

**Etudes d'association pangénomiques appliquées à
la recherche de nouveaux facteurs de risque
génétique de la maladie d'Alzheimer**

Études d'association génétique

Les études d'association comparent la distribution des allèles ou des génotypes de marqueurs génétiques chez des individus malades et témoins d'une population de sujets non apparentés. On s'attend en effet à ce que la fréquence d'un variant impliqué dans une maladie soit significativement différente dans l'échantillon de malades comparativement aux témoins.

La détection d'une association entre un marqueur et la maladie suggère que le marqueur est soit directement impliqué dans la maladie (variant causal), soit en déséquilibre de liaison avec un variant causal.

Les études d'association sont réalisées dans la plupart des cas à partir de méthodes utilisées en épidémiologie classique : test du χ^2 , test de rapport de vraisemblance ou régression logistique par exemple.

Les premières études d'association ont donc été réalisées en complément des études de liaison, avec des variants situés dans un pic de liaison génétique et à proximité de gènes dont la fonction supposée suggérait leur implication dans le développement de la maladie. On parle alors de stratégie « gène candidat ».

À partir de 2005, les connaissances apportées dans les modèles génétiques et la structure du génome humain et les progrès réalisés dans les techniques de génotypage ont permis la réalisation d'études d'association pangénomiques à des coûts raisonnables.

Ces dernières, mieux adaptées à la recherche de variants fréquents supposés impliqués dans les maladies multifactorielles, ont remplacé les analyses de liaison pour la recherche de loci d'intérêt dans les maladies complexes.

Études d'association pangénomiques

A- Contexte

Deux articles fondateurs publiés en 1996 sont à l'origine de l'intérêt et de l'utilisation des études d'association pangénomiques (Lander ES. Science,1996; Risch N. Merikangas K,1996).

1- Modèle maladie fréquente, variant fréquent

Les maladies fréquentes sont liées à des variants fréquents et que ces derniers seront plus facilement détectés en utilisant des études d'association en population plutôt que des études de liaison.

Les études d'association permettent d'avoir une puissance plus grande que les études de liaison pour des variants fréquents. De plus, pour des variants associés à une modulation du risque faible à modérée, seule une étude d'association est possible. En effet, les effectifs nécessaires pour obtenir une puissance statistique équivalente dans une étude de liaison seraient de l'ordre du million de sujets.

Ces calculs de puissance sont basés sur un nombre important de marqueurs. L'utilisation de cartes denses de SNPs basées sur le déséquilibre de liaison est proposée pour limiter le nombre de SNPs nécessaires tout en assurant une couverture complète du génome. Ces pré-requis sont techniquement hors de portée en 1996.

Néanmoins, l'hypothèse « maladie fréquente, variant fréquent » (“Common Disease Common Variant”) prend forme et entraîne des réorientations stratégiques en vue de rendre possible la réalisation d'études d'association pangénomiques.

2. Identification des variants fréquents : le projet Hapmap

En 2002, le projet HapMap a pour objectif d'identifier les polymorphismes fréquents présents sur le génome ainsi que les structures de déséquilibre de liaison entre ces polymorphismes afin d'aider à la planification et à l'analyse des études génétiques sur les maladies humaines.

Un des buts recherchés est de permettre la réalisation d'études pangénomiques grâce à la sélection de tagSNPs sur l'ensemble du génome.

L'objectif initial est de génotyper 1 million de SNPs de MAF > 5% chez 269 individus appartenant à 4 populations différentes, européenne, africaine, chinoise et japonaise. En complément, une étude plus exhaustive sur certaines régions du génome (régions ENCODE) par séquençage est réalisée. Les SNPs identifiés dans ces régions sont ensuite génotypés dans l'ensemble de l'échantillon HapMap. L'objectif est atteint en 2005.

Par ailleurs, des millions de nouveaux SNPs sont identifiés et la faisabilité des études d'association pangénomiques est démontrée. Le projet HapMap se poursuit avec la publication régulière de nouvelles cartes des blocs haplotypiques. Plus récemment, ces efforts ont été rejoints par le projet 1000 Genomes.

3. Puces de génotypage

Le génotypage utilisé dans les études pangénomiques se base sur les puces à ADN et des technologies comme Infinium et BeadArray. Le coût du génotypage d'un SNP est passé de 1 dollar par échantillon en 2001 à 0,05 dollar en 2008. Actuellement, le coût du génotypage de 2,5 millions de SNPs coûte environ 500 dollars par échantillon, soit 0,0002 dollar par SNP.

B. Conception

Une étude d'association pangénomique est une étude cas-témoin réalisée dans une population de sujets non apparentés. Le but est d'identifier les loci dont la fréquence allélique est significativement différente entre cas et témoins. En ce sens, elle diffère peu de l'approche gène candidat. Ses spécificités sont liées à son caractère de recherche systématique, sans à priori, de loci associés aux maladies et au grand nombre de marqueurs génétiques testés qui en découle.

Tableau 3.1 – Choix et contraintes dans la conception d'une étude pangénomique

Choix		Contrainte
Puissance maximale	↔	Coût
Taille d'échantillon	↔	Limite supérieure (temps et coût)
Quels SNPs génotyper	↔	Disponibilité des puces commerciales
Analyses statistiques	↔	

1- Collecte des échantillons

La première étape d'une étude pangénomique consiste à collecter des échantillons adéquats, le plus souvent des cas et des témoins. Si le trait est quantitatif, un tirage au sort en population ou la sélection des valeurs extrêmes d'une distribution peuvent être envisagés.

Assurer une puissance statistique suffisante est un enjeu capital des études d'association pangénomiques. Celle-ci dépend de la taille d'effet et de la fréquence allélique du variant fonctionnel, paramètres inconnus avant la réalisation de l'étude. L'effet et la fréquence attendues du variant causal peuvent être estimées par les études antérieures.

Par exemple, pour un variant causal ayant un odds ratio de 1,25 et une MAF de 40%, une erreur de type I $\alpha = 5 \times 10^{-7}$ et un marqueur ayant un LD = 0,8, 3125 cas et 3125 témoins sont nécessaires pour obtenir une puissance de 80%.

L'enrichissement génétique consiste en la sélection de cas qui présentent en plus de la maladie considérée, d'autres caractéristiques à forte composante génétique, comme un âge d'apparition plus bas, des antécédents familiaux, ou une forme plus grave de la maladie. L'avantage de cette méthode est un gain de puissance statistique. Néanmoins, les inconvénients sont nombreux. L'information est plus difficile à retrouver et souvent de moins bonne qualité.

Le gain de puissance attendu est donc contrebalancé par la taille plus faible des échantillons. De plus, rien ne garantit que la caractéristique considérée apporte un enrichissement génétique. Enfin, la généralisation des résultats peut être compromise.

2. Génompage

Une fois que les échantillons ont été collectés, l'étape suivante est de les génotyper. Après traitement du signal, les résultats des puces indiquent pour chaque SNP, le génotype des individus en terme d'intensité des deux allèles du SNP.

Si on représente graphiquement les deux intensités pour tous les individus, on voit apparaître, dans le meilleur des cas, trois groupes correspondant aux génotypes (deux homozygotes, un hétérozygote).

L'utilisation de « calling algorithms » permet d'attribuer un génotype aux groupes définis par l'intensité.

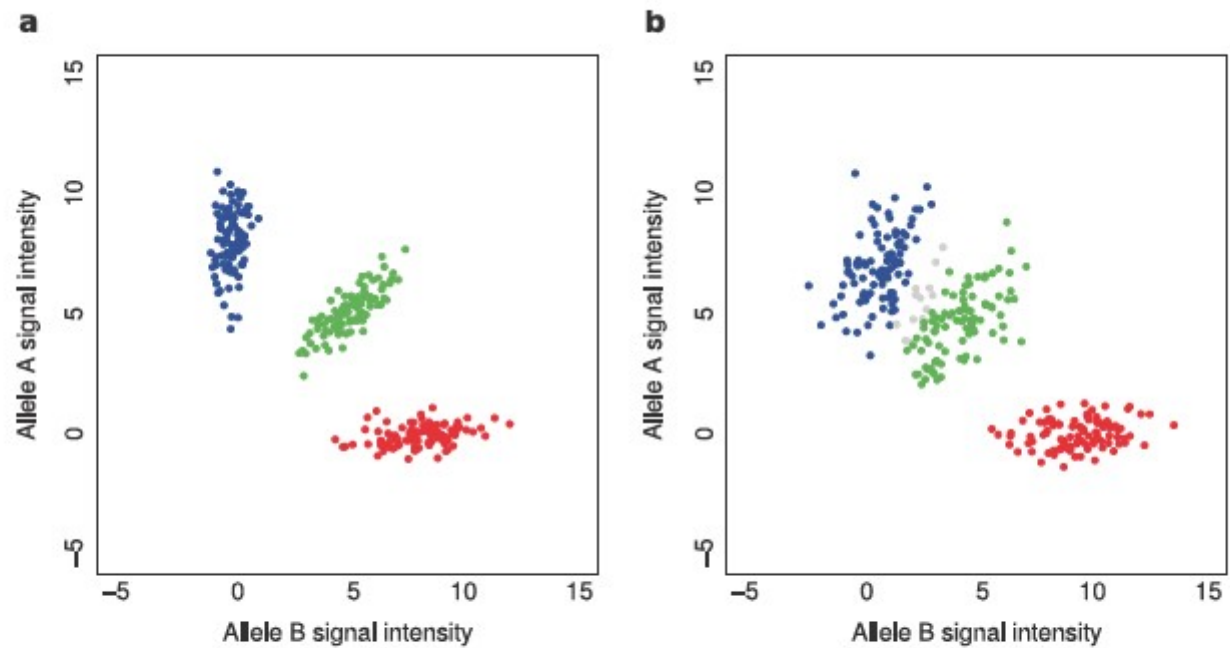


FIGURE 3.3 – Exemple de graphe d'intensité. (a) Les trois génotypes apparaissent distinctement et l'algorithme attribue correctement les génotypes. (b) En raison de leur proximité, une partie des groupes AA et AB est classée en données manquantes par l'algorithme (en gris). Tiré de [125]

3. Contrôle qualité

De nombreux problèmes, d'ordre technique ou statistique, peuvent être à l'origine de résultats faussement positifs et perturber la détection des marqueurs génétiques associés à la maladie considérée.

Afin de limiter au maximum ces faux positifs, des critères qualité doivent être systématiquement appliqués.

Les critères utilisés pour les SNPs visent à détecter d'éventuelles erreurs de génotypage. Les critères utilisés chez les individus cherchent à identifier une mauvaise qualité de l'ADN ou une structure cachée de la population (stratification, apparentement).

Contrôle qualité sur les SNPs.

- ✓ **Taux de données manquantes (Call Rate).** Les SNPs ayant un pourcentage élevé de données manquantes peuvent avoir une mauvaise qualité pour les données non Manquantes.
- ✓ **Équilibre de Hardy-Weinberg (HWE).** Une déviation de l'HWE peut être observée en cas de stratification de la population, d'apparement entre individus ou d'erreur de génotypage ce qui permet d'identifier des SNPs à problèmes. Une association positive entre un marqueur et une maladie pouvant être à l'origine d'un léger déséquilibre, des seuils de significativité très bas sont utilisés ($p < 1 \times 10^{-6}$).

✓ **Fréquence Allélique.**

Pour les SNPs ayant une MAF basse, l'attribution des génotypes par les algorithmes est plus difficile en raison des faibles effectifs potentiels dans les différentes catégories (le groupe homozygote pour l'allèle mineur peut être manquant et le groupe hétérozygote de faible taille) et le risque de données manquantes et d'erreurs de génotypage est plus élevé.

Enfin, les SNPs ayant une MAF basse ne permettent pas d'obtenir une puissance statistique suffisante à leur détection.

En conséquence, les SNPs ayant une MAF basse sont exclus des résultats.

- ✓ Tout autre donnée à disposition (chronologie des plaques de génotypage ou différence de call rate entre cas et témoins par exemple) doit être utilisée en cas d'observation inhabituelle.

4. Analyse statistique

L'approche la plus couramment utilisée consiste en l'analyse individuelle de tous les SNPs disponibles.

On considère souvent les allèles dans un modèle génétique additif, et testé par un test du χ^2 de tendance ou dans une régression logistique.

Les SNPs peuvent également être utilisés sous formes d'haplotypes. Les avantages par rapport à l'analyse d'un SNP seul sont une meilleure couverture de l'information du génome et une identification plus facile des variants rares.

Ces analyses nécessitent des techniques statistiques spécifiques car les haplotypes des sujets sont inconnus et doivent donc être inférés à partir des génotypes.

Une étude d'association pangénomique peut également chercher à identifier des gènes qui influencent des traits quantitatifs.

Ces études sont alors basées sur des échantillons aléatoires réalisés en population. Comme dans les études cas témoins, l'association dépend du LD entre un marqueur et un locus impliqué dans un trait.

Elles peuvent être indépendantes ou complémentaires d'une étude cas-témoin. L'analyse fait alors appel à une régression linéaire.

L'interprétation du test se fait classiquement via le calcul de la P-value et sa comparaison à un seuil de significativité. Les seuils utilisés sont souvent bas en raison du nombre important de tests réalisés. Une simple correction de Bonferroni est conservatrice, les tests n'étant pas tous indépendants en raison du LD.

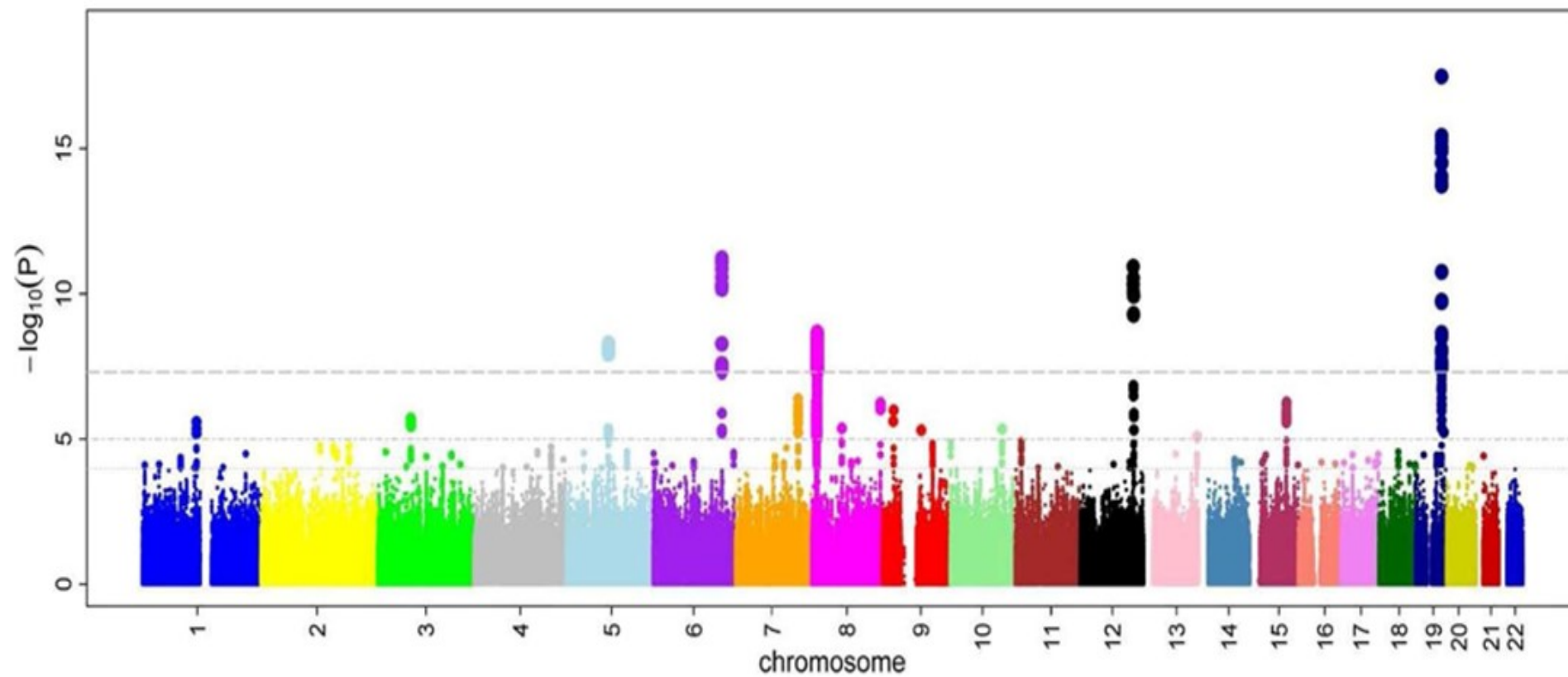


FIGURE 3.7 – Exemple de Manhattan Plot

En cas de résultats significatifs, trois interprétations possibles peuvent être faites :

- ❑ Le locus est causal et influence directement le risque de maladie. Des études fonctionnelles sont nécessaires,
- ❑ Les allèles sont corrélés aux allèles à risque via le LD mais n'influencent pas directement le risque de maladie,
- ❑ Le résultat est un faux positif,

Il est intéressant d'estimer le risque via le calcul d'un odds ratio, d'une mesure d'impact comme la fraction de risque attribuable en population ou d'un pourcentage de variance génétique expliquée. Néanmoins, ces mesures sont difficiles à interpréter et dépendent de la population étudiée.

Tableau 3.2 – Éléments de preuve dans une étude d'association pangénomique

Preuve Forte	Preuve Faible
L'association n'est pas liée à un artefact (erreur de génotypage, stratification)	SNP suspicieux
P-value faible	P-value haute
Puissance de l'étude élevée	Puissance faible
MAF élevé	MAF faible
Gène candidat	Région intergénique
SNP connu pour être associé	SNP inconnu

Malgré les précautions prises pendant la conception de l'étude et la collecte, le nettoyage et l'analyse des données, le risque de faux-positifs reste élevé.

La meilleure validation reste donc la réplication des résultats obtenus dans une population indépendante.

Cette étape permet également d'estimer plus précisément le risque associé au SNP, celui-ci ayant tendance à être initialement surestimé (« winner's curse »).

Maladie d'Alzheimer est une maladie multifactorielle

Définition

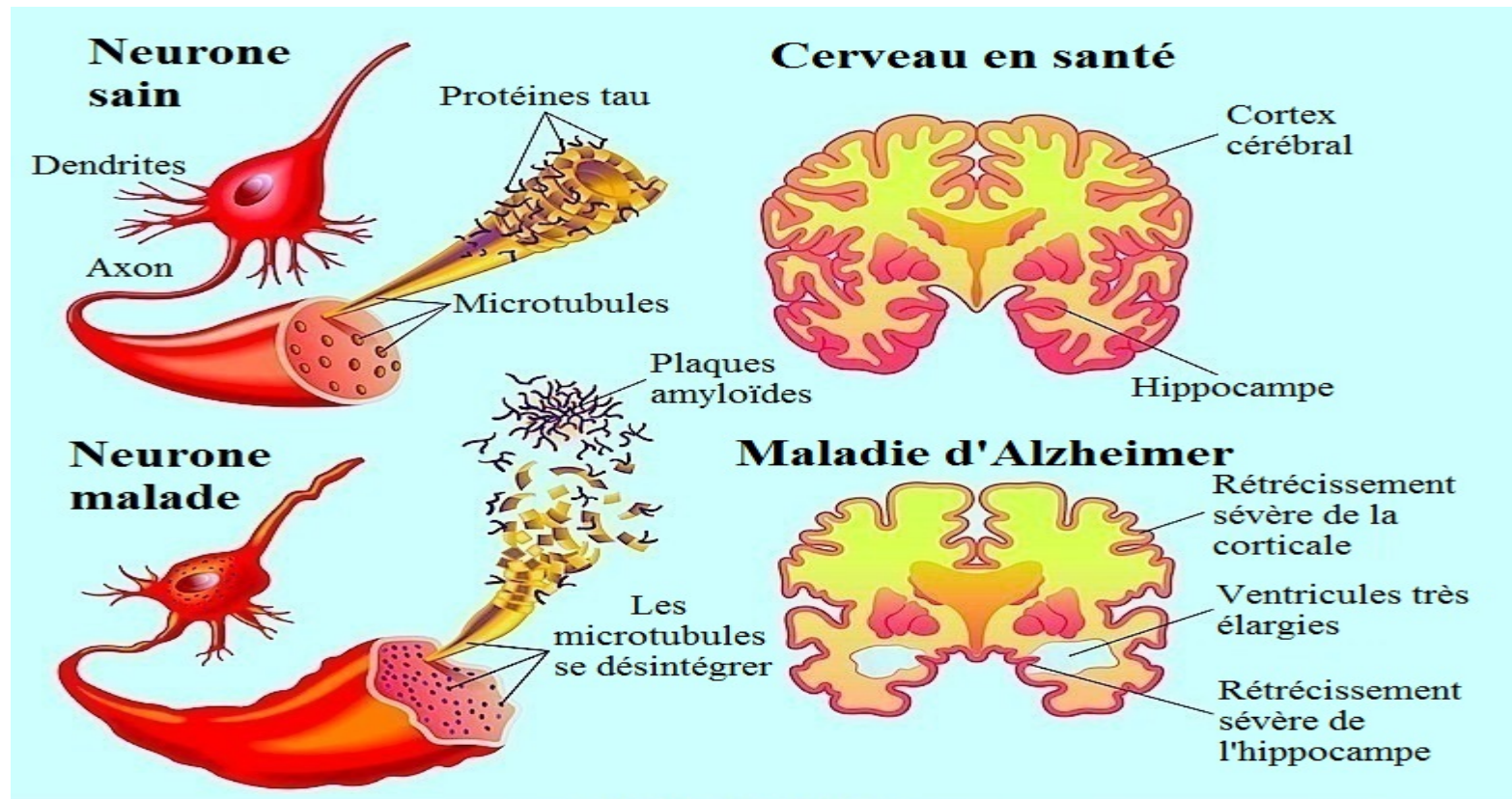
La maladie d'Alzheimer est une maladie neurodégénérative qui se caractérise par le développement d'une démence progressive (du latin demens, fou, obnubilé). Elle se manifeste principalement à un âge avancé et est associée à une diminution de la performance du cerveau et à des troubles de la mémoire. Elle a été décrite pour la première fois au début du 20e siècle par le médecin allemand, Dr. Alois Alzheimer de Tübingen, en se basant sur le tableau clinique et les découvertes anatomopathologiques sur une de ses patientes.

Des dépôts de protéines au niveau cérébral, appelés plaques, engendrent des troubles du langage, de la pensée et de la mémoire. Des neurotransmetteurs importants tels que l'acétylcholine ne sont plus synthétisés en quantité suffisante, ce qui conduit à un affaiblissement des capacités cérébrales.

Avec l'augmentation de l'espérance de vie, surtout dans les pays industrialisés, l'incidence de la maladie y connaît une forte augmentation. Le risque de développer la maladie d'Alzheimer augmente avec l'âge. Dans de rares formes génétiques, la maladie peut déjà survenir dès l'âge de 30 ans.

CAUSES

Les vraies causes de la maladie sont encore inconnues et tout porte à croire que divers facteurs pourraient être à l'origine de la maladie. Des dépôts amyloïdes au niveau cérébral entraînent la mort de cellules nerveuses. Les cellules du cerveau, principalement celles du centre de la mémoire, du centre du langage et du centre des activités cognitives sont avant tout touchées.



Facteurs de risque de la maladie

A- Facteurs de risque modifiables

Les revues systématiques et méta-analyses résultats suggèrent que la réserve cognitive, un concept combinant les bénéfices de l'éducation, du travail et des activités intellectuelles, une activité physique régulière, l'obésité chez les personnes d'âge moyen, la consommation d'alcool et le tabagisme sont les plus importants facteurs modifiables de maladie d'Alzheimer.

Les accidents vasculaires cérébraux, le diabète et l'hypercholestérolémie et l'hypertension chez les personnes d'âge moyen sont eux associés à un risque augmenté de maladie d'Alzheimer.

Selon une autre étude, 50% des cas de maladie d'Alzheimer dans le monde pourrait être prévenu par la suppression de sept facteurs de risque : diabète, hypertension et obésité chez les personnes d'âge moyen, dépression, sédentarité, tabagisme et bas niveau d'éducation. Néanmoins, l'étude des facteurs de risque associés au risque de maladie d'Alzheimer est complexe et difficile.

Synthèse des facteurs modulant le risque de maladie d'Alzheimer

Sens de l'association	Facteurs	Niveau de preuve
Augmentation du risque	Allèle $\epsilon 4$ de l'APOE	Modéré
	Association œstrogène conjugué équin / méthyl-progestérone Certains anti-inflammatoires non-stéroïdiens Dépression Diabète de type 2 Hyperlipidémie en milieu de vie Lésion cérébrale traumatique chez les hommes Exposition aux pesticides Jamais marié, bas support social Tabagisme actif	Bas
Diminution du risque	Régime méditerranéen Acide folique Inhibiteurs de la HMG-CoA reductase (statines) Haut niveau d'éducation Consommation basse à modérée d'alcool Activités intellectuelles Activités physiques	Bas
	Pas d'association	Élevé Modéré
Pas d'association	Ginkgo Biloba Vitamine E Inhibiteurs de la cholinestérase Traitement antihypertenseur Œstrogènes conjugués équins Acides gras $\omega 3$ Vitamines B12, C, β -carotène Homocystéine Hypertension Obésité Syndrome métabolique Facteurs précoces de l'enfance Catégorie socio-professionnelle Plomb	Bas
	Manque de preuves	(Non applicable)
Manque de preuves	Acide gras saturés Apport en fruits et légumes Métaux Apports caloriques élevés Mémantine Apnée du sommeil Troubles anxieux Résilience Loisirs Agent orange, syndrome de la guerre du Golfe Solvants, aluminium Facteurs génétiques autres que APOE	(Non applicable)

Génétique

A- Composante génétique de la maladie d'Alzheimer

Après l'âge, les antécédents familiaux de maladie d'Alzheimer représentent un des ses plus importants facteurs de risque. Le risque de développer la maladie d'Alzheimer est ainsi deux fois plus élevé chez les sujets apparentés au premier degré à des patients atteints de la maladie d'Alzheimer comparé au risque en population générale.

Dans une étude réalisée chez 11884 paires de jumeaux, l'héritabilité de la maladie d'Alzheimer est estimée entre 58% et 79%. Ces résultats sont en faveur d'une composante génétique importante de la maladie d'Alzheimer.

B- Formes génétiques de la maladie d'Alzheimer

Sur le plan génétique, deux formes de maladie d'Alzheimer sont classiquement identifiées, même si en réalité, les différences sont loin d'être aussi marquées.

Les formes familiales représentent moins de 1% des cas de maladie d'Alzheimer. Elles sont caractérisées par un âge de survenue précoce, avant 60 ans, et un mode de transmission mendélien, autosomique dominant le plus souvent. Ces formes sont causées par des mutations rares, à pénétrance élevée, situées dans trois gènes, APP, PSEN1 et PSEN2 codant respectivement pour la protéine précurseur du peptide amyloïde et les présénilines 1 et 2.

Les formes dites « sporadiques » constituent la grande majorité des cas de maladie d'Alzheimer. Survenant après l'âge de 60 ans, ces formes n'ont pas de mode de transmission précis ni d'agrégation familiale.

B- Formes génétiques de la maladie d'Alzheimer

En raison de leur composante génétique supposée forte, on préfère actuellement parler de formes à mode de transmission complexe ou non-mendélien plutôt que de formes sporadiques. Leur survenue implique également une interaction avec des facteurs environnementaux, bien que celle-ci reste difficile à établir.

L'identification de ces facteurs de risque génétique constitue un enjeu important à plusieurs titres : apporter de nouvelles connaissances sur la physiopathologie de la maladie, identifier de nouvelles cibles thérapeutiques et de nouveaux modes d'intervention et développer de nouveaux marqueurs diagnostiques et de dépistage.

Facteurs de risque génétique de maladie d'Alzheimer

A- Facteurs génétiques impliqués dans les formes familiales de maladie d'Alzheimer

Historiquement, l'identification des premières mutations responsables de maladie d'Alzheimer fait suite à la découverte du peptide A β et le constat que ce peptide est retrouvé à la fois chez des patients atteints de maladie d'Alzheimer et chez des sujets trisomiques 21, qui présentent presque systématiquement un syndrome démentiel à partir de 40 ans.

En 1987, le gène APP codant pour le précurseur du peptide amyloïde est identifié. Ce gène constitue un bon candidat car situé à proximité d'un pic de liaison avec des formes familiales de maladie d'Alzheimer.¹⁴⁴ Ceci est confirmé par l'identification des premières mutations causant la maladie d'Alzheimer dans le gène de l'APP.

Les mutations du gène APP n'expliquent qu'une partie des formes familiales de maladie d'Alzheimer. En 1995, d'autres mutations associées à ces formes sont découvertes dans les gènes PSEN1 et PSEN2, codant respectivement pour les présénilines 1 et 2.

Facteurs de risque génétique de maladie d'Alzheimer

A- Facteurs génétiques impliqués dans les formes familiales de maladie d'Alzheimer

Gènes impliqués dans les formes familiales précoces de maladie d'Alzheimer et leurs effets pathogènes

Gène Protéine	Localisation génomique	Mutations (clinique/ indéterminée/ bénigne)	Familles ou individus	Phénotype moléculaire
APP Précurseur du peptide amyloïde β	21q21.2	32/1/6	90	Ratio $A\beta_{1-42}/A\beta_{1-40}$ augmenté Production d' $A\beta$ augmentée Agrégation d' $A\beta$ augmentée
PSEN1 Préséniline 1	14q24.3	185/8/4	411	Ratio $A\beta_{1-42}/A\beta_{1-40}$ augmenté
PSEN2 Préséniline 2	1q42.13	13/7/5	34	Ratio $A\beta_{1-42}/A\beta_{1-40}$ augmenté

Facteurs de risque génétique de maladie d'Alzheimer

B- Facteurs génétiques impliqués dans les formes sporadiques de maladie d'Alzheimer

Historiquement, le gène APOE codant pour l'apolipoprotéine E a été un des premiers candidats à être étudié. En effet, un pic de liaison avait été détecté dans la région correspondante sur le chromosome 19. De plus, une colocalisation de APOE dans les plaques amyloïdes avait déjà été retrouvée.

Les deux variants les plus fréquents du gène de l'APOE déterminent un haplotype à trois allèles, $\epsilon 2$, $\epsilon 3$ et $\epsilon 4$. Ces allèles sont responsables de modifications dans la séquence d'acides aminés de la protéine et de la synthèse de trois isoformes de l'APOE : APOE2 (cys112, cys158), APOE3 (cys112, arg158) et APOE4 (arg112, arg158).

Facteurs de risque génétique de maladie d'Alzheimer

B- Facteurs génétiques impliqués dans les formes sporadiques de maladie d'Alzheimer

Risque de maladie d'Alzheimer associé aux allèles du gène APOE chez des sujets de descendance européenne

Génotype APOE	Études hospitalière/autopsique			Études en population		
	N	OR	(CI 95%)	N	OR	(CI 95%)
$\epsilon 3/\epsilon 3$	2854	1.0	(référence)	2683	1.0	(référence)
$\epsilon 2/\epsilon 2$	21	0.6	(0.2-2.0)	36	0.9	(0.3-2.8)
$\epsilon 2/\epsilon 3$	447	0.6	(0.5-0.8)	568	0.6	(0.5-0.9)
$\epsilon 2/\epsilon 4$	141	2.6	(1.6-4.0)	152	1.2	(0.8-2.0)
$\epsilon 3/\epsilon 4$	2171	3.2	(2.8-3.8)	1226	2.7	(2.2-3.2)
$\epsilon 4/\epsilon 4$	671	14.9	(10.8-20.6)	193	12.5	(8.8-17.7)

Facteurs de risque génétique de maladie d'Alzheimer

B- Facteurs génétiques impliqués dans les formes sporadiques de maladie d'Alzheimer

Entre 1996 et 2005, plus de 1000 articles portant sur plus de 500 gènes sont publiés. Devant cette masse de résultats, la base Alzgene est créée afin d'évaluer chaque gène-candidat selon des critères reproductibles et réaliser si possible des méta-analyses. Cette démarche confirme l'APOE comme plus important facteur de risque génétique des formes tardives de la maladie d'Alzheimer. Elle permet également d'objectiver l'absence de répliation de la plupart des autres gènes-candidats.

Principaux variants génétiques significativement associés à la maladie d'Alzheimer dans les études pangénomiques ($P < 5 \times 10^{-8}$)

Phénotype	Auteur	Année de publication	Gènes proches	Région chromosomique	Meilleur SNP -allèle à risque	P
Maladie d'Alzheimer						
	Lambert JC ¹⁶⁶	2009	CR1	1q32.2	2-SNP haplotype	3e-10
	Naj AC ¹⁶⁹	2010	Intergenic	2q12.3	rs4676049-A	4e-08
	Hu X ¹⁷⁰	2011	BIN1	2q14.3	rs12989701-?	3e-10
	Hu X ¹⁷⁰	2011	BIN1	2q14.3	rs744373-?	1e-10
	Naj AC ¹⁶⁹	2010	MTHFD1L	6q25.1	rs11754661-A	2e-10
	Harold D ¹⁶⁷	2009	CLU	8p21.1	rs11136000-?	9e-10
	Lambert JC ¹⁶⁶	2009	CLU	8p21.1	3-SNP haplotype	6e-10
	Reiman EM ¹⁷¹	2007	GAB2	11q14.1	rs2373115-G	1e-10
	Harold D ¹⁶⁷	2009	PICALM	11q14.2	rs3851179-?	1e-09
	Abraham R ¹⁷²	2008	APOE, PVRL2, TOMM40	19q13.32	rs6859-A	6e-14
	Feulner TM ¹⁷³	2010	APOE, TOMM40	19q13.32	rs157580-?	1e-40
	Harold D ¹⁶⁷	2009	APOE, TOMM40	19q13.32	rs2075650-?	2e-157
	Lambert JC ¹⁶⁶	2009	APOE	19q13.32	rs2075650-?	2e-16
	Heinzen EL ¹⁷⁴	2010	APOE, TOMM40	19q13.32	rs2075650-?	3e-11
	Seshadri S ¹⁶⁸	2010	APOE	19q13.32	rs2075650-G	1e-295
	Naj AC ¹⁶⁹	2010	APOE, TOMM40	19q13.32	rs2075650-G	5e-36
	Coon KD ¹⁷⁵	2007	APOE	19q13.32	rs4420638-?	1e-39
	Webster JA ¹⁷⁶	2008	APOE	19q13.32	rs4420638-?	1e-39
	Li H ¹⁷⁷	2008	APOC1, APOE	19q13.32	rs4420638-?	2e-44
Biomarqueurs (LCR)						
AB1-42	Han MR ¹⁷⁸	2010	CYP19A1	15q21.2	rs2899472-A	2e-09
T-tau	Han MR ¹⁷⁸	2010	Intergenic	12q24.23	rs1997111-T	1e-08
Imagerie (IRM)						
Atrophie cérébrale	Furney SJ ¹⁷⁹	2011	ZNF292	6q14.3	rs1925690-?	3e-08
Volume cérébral	Seshadri S ¹⁸⁰	2007	CDH4	20q13.33	rs1970546-?	4e-08
Vieillessement						
	Lunetta KL ¹⁸¹	2007	Intergenic	2p12	rs10496265-?	1e-08
	Poduslo SE ¹⁸²	2010	LRP1B	2q22.1	rs12474609-?	6e-09