



## Qu'est-ce que le Projet international HapMap?

Le Projet international HapMap est le fruit d'une collaboration de plusieurs pays qui vise à identifier et à cataloguer les similitudes et les différences génétiques chez les êtres humains, donc à décrire les modèles des variations courantes dans les séquences de l'ADN humain. HapMap est une ressource clé pour les scientifiques à la recherche des gènes associés à la santé, à la maladie, et à la réponse aux médicaments et aux facteurs environnementaux.

Les chercheurs partout dans le monde ont maintenant librement accès à l'information produite par le projet HapMap et versée dans des bases de données publiques.

Le Projet international HapMap a officiellement commencé par une réunion qui s'est tenue du 27 au 29 octobre 2002. Il a réalisé son objectif qui était de dresser une carte en l'espace de trois ans. Les chercheurs de centres universitaires, de groupes de recherche biomédicale et sans but lucratif et d'entreprises privées au Canada, en Chine, au Japon, au Nigeria, au Royaume-Uni et aux États-Unis ont collaboré pour mener à terme le projet.

Vous trouverez à l'adresse suivante une liste des établissements participants et des sources de financement : [www.hapmap.org/groups.html](http://www.hapmap.org/groups.html).

## Pourquoi est-il important d'étudier les variations génétiques?

La plupart des maladies courantes, notamment le diabète, le cancer, les accidents cérébrovasculaires, les maladies cardiaques, la dépression et l'asthme, sont attribuables à des ensembles de gènes et à des facteurs environnementaux. Bien qu'environ 99,9% de la séquence d'ADN soit la même chez tous les humains, la fraction de la séquence qui varie joue un rôle important, car elle détermine les différences entre les individus en termes de risque de maladie et de réponse aux médicaments.

La découverte de ces variations de séquence d'ADN (SNP) qui contribuent aux risques de maladies courantes est l'une de nos meilleures chances d'élucider les causes complexes des maladies chez les humains et de mieux connaître les moyens de prévenir, de diagnostiquer et de traiter ces maladies.

## Que sont les polymorphismes nucléotidiques simples (SNP), les allèles et les génotypes?

Les sites où la séquence d'ADN du génome peut varier d'une seule base d'un individu à l'autre se nomment polymorphismes nucléotidiques simples (SNP). Par exemple, certaines personnes peuvent avoir, dans un chromosome, une base A en un site particulier, alors que d'autres ont une base G à cet endroit. Ces différentes formes sont nommées allèles.

...TAGC...  
...TGGC...

*Une séquence d'ADN contenant un site SNP. Les allèles A et G sont illustrés.*

Chaque individu possède deux copies de tous les chromosomes, exception faite des chromosomes sexuels. L'ensemble d'allèles appartenant à un individu se nomme le génotype. Le terme génotype peut désigner une paire d'allèles SNP en un site particulier ou en plusieurs sites dans le génome. La méthode de détermination du génotype d'un individu se nomme génotypage.

### Qu'est-ce qu'un haplotype?

Il existe environ 10 millions de SNP chez les humains. Les allèles SNP les plus rares ont une fréquence d'au moins 1%. Les allèles SNP rapprochés tendent à être demeurer regroupés lors de la transmission héréditaire. Un tel groupe d'allèles SNP associés dans une région d'un chromosome se nomme haplotype. La plupart des régions chromosomiques comptent seulement quelques haplotypes communs, qui représentent la majorité des variations génétiques entre les individus d'une population. Bien qu'une région chromosomique puisse contenir de nombreux SNP, la localisation de quelques SNP marqueurs suffit pour connaître la nature de la variation génétique dans une région.

..A..C..A..T..G..T..  
..A..C..C..G..C..T..  
..G..T..C..G..G..A..

*Une région chromosomique où seuls les SNP sont montrés. Trois haplotypes sont illustrés. Les deux SNP colorés suffisent à identifier (marquer) chacun des haplotypes. Par exemple, si les deux sites SNP marqueurs du chromosome portent les allèles A et T, on peut déduire qu'il s'agit du premier haplotype.*

HapMap vise à cartographier les variations courantes du génome humain. Cette cartographie comprendra les régions chromosomiques comportant des SNP fortement associés, les haplotypes de ces régions et les SNP marqueurs correspondants. Elle décrira également les régions chromosomiques où les associations entre les SNP sont plus faibles.

Les scientifiques à la recherche des gènes liés aux maladies, comme le diabète, compareront un groupe de personnes affectées par la maladie et un groupe de personnes saines. Les régions chromosomiques où la fréquence des haplotypes différera entre les deux groupes seront des candidates pour contenir des gènes liés à la maladie. En théorie, les chercheurs pourraient tenter de découvrir ces régions en entreprenant le génotypage des 10 millions de SNP, mais le coût des méthodes de génotypage actuelles est prohibitif. HapMap identifiera de 250 000 à 500 000 SNP marqueurs qui contiendront presque autant d'information de cartographie que les 10 millions de SNP.

## Quelles populations ont été échantillonnées?

La plupart des haplotypes courants se retrouvent dans toutes les populations humaines; cependant, leur fréquence peut varier d'une population à l'autre. C'est pourquoi on doit échantillonner plusieurs populations avant de faire le choix des SNP marqueurs. Des études préliminaires ont révélé suffisamment de variations de fréquence d'haplotypes entre des populations du Nigeria (Yorubas), du Japon, de la Chine et des États-Unis (échantillons de résidents de l'Utah originaires d'Europe du Nord et de l'Ouest) pour justifier le développement de HapMap au moyen d'une analyse à grande échelle des haplotypes de ces populations.

Forte de l'information recueillie auprès de ces populations, HapMap devrait être pertinente pour toutes les populations du monde. Cependant, pour évaluer le gain d'information que représenterait l'inclusion d'autres populations, une étude parallèle analysera les haplotypes d'une série de régions chromosomiques dans des échantillons de plusieurs populations additionnelles.

Plus précisément, les échantillons d'ADN pour le projet HapMap proviennent de 269 individus : Yorubas d'Ibadan, au Nigeria (30 trios parents-enfant adulte), Japonais de Tokyo (45 individus sans lien de parenté), Chinois Han de Beijing (45 individus sans lien de parenté) et résidents de l'Utah originaires de l'Europe du Nord et de l'Europe de l'Ouest (30 trios). Ce nombre d'échantillons a permis au projet d'identifier presque tous les haplotypes d'une fréquence de 5% ou plus.

La collecte des échantillons a suivi des protocoles approuvés par les comités d'éthique compétents. Elle a été soumise à un processus d'engagement dans la collectivité ou de consultation publique respectant la culture locale et a été conditionnelle au consentement éclairé des individus. Le processus d'engagement dans la collectivité a été conçu pour répondre aux questionnements propres à chaque culture à l'égard du projet et donner l'occasion aux collectivités participantes de contribuer aux procédures de consentement éclairé et au protocole de collecte des échantillons.

## Comment le projet a-t-il résolu les questions de nature éthique?

Le projet soulève de nombreuses questions d'éthique. Étant donné qu'aucun renseignement personnel n'est associé aux échantillons, le risque de violation de la vie privée des donateurs est minimal. Cependant, la population d'origine figure sur les échantillons pour permettre aux chercheurs de choisir les marqueurs SNP les plus utiles à la recherche ultérieure dans chaque population. Les marqueurs SNP ont été sélectionnés selon la fréquence des haplotypes. Les marqueurs SNP de certaines régions peuvent varier d'une population à l'autre, si les fréquences des haplotypes diffèrent considérablement. Les SNP et la fréquence des haplotypes de chaque population seront donc calculés aux fins de comparaison. Cette procédure comporte un risque de stigmatisation ou de discrimination, dans la mesure où il serait possible de découvrir qu'une variante associée à une maladie a une fréquence plus élevée dans une population. Les risques associés à cette variante pourraient alors être indûment généralisés à l'ensemble de la population.

En outre, on doit prendre garde au fait que les « populations » décrites dans le projet, dont la définition repose sur l'ascendance géographique, pourraient être interprétées comme des « races » (un concept surtout social) et considérées à tort comme des concepts précis et significatifs fondés sur la biologie. Des consultations publiques ont été entamées pour que les collectivités fassent entendre leurs inquiétudes à cet égard.

## Quelle a été la stratégie scientifique du projet?

Pour établir la carte HapMap, les échantillons ont fait l'objet de génotypage pour trois millions de SNP dans le génome humain. Au moment du lancement du projet, on comptait 2,6 millions de SNP dans la base de données publique dbSNP [<http://www.ncbi.nlm.nih.gov/SNP>]. Cependant, trop peu de SNP ont été identifiés dans de nombreuses régions chromosomiques, et de nombreux autres sont trop rares pour revêtir une quelconque utilité. Il s'est donc révélé nécessaire d'identifier des millions de SNP additionnels pour créer HapMap. Les chercheurs ont découvert six millions de nouveaux SNP.

Le génotypage a été effectué dans 11 centres au Canada, en Chine, au Japon, au Royaume-Uni et aux États-Unis. Dans la Phase I, chaque centre a génotypé tous les échantillons des chromosomes qui lui avaient été assignés. Les centres ont utilisé six techniques de génotypage différentes. La carte de la Phase I initiale a produit des données sur un million de SNP dans les échantillons HapMap, espacés uniformément sur tout le génome. Dans la Phase II, 2,1 autres millions de SNP ont été génotypés dans les échantillons. La qualité du génotypage a été contrôlée par la duplication des échantillons, le génotypage d'un ensemble de SNP de contrôle par tous les centres, la vérification, par un centre, d'un échantillon des génotypes produits par un autre centre et la comparaison des données qui chevauchent dans d'autres projets.

Le groupe canadien, dirigé par M. Tom Hudson et son équipe au Centre d'innovation de Génome Québec et de l'Université McGill, a procédé à l'analyse des chromosomes 2 et 4p, qui correspondent à environ 10 % du génome humain.

## Quelles sont les politiques concernant l'accès aux données et la propriété intellectuelle?

Le projet a publié dans le domaine public toutes les données qu'il a produites, de manière à ce que tous les chercheurs dans le monde puissent les consulter gratuitement.

Le projet ne prévoit pas d'études établissant de liens entre la variation génétique et des phénotypes comme le risque de contracter une maladie et la réponse à des médicaments, c'est-à-dire une « utilité spécifique ». Les participants au projet sont d'avis que des SNP, des génotypes ou des haplotypes pour lesquels aucune utilité spécifique n'a été identifiée ne représentent pas une invention brevetable. Les politiques d'accès aux données n'interdisent pas toutefois aux utilisateurs de faire des demandes de brevet pour des SNP ou des haplotypes pour lesquels ils ont démontré une utilité spécifique, dans la mesure où ils n'empêchent pas d'autres parties d'avoir accès aux données du projet.