

Glossaire général

Allèle : une des formes alternatives d'un locus. Dans une cellule diploïde, il y a deux allèles pour chaque locus (un allèle transmis par chaque parent), qui peuvent être identiques. Dans une population, on peut avoir plusieurs allèles pour un locus.

Annotation structurale : repérage des coordonnées des diverses structures dans le génome, telles que les gènes.

Annotation fonctionnelle : renseignements sur les fonctions des séquences, le plus souvent pour les gènes.

BAC : *Bacterial Artificial Chromosome*. Vecteur de clonage permettant l'obtention de clones bactériens contenant un grand fragment d'ADN génomique (taille > 100 kb*). Les BAC assemblés en contigs* sont à la base des cartes physiques du génome.

Carte cytogénétique : carte des chromosomes. Réalisée par localisation visuelle (FISH*) au microscope de fragments d'ADN sur les chromosomes au stade métaphase de la mitose.

Carte d'hybrides irradiés : réalisée en testant par PCR la présence ou l'absence de fragments d'ADN dans une collection de clones d'hybrides irradiés (RH*). Deux fragments d'ADN sont proches sur le génome s'ils sont trouvés fréquemment dans les mêmes clones.

Carte génétique : obtenue par l'étude de la ségrégation dans des familles ou des populations, de marqueurs polymorphes, soit moléculaires, soit phénotypiques, deux séquences étant d'autant plus proches qu'elles sont souvent transmises ensemble lors de la méiose.

Clonage positionnel : stratégie visant à identifier un gène responsable de l'expression d'un phénotype en utilisant des informations de position sur le génome.

Contig : ensemble de clones (le plus souvent des BAC*) ou de lectures de séquence ordonnés grâce à des informations sur leur parties chevauchantes.

Cosmide : vecteur de clonage permettant l'obtention de clones bactériens contenant des fragments d'ADN génomique de taille avoisinant les 50 kb*.

CNV : *Copy Number Variation* ; polymorphisme du génome correspondant à la variation du nombre de copies d'une séquence, pouvant dans certains cas contenir un ou plusieurs gènes.

Déséquilibre gamétique : pour deux loci quelconques, c'est le fait que la fréquence des haplotypes* estimée pour tous les gamètes est différente de celle attendue à partir du produit des fréquences alléliques de chaque locus. *Synonyme* : déséquilibre de liaison. *Contraire de* : équilibre gamétique.

Dominance : qualificatif de l'effet d'un allèle, dont une copie suffit à l'expression du phénotype* approprié. L'allèle A est dominant sur l'allèle a si l'hétérozygote* Aa a le même phénotype* que l'homozygote AA.

EST : *Expressed Sequence Tag* : séquences étiquettes (partielles) de transcrit, obtenues par séquençage aléatoire d'ARN.

Evaluation génomique : évaluation de la valeur génétique d'individus d'après leurs génotypes pour un ensemble de loci distribués sur le génome, d'après des équations établies à partir des performances d'individus de référence phénotypés et génotypés.

* = Voir définition.

Expression génique : études visant à estimer le niveau de production (expression) des gènes en fonction d'états physiologiques ou de tissus différents.

Exon : fraction de la partie codante d'un gène eucaryote. Les gènes des organismes eucaryotes sont le plus souvent fractionnés en plusieurs séquences d'ADN dans le génome, les exons, séparés entre eux par d'autres séquences (introns*).

FISH : *Fluorescent In Situ Hybridisation*. Hybridation de sondes d'ADN marquées à l'aide d'un fluorochrome, sur des chromosomes au stade métaphase de la mitose. Permet la réalisation de la carte cytogénétique.

Fingerprinting : technique permettant d'estimer très grossièrement la similarité entre des séquences d'ADN sans les séquencer, par la comparaison des longueurs de bandes produites par des enzymes de restriction coupant l'ADN à des sites précis.

Fosmide : vecteur de clonage permettant l'obtention de clones bactériens contenant des fragments d'ADN génomique de taille déterminée et égale à 40 kb*.

FPC : *FingerPrint Contig** ; contig* de clones (généralement des BAC*) ordonnés par la technique du fingerprinting, afin d'obtenir une carte physique du génome.

Génotype 1 : constitution génétique d'un individu. 2. Combinaison allélique* à un locus particulier, ex: *Aa* ou *aa*.

Haplotype : combinaison allélique spécifique pour des loci appartenant à un fragment de chromosome défini.

Héritabilité au sens strict : proportion de la variance phénotypique due à la variabilité des valeurs génétiques = proportion de la variance phénotypique due à la variance génétique additive.

Hétérozygote : individu ayant des allèles non identiques pour un locus* particulier ou pour plusieurs loci. Cette condition définit l'«hétérozygotie». *Contraire de*: homozygote.

Homologues : séquences similaires en raison d'une origine évolutive commune.

Hybride irradié : cellule hybride obtenue par fusion entre cellules hôte d'une espèce et donneuse d'une autre espèce, contenant une fraction aléatoire du génome de l'espèce donneuse, après cassures par irradiation, reconstitution aléatoire de chromosomes ou insertion dans des chromosomes de la cellule hôte et rétention partielle. Deux séquences proches sur le génome sont en probabilité dans les mêmes clones RH*, tandis que deux séquences distantes ont une probabilité faible d'être conservées ensemble.

IBD : pour *identity by descent*. Identité entre deux chromosomes (ou parties de chromosomes), liée à leur descendance d'un même chromosome ancestral.

Indel : Insertion – deletion ; polymorphisme de présence ou absence d'un ou plusieurs nucléotides.

Intron : séquence non-codante dans les gènes, séparant les exons, qui codent pour une protéine.

Kb : kilobase ; séquence de mille paires de bases (pb*).

Locus (pl. : loci) : Site sur un chromosome. Par extension, emplacement d'un gène ou d'un marqueur génétique sur un chromosome.

Marqueur génétique : séquence d'ADN dont le polymorphisme est employé pour identifier un emplacement particulier (locus) sur un chromosome particulier.

Mate-pair : séquences appariées (1 à 10 kb* de distance), produites en circularisant les fragments d'ADN, puis par séquençage à travers le point de jointure.

Mb : mégabase ; séquence d'un million de paires de bases (pb*) de longueur.

Orthologues : séquences homologues* entre deux espèces.

Paired-end : séquences appariées produites par la lecture des deux extrémités de courts fragments d'ADN (moins de 500 pb*) dans le cas des nouvelles technologies de séquençage.

Paralogues : séquences homologues* résultat de la duplication d'une séquence ancestrale dans le génome. Il s'agit de deux (ou plus) séquences similaires par homologie dans un même génome.

Pb : paire de base ; unité de séquence d'ADN, représentée par une base et sa complémentaire-inverse sur l'autre brin.

Phénotype : caractère observable d'un individu résultant des effets conjugués du génotype et du milieu.

Phylogénomique : utilise les méthodes de la génomique et de la phylogénie. Par la comparaison de génomes entiers, permet de mettre en évidence des pertes et gains de gènes dans les génomes, ainsi que leur variabilité moléculaire, afin (entre autres buts) d'aider à prédire leur fonctions.

Plasmide : vecteur de clonage permettant l'obtention de clones bactériens contenant des fragment d'ADN génomique de taille allant de 500 pb* à 10 kb* environ.

Polymorphisme d'ADN : existence de deux ou de plusieurs allèles* alternatifs à un locus.

Puce à ADN ou puce pangénomique : Système permettant pour un individu le génotypage simultané de très nombreux marqueurs génétiques (de quelques milliers à quelques centaines de milliers).

QTL : abréviation de locus à effets quantitatifs (de l'anglais *Quantitative Trait Locus*).

Récessivité : qualificatif de l'effet d'un allèle, où l'homozygotie* est nécessaire pour l'expression du phénotype* approprié. *opposé de* : dominance*.

RH : *Radiation Hybrid* (hybride irradié*)

Sanger (méthode de) : méthode de séquençage publiée en 1977 (Sanger *et al* 1977) et encore utilisée de nos jours avec les séquenceurs à électrophorèse capillaire.

Scaffold : ensemble de contigs* de séquence reliés entre eux par des informations apportées par des lectures appariées (mate-pairs* ou paired-ends*).

Sélection assistée par marqueurs (abréviation : SAM) : utilisation d'un jeu restreint de marqueurs de l'ADN pour améliorer la réponse à la sélection dans une population : les marqueurs sont choisis comme étroitement liés à un ou plusieurs loci cibles, qui sont souvent des loci à effets quantitatifs ou QTL*.

SNP : polymorphisme d'un seul nucléotide à une position particulière de la séquence d'ADN (abréviation de l'anglais *Single Nucleotide Polymorphism*).

Supercontig : nom alternatif pour les scaffolds*.

WGS : *Whole Genome Shotgun* ; production de lectures de séquence d'un génome entier de manière aléatoire.

