89.1	55
<hr/>					
<b>FL/P</b>					
Enfant : OR*; IC 95%	1.12	0.63-1.98	1.00		
Mère : OR†; IC 95%	1.56	0.59-4.11	1.00		
<b>FP</b>					
Enfant : OR*; IC 95%	1.24	0.59-2.61	1.00		
Mère : OR†; IC 95%	1.96	0.65-5.91	1.00		
<b>Fentes orales non syndromiques</b>					
Enfant : OR*; IC 95%	1.17	0.69-1.98	1.00		
Mère : OR†; IC 95%	1.61	0.63-4.10	1.00		
<b>Fentes orales non syndromiques : analyses restreintes aux familles d'origine européenne</b>					
Enfant : OR*; IC 95%	1.50	0.83-2.70	1.00		
Mère : OR†; IC 95%	1.26	0.48-3.29	1.00		

OR : Odds-ratio ; IC : Intervalle de Confiance

\* ajusté sur l'âge de la mère (3 classes), le centre (3 catégories) (et l'origine ethnique de la mère (4 catégories))

† ajusté sur l'âge de la mère (3 classes), le centre (2 catégories) (et l'origine ethnique de la mère (2 catégories))

## Annexe 6 : Effectifs du tableau 42

Exposition (% colonne)	Solvants oxygénés				Solvants chlorés n=20	Solvants pétroliers n=30
	Alcools alipha- tiques n=136	Ethers de glycol n=96	Aldeh., est.,cét. aliph.* n=130	Au moins un solvant oxygéné n=157		
Médecin, dentiste, vétérinaire, pharmacien	14	0	14	14	0	0
Infirmière, aide-soignant	26	4	26	26	0	0
Personnel enseignant et animateur	4	5	7	8	0	1
Ingénieur, technicien	6	5	9	11	2	5
Employé, secrétaire et métiers de bureau	13	9	5	15	3	5
Commerçant et vendeur	15	16	10	16	1	4
Femmes de ménage	27	26	27	27	0	1
Nourrice	5	5	0	5	0	0
Autres employés de maison	3	3	3	3	0	0
Coiffeur	5	5	5	5	1	0
Serveur, cuisinier	7	7	7	7	0	0
Esthéticien	3	3	3	3	3	4
Ouvrier du cuir	0	0	1	1	1	1
Autres ouvriers industriels	8	8	10	13	9	9
Ouvriers agricoles	0	0	3	3	0	0

\* Aldéhyde, esters et cétones aliphatiques

# Annexe 7 : Puissance statistique *a posteriori* des tests d'association pour chaque facteur environnemental et chaque gène, par le logiciel QUANTO

Référence du logiciel : Gauderman WJ. Sample size requirements for matched case-control studies of gene-environment interaction. Stat Med. 2002;21:35-50

## Puissance statistique pour les tests d'association avec chaque facteur environnemental étudié

Paramètres requis par QUANTO	Consommation maternelle d'alcool		Exposition professionnelle aux solvants	
	En période péri-conceptionnelle	Au premier trimestre de grossesse	Tout solvant	Solvants chlorés
Effectif <sup>†</sup>	236 témoins	236 témoins	187 témoins	187 témoins
Prévalence de l'exposition dans la population*	0.29	0.13	0.40 <sup>‡</sup>	0.02 <sup>‡</sup>
Risque de base de la maladie dans la population	0.0014	0.0014	0.0014	0.0014
<b>Puissance statistique estimée pour détecter un risque relatif de maladie associé à l'exposition de :</b>				
1.25	20%	14%	19%	
1.50	54%	36%	49%	
1.75	82%	61%	77%	
2.00	95%	81%	91%	20%
5.00				88%
8.00				99%
Estimation du risque dans l'application	≈1.5	≈1.7	≈1.7	≈8

Note : Compte-tenu des contraintes de QUANTO, les valeurs choisies des paramètres sont approximées par rapport aux valeurs observées dans chaque application. Le risque d'erreur de première espèce est fixé à 0.05.

\* Calculée à partir de la population de témoins

<sup>†</sup> Pour un schéma d'étude cas-témoins non apparié, le ratio cas : témoin est 1 :1

<sup>‡</sup> Sur la population des femmes ayant une activité professionnelle au 1er trimestre de grossesse

## Puissance statistique pour les tests d'association avec chaque gène étudié

Paramètres requis par QUANTO	<i>ADH1C</i>	<i>MTHFR</i>		<i>CYP2E1</i>	<i>GSTM1</i>	<i>GSTT1</i>
Schéma d'étude*	Cas-Parents	Cas-témoins / Cas-Parents (non appariés)		Cas-Parents	Cas-témoins (non appariés)	Cas-témoins (non appariés)
Effectif	174 triades	146 témoins / 124 triades		175 triades	165 témoins	165 témoins
Mode de transmission de l'allèle†	Dominant	Récessif	Dominant	Dominant	Récessif	Récessif
Fréquence de l'allèle d'intérêt dans la population‡	0.40	0.44	0.44	0.05	0.75§	0.42§
Risque de base de la maladie dans la population	0.0014	0.0008	0.0008	0.0014	0.0014	0.0014
<b>Puissance statistique estimée pour détecter un risque relatif de maladie associé à l'allèle d'intérêt de :</b>						
0.20	99%	98% / 97%	99% / 99%	88%	99%	98%
0.50	92%	55% / 54%	82% / 81%	39%	88%	56%
1.50	45%	31% / 31%	34% / 32%	23%	43%	32%
2.00	86%	73% / 74%	72% / 69%	59%	85%	75%
Estimation du risque dans l'application	≈0.6	≈0.9 / ≈0.9	≈0.6 / ≈0.75	≈0.3	≈0.9	≈1.2

Note : Compte-tenu des contraintes de QUANTO, les valeurs choisies des paramètres sont approximées par rapport aux valeurs observées dans chaque application. Nous considérons ici le gène de l'enfant. Le risque d'erreur de première espèce est fixé à 0.05.

\* Pour un schéma d'étude cas-témoins non apparié, le ratio cas :témoins est 1 :1

† Dominant, récessif ou hypothèse multiplicative

‡ Calculée à partir de la population de témoins

§ Sachant la fréquence du génotype homozygote *Null* dans la population témoins, la fréquence de la délétion est calculée à partir de l'équilibre  $p^2+2pq+q^2$

# Sur la population des femmes ayant une activité professionnelle au 1er trimestre de grossesse

## Résumé

Les fentes orales sont des malformations congénitales d'origine complexe, dépendant à la fois d'expositions environnementales et d'influences génétiques. L'objectif général de ce travail est la recherche d'interactions gène-environnement dans leur étiologie.

Ce projet s'appuie sur une étude épidémiologique cas-témoins menée en France (1998-2001, 240 cas-236 témoins) intégrant un schéma d'étude cas-parents. Il présente les méthodes statistiques à ce jour disponibles pour évaluer une interaction gène-environnement selon les deux schémas d'étude. Un modèle spécifique est utilisé pour différencier les effets du génotype de l'enfant de celui de la mère.

Les applications de ce projet s'intéressent au gène *ADH1C* et à la consommation maternelle d'alcool, au gène *MTHFR* et à l'apport en folates de la mère, et à l'exposition professionnelle de la mère aux solvants organiques en interaction avec des gènes de métabolisme/détoxication tels que *CYP2E1*, *ADH1C*, *GSTM1* et *GSTT1*.

**Mots-clés :** fente orale, environnement, gène, interaction, cas-témoins, cas-parents, épidémiologie

## Abstract

Oral clefts are congenital anomalies with a complex origin, which appears to be determined both by environmental exposures and genetic influences. The purpose of this work is to explore gene-environment interactions in their etiology.

This epidemiological project is based on a case-control study, conducted in France between 1998 and 2001 (240 cases, 236 controls), which incorporates a case-parents design. Recent statistical models available for the evaluation of gene-environment interactions are described. A specific model is used to distinguish the effects of the child's and maternal genotypes. .

Applications of this project focus on *ADH1C* gene and maternal alcohol consumption, *MTHFR* gene and maternal folate intake, and maternal occupational exposure to organic solvents and genes involved in the metabolism/detoxification processes such as *CYP2E1*, *ADH1C*, *GSTM1* and *GSTT1*.

**Key-words :** oral cleft, environment, gene, interaction, case-control, case-parent, epidemiology