

MÉTHODOLOGIE ACTUELLE DE LA BIOLOGIE MOLÉCULAIRE POUR LA DÉTECTION DES MUTATIONS

LUCIAN NEGURĂ^{1*}, EUGEN CARASEVICI²,
ANCA-MIHAELA HUMĂ¹, VLAD ARTENIE¹

Mots-clef : détection des mutations, PCR-multiplex, screening moléculaire, séquençage

Résumé : De plus en plus de maladies héréditaires ou non s'avèrent être causées par le dysfonctionnement des enzymes impliquées dans les voies métaboliques, à leur tour défectueuses à cause des mutations essentielles au sein des gènes. La détection précise de ces mutations les plus fréquentes représente actuellement une priorité pour le génie génétique et pour la médecine moléculaire. Nous présentons ici quelques-unes des méthodes de détection modernes.

INTRODUCTION

La médecine moderne s'est vue ces dernières années confrontée à de nombreux dilemmes cliniques et thérapeutiques dont la compréhension et la résolution nécessitent d'une manière quasi-ubiquitaire les outils de la biologie moléculaire. De plus en plus de maladies, génétiques ou spontanées trouvent l'explication de leur causalité dans les processus métaboliques complexes dans lesquels sont impliquées de nombreuses enzymes. Certaines altérations moléculaires au niveau des gènes codant pour ces enzymes ont souvent été désignées comme responsables de l'absence ou du dysfonctionnement des molécules participant aux réactions du métabolisme. Appelées couramment des mutations, ces altérations, d'une grande variabilité au niveau du génome, nécessitent à l'heure actuelle une détection précise et une caractérisation adéquate, dans le but d'élaborer une interprétation pertinente et ciblée de leur implication dans les dérèglements métaboliques et fonctionnels. La compréhension du rôle des mutations au niveau moléculaire de l'ADN nous permettra à long terme l'élaboration de stratégies thérapeutiques pour leur correction ou même leur utilisation pharmaceutique dans le cadre de la nouvelle médecine moléculaire.

LES MUTATIONS GÉNÉTIQUES

Tout chercheur, aussi bien dans le domaine fondamental qu'appliqué, se voit un jour rencontrer la notion de mutation. On entend par mutation toute modification qui touche le matériel génétique ou génome d'une cellule, contenant le message héréditaire. Nous pouvons aborder ces modifications de deux points de vue distincts, d'un côté les mutations apparues spontanément dans la nature, transmises ou non aux générations suivantes ; d'un autre côté, nous pouvons envisager la possibilité d'induire nous-mêmes ces mutations. Trois aspects concernant les mutations sont essentiels du point de vue du scientifique.

Premièrement, elles peuvent entraîner la perte totale d'une information et la suppression de la protéine codée par le gène muté, ou alors une altération de la structure de cette protéine. Une protéine modifiée représente une molécule non fonctionnelle et ce dysfonctionnement est d'une importance capitale dans le **diagnostique moléculaire**, pouvant être à l'origine de la plupart des pathologies.

Il existe également des mutations n'affectant pas radicalement la structure ou la fonction de la protéine concernée, mais cependant variées et très répandues au sein des populations. Dès lors on peut parler de polymorphisme des individus, et il s'agit souvent de variations d'un seul nucléotide connues sous le nom de SNP ou SNIps (Single Nucleotide Polymorphism). Quoique non essentielles pour le fonctionnement de la cellule, ces variations présentent un grand intérêt dans la médecine légale et sont utilisées pour l'établissement de la paternité ou de l'identité.

Troisièmement, nous pouvons apercevoir les mutations comme un « carburant » de l'évolution des espèces dont le moteur serait la pression de sélection. Les mutations spontanées peuvent s'avérer à un moment donné utiles à l'organisme qui les porte et représenter par leur conservation un atout de survie, ou, au contraire, elles peuvent nuire à ce même organisme et conduire à l'extinction de l'espèce. Dans la recherche fondamentale, l'évolutionnisme moléculaire peut apporter l'explication de l'origine et des interrelations des espèces et des molécules qui les composent.

Nous nous intéressons dans cette étude, d'un point de vue diagnostique moléculaire médical, à ces mutations qui nuisent effectivement au bon déroulement du métabolisme en affectant l'intégralité et l'activité des protéines, ainsi qu'aux moyens existants pour leur détection et leur caractérisation.

Comme nous l'avons déjà mentionné, les altérations du patrimoine génétique peuvent être diverses et nombreuses, créant ainsi un immense éventail de mutations possibles, classifiées à l'heure actuelle selon de nombreux critères. Du point de vue épidémiologique et clinique, nous devons faire les distinctions suivantes :

- Il y a une grande différence entre les mutations dites essentielles (affectant l'intégralité/fonctionnalité de la protéine) et celles non essentielles (intervenant souvent dans des régions non codantes et non régulatrices de l'ADN). Il est bien évident que notre intérêt se porte sur la première de ces deux catégories.
- On distingue les mutations ponctuelles, affectant une seule base azotée (délétions, insertions, mutations mis-sens ou non-sens) des mutations larges, affectant jusqu'à des régions chromosomiques entières (délétions, insertions, inversions, répétitions, translocations) ; les deux catégories sont intéressantes pour le clinicien.
- Une nette distinction doit être faite entre mutations somatiques et germinales. Si une mutation a lieu dans une cellule de la lignée germinale (faisant partie de l'ensemble des cellules qui donnent les gamètes), elle sera transmise à la descendance de ces cellules. La mutation est alors héréditaire. La majorité de ces mutations empêchent le développement de la cellule œuf. Si l'œuf est en mesure de se développer, l'individu ainsi généré souffre de malformations diverses. Si une mutation a lieu dans une cellule somatique (ne faisant pas partie de la lignée germinale), elle ne peut être transmise à un descendant, elle n'est pas héréditaire. Si les mutations somatiques présentent un grand intérêt pour la thérapie génique en oncologie, à celles germinales s'intéressent l'épidémiologie moléculaire, la statistique et surtout la génétique médicale humaine.
- Finalement, et c'est le point de départ du choix des méthodes de détection, il existe pour chaque gène des mutations connues, répétitives d'un individu à un autre et stockées dans des bases de données internationales, mais surtout énormément de mutations inconnues, apparues de novo ou caractéristiques d'une seule lignée familiale ou d'un seul individu. Pour la détection de ces dernières, l'association de plusieurs techniques de détection est absolument nécessaire.

La détection de mutations dans l'ADN représente une étape essentielle de la biologie moléculaire, pour la recherche fondamentale comme pour les applications médicales. Les stratégies actuellement disponibles pour déceler les allèles mutés sont fondées, soit sur l'utilisation de méthodes rapides mais limitées à l'identification d'un nombre restreint de mutants déterminés, soit sur l'utilisation de techniques coûteuses et laborieuses mais capables de déceler la moindre altération de séquence dans les portions essentielles des gènes.

DÉTECTION D'UNE MUTATION CONNUE PAR LA RÉACTION PCR-MULTIPLEX

Lorsque nous connaissons exactement la nature et l'emplacement d'une certaine mutation à détecter, nous pouvons cibler sur le génome la région concernée, voire le gène d'intérêt, et utiliser une technique simple et extrêmement efficace. Nous n'insisterons pas ici sur la description détaillée de la réaction PCR (Polymerase Chain Reaction). Depuis sa description en 1985 et la découverte de la Taq DNA Polymérase thermorésistante, la PCR a totalement révolutionné les méthodes d'analyse moléculaire. A tel point que cette dernière est aujourd'hui appliquée à l'étude de la quasi totalité des organismes (virus, bactéries, parasites, plantes, cellules et tissus animaux et humains). En l'espace de deux décennies, la PCR s'est imposée comme une technologie particulièrement performante et accessible pour un grand nombre d'applications et de développements. De plus la PCR a été depuis 1988 déclinée en une multitude de variantes dont la liste ne cesse de s'allonger chaque jour: RT-PCR, PCR-Multiplex, PCR-in Situ, Real-Time-PCR, etc.

Sans nécessiter de l'appareillage ou des réactifs supplémentaires comparatif à la réaction classique, la PCR-multiplex présente l'avantage de l'utilisation de plusieurs amorces (primers) au sein de la même réaction, en conséquent la possibilité de multiples amplifications à sélectivité dirigée. La difficulté initiale dans l'établissement d'un tel protocole réside du fait même des différences de propriétés physico-chimiques entre les amorces utilisées, différence croissante avec le nombre de ces dernières. Le paramètre essentiel à prendre en compte est le T_m (Melting temperature) de chaque amorce, valeur caractérisant sa capacité de fixation plus ou moins spécifique fonction des conditions de réaction et qui influence directement la température d'hybridation. Dans l'élaboration du protocole, le but est la réalisation d'un contexte de conditions intermédiaires optimales pour chacun des participants à la réaction, ceci par modulation des températures, des concentrations et du rapport matrice/amorce, des quantités de dNTP, polymérase ou magnésium

Quoique la réaction PCR-Multiplex est le plus souvent utilisée préalablement au séquençage, cas où les amorces vont par paires et où la notion de multiplex concerne le nombre d'amplicons utilisés simultanément, il existe des variantes de la même technique, utilisées dans le but précis de détection d'une mutation, mettant en compétition des amorces complémentaires au même locus, mais de séquences différentes. Parmi toutes les possibilités proposées dans la littérature de spécialité, on peut mentionner :

- La variante avec trois amorces simultanément [Chan, 1999], dont l'amorce sens est commune et on utilise une amorce antisens spécifique de la région normale, non mutée, ainsi qu'une autre dégénérée mutation-spécifique. L'interprétation du résultat se fait dans ce cas en fonction de la taille des fragments obtenus.
- La variante avec quatre amorces simultanément [Fitzgerald, 1996], deux extérieures au locus d'intérêt, une spécifique au locus non muté et une mutation-spécifique. Trois fragments différents peuvent être obtenus par une telle combinaison, l'interprétation des résultats (mutation/non, homozygotie/hétérozygotie) se faisant par la taille, mais également par le nombre de fragments obtenus.

Dans la figure 1 nous présentons une expérience de Multiplex à 4 amorces effectuée dans la région de l'exon 2 du gène BRCA1. Quoique la mutation recherchée initialement (185delAG) n'a pas été retrouvée chez les patients investigués, on observe une bande supplémentaire (pointée dans la figure par une flèche) correspondant à toute évidence à une insertion/duplication dans la

région d'intérêt. On a probablement à faire à une mutation inactivant le gène BRCA1, ce qui pourrait être à l'origine du cancer de sein développé par le patient concerné.



Figure 1. Détection d'une possible mutation sur BRCA1 par PCR-Multiplex

Cet exemple nous emmène devant les limites de cette technique, à savoir qu'aucun résultat PCR ne puisse être validé sans une confirmation par des techniques de séquençage. Néanmoins, cette première expérience nous prouve la nécessité de continuer l'investigation par d'autres techniques et nous indique plus précisément la région du gène à investiguer.

DÉTECTION DES MUTATIONS INCONNUES : LE SCREENING MOLÉCULAIRE

La plupart des mutations affectant les gènes ne sont pas répertoriées dans les bases de données, leur identification étant bien souvent une première, caractéristique d'un individu ou d'une lignée familiale. La détection de telles mutations est une démarche complexe et aucune des méthodes disponibles n'est applicable seule à toutes les situations, leur choix dépendant de nombreux critères (nature des mutations, taille et structures du gène, accès à l'ARNm, degrés d'efficacité et de sensibilité recherchés). Parmi les techniques courantes de balayage d'un fragment d'ADN ou d'ARN on peut mentionner l'électrophorèse en gel de gradient dénaturant (DGGE), l'analyse d'hétéro duplex (HA), le polymorphisme de conformation de l'ADN en simple brin (SSCP) ou le clivage chimique ou enzymatique de mésappariements (CCM/EMC). Des techniques plus ciblées comme le test de troncation protéique (PTT) et bien évidemment le séquençage systématique peuvent y être ajoutées.

LE TEST DE TRONCATION DES PROTÉINES (PROTEIN TRUNCATION TEST) (PTT)

Le test de troncation des protéines (PTT, *protein truncation test*) s'avère particulièrement précieux pour la détection spécifique des mutations non-sens ou qui décalent le cadre de lecture de l'ARNm et provoquent une terminaison de traduction prématurée [Hogervorst, 1995]. À la différence des autres techniques de criblage, les mutations ne sont pas détectées au niveau de l'ADN, mais au niveau des produits polypeptidiques obtenus après transcription et traduction *in vitro* des fragments d'ADN génomique ou d'ADNc amplifiés à partir des régions d'intérêt du

gène étudié. Le principe de base de la technique est la différence de taille (et par conséquent de migration électrophorétique) entre une protéine normale et une variante mutante, tronquée.

Les produits PCR sont transcrits et traduits *in vitro*. Après séparation électrophorétique des produits de traduction et autoradiographie, la visualisation d'un peptide tronqué de poids moléculaire inférieur au produit normal indique la présence d'une mutation terminatrice dans le fragment d'ADN ou d'ARN constituant le matériel de départ. En général, l'estimation de la taille de la bande anormale permet de localiser approximativement le site muté dans le fragment analysé, de sorte que le séquençage destiné à identifier la mutation sera limité finalement à une très petite portion du gène. La taille des exons étant généralement de l'ordre de 100 à 300 paires de bases, il est évident qu'il est plus attirant de cribler la version ADNc en quelques fragments de RT-PCR plutôt que chacun des exons séparément. Cependant, la méthode peut aussi s'appliquer à l'exploration directe d'ADN génomique amplifié lorsqu'on recherche des mutations introduisant un signal prématuré de terminaison de traduction dans des exons dont la taille est exceptionnelle.

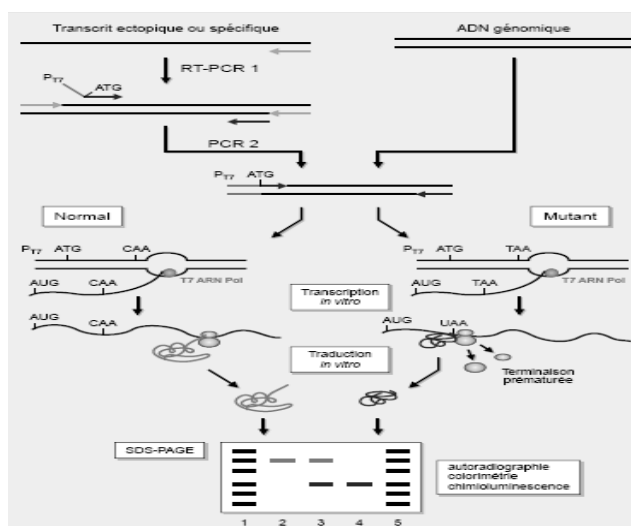


Figure 2. Les différentes étapes du Test de Troncation Protéique

Les mutations pouvant être détectées par le PTT sont de trois types distincts mais ayant la même conséquence : mutations non-sens (menant à l'apparition d'un codon STOP), mutations frameshift (menant à un décalage du cadre de lecture et souvent à l'apparition d'un codon STOP) et mutations dans les sites d'épissage (menant à des variations d'élimination des introns au niveau des ARNm).

L'élément clé du PTT est représenté par l'amorce sens de l'amplification, une construction complexe constituée de :

- 5' séquence promoteur pour l'ARN Polymérase (habituellement la T7 ARN Polymérase)
- séquence de Kozak pour l'initiation de la translation + un espaceur de 3-6 nucléotides
- séquence complémentaire du gène d'intérêt, en phase avec le codon initiateur ATG
- 3' site de restriction pour un éventuel clonage par des vecteurs.

Tableau nr. 1. Avantages et inconvénients de la technique PTT

AVANTAGES	INCONVÉNIENTS
<ul style="list-style-type: none"> - Mise en évidence des protéines incomplètes, donc non fonctionnelles ; - Analyse de séquences codantes jusqu'à 5 kilobases ; - Localisation du site muté fonction de la longueur de la protéine tronquée. 	<ul style="list-style-type: none"> - Limitée à la détection uniquement des mutations inactivatrices (10% des mutations pour certains gènes ; - Plus fiable dans le cas de l'utilisation de l'ARNm, mais qui est instable et plus difficile à manipuler ; - Ne tient pas compte de l'épissage alternatif et des populations hétérogènes d'ARNm.

LE POLYMORPHISME DE LONGUEUR DES FRAGMENTS DE RESTRICTION (RESTRICTION FRAGMENT LENGTH POLYMORPHISM) (RFLP)

Les applications du RFLP en biologie moléculaire ont à la base deux notions fondamentales :

- ❖ **Les Enzymes de Restriction** – endonucléases qui reconnaissent des séquences spécifique de l'ADN double brin et scindent spécifiquement au niveau de ces sites ;
- ❖ **Le Polymorphisme** – différence génétique héréditaire au sein d'une population. Peut exister au niveau d'un seul nucléotide, cas des SNP (Single Nucleotide Polymorphism).

Les SNP ou des mutations peuvent causer l'apparition/disparition d'un site de restriction et donc une modification du profile de restriction.

De ces notions émerge la notion de RFLP, qui définit concrètement certains fragments d'ADN génomique humain cloné qui, lorsqu'ils sont utilisés comme sondes sur de l'ADN génomique humain digéré par un enzyme de restriction, permettent la détection, pour un locus donné, des fragments dont la taille est différente selon les individus (correspondant aux différents allèles de ce locus). Ce polymorphisme peut être dû soit à la répartition des sites de restriction, soit à la présence de minisatellites diversement répétés. La taille des fragments reconnus dépend de la position des sites de coupures qui encadrent ou sont internes à la région reconnue par la sonde, ou du nombre de répétitions du minisatellite.

Le principe du RFLP est assez simple. Après amplification de la région d'intérêt par PCR, on réalise la digestion des fragments obtenus par l'enzyme (les enzymes) de restriction et on effectue une migration en gel d'agarose ou polyacrylamide, suivie de la visualisation au bromure d'éthidium sous UV. Pour les fragments génomiques de grande taille, l'élaboration d'une sonde à séquence complémentaire au fragment « cible » et l'analyse par Southern Blot seront nécessaires [Horning, 2006].

Les applications de la technique RFLP sont nombreuses et variées, parmi les plus utilisées étant la détermination de la paternité (comme le montre l'exemple dans la figure 3), analyse de l'identité en criminalistique, analyse de la transmission des maladies héréditaires ou établissement de cartes génétiques.

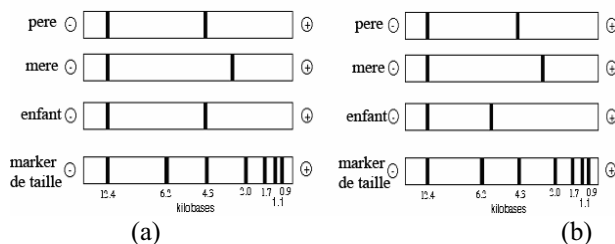


Figure 3. Détermination de la paternité par RFLP. (a) Potentiellement positif ; (b) négatif

L'ÉLECTROPHORÈSE EN GEL AU GRADIENT DÉNATURANT (DENATURING GEL GRADIENT ELECTROPHORESIS) (DGGE)

Cette méthode est basée sur la différence de mobilité électrophorétique entre deux molécules d'acides nucléiques qui diffèrent par une variation de séquence. La mobilité électrophorétique d'un fragment d'ADN double brin dans un gel de polyacrylamide est affectée par la séparation partielle de ses deux brins. Une molécule composée en partie de la double hélice classique et en partie de deux simples brins est fortement ralentie par rapport à une molécule double brin ou complètement dénaturée. Dans des conditions dénaturantes (urée-formamide), le T_m du double brin d'ADN dépend de la séquence et la mobilité électrophorétique sera d'autant réduite qu'il atteindra le T_m tôt lors de la migration [Kuperstein, 2006].

Après amplification PCR des régions susceptibles de contenir une mutation, l'ADN de ces fragments migre dans un gel d'acrylamide où il rencontre des conditions de plus en plus dénaturantes (gradient en acrylamide). Cependant, si le gradient augmente de façon linéaire, la dénaturation de l'ADN, elle, ne se fait pas de façon progressive tout au long de la molécule, mais de domaine en domaine, qui s'ouvrent plus ou moins rapidement, en fonction de leur composition en bases AT et GC. La stabilité de chacun des domaines dépend de la structure primaire de la molécule. Deux molécules dont les séquences ne différeront que d'une paire de base pourront donc ne pas s'ouvrir au même moment et migreront alors différemment à travers un gradient précis.

LE POLYMORPHISME DE CONFORMATION DES SIMPLE BRINS (SINGLE STRAND CONFORMATION POLYMORPHISM) (SSCP)

La technique [Kozlowski, 2001] est basée sur la capacité des simple brins d'ADN de pouvoir prendre dans des conditions NON dénaturantes différentes conformations 3D selon la séquence de départ. Deux brins ADN différents en structure primaire adopteront des structures spatiales différentes, ce qui se traduit en une différence de mobilité électrophorétique.

Après dénaturation thermique des fragments ADN amplifiés, on procède à un refroidissement brusque, pour que les molécules restent monocaténares et adoptent les conformations les plus stables pour chacune d'entre elles. Les profils de migration en gel d'acrylamide non dénaturant sont comparés en partant de l'idée qu'une différence de migration signifie une différence de séquence et par conséquent un polymorphisme ou une mutation (figure 4).

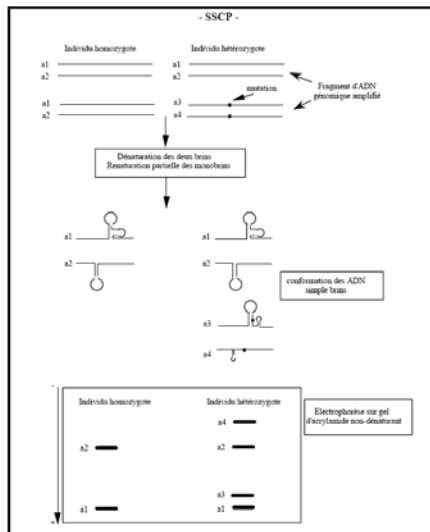


Figure 4. Les étapes du SSCP

CONCLUSIONS ET PERSPECTIVES

La multitude des méthodes existantes à l'heure actuelle rend d'autant plus complexe le choix judicieux de la méthode la plus adaptée au gène d'intérêt, l'association de plusieurs de ces méthodes en systèmes de détection performantes et fiables s'avérant très souvent nécessaire. Pour chacune des méthodes utilisées, il est indispensable d'effectuer le séquençage ultérieur de la région susceptible de contenir une mutation, afin de confirmer de façon indubitable la présence de cette mutation, spécialement dans le cas du diagnostic clinique. Généralement, ces méthodes sont proposées par les fournisseurs sous forme de kits, le prix élevé et la complexité de mise au point de ces derniers étant compensés par la précision et la fiabilité des systèmes modernes de détection.

RÉFÉRENCES

1. Chan P., Wong B., Ozcelik H. and Cole D., 1999, Simple and rapid detection of BRCA1 and BRCA2 mutations by multiplex mutagenically separated PCR. *Clin Chem.*, **45(8)**, 1285-1287.
2. Fitzgerald M.G. et al., 1996, Germ-line BRCA1 mutations in Jewish and non-Jewish women with early-onset breast cancer. *N Engl J Med.*, **334(3)**, 143-149.
3. Hogrevorst F.B. et al., 1995, Rapid detection of BRCA1 mutations by the protein truncation test. *Nat Genet.*, **10(2)**, 208-212.
4. Horning M.E. and Cronn R.C., 2006, Length polymorphism scanning is an efficient approach for revealing chloroplast DNA variation. *Genome.*, **49(2)**, 134-142.
5. Kuperstein G., Jack E. and Narod S.A., 2006, A fluorescent multiplex-DGGE screening test for mutations in the BRCA1 gene. *Genet Test.*, **10(1)**, 1-7.
6. Kozłowski P. and Krzyżosiak W., 2001, Combined SSCP/duplex analysis by capillary electrophoresis for more efficient mutation detection. *Nucleic Acids Res.*, **29(14)**, E71.

1) Université «Alexandru Ioan Cuza», Iași, Roumanie

2) Université de Médecine et Pharmacie « Gr.T.Popa », Iași, Roumanie

* correspondance à adresser à : luciannegura@yahoo.fr