

4. Cho CH, Kammerer RA, Lee HJ, et al. Designed angiopoietin-1 variant, COMP-Ang1, protects against radiation-induced endothelial cell apoptosis. *Proc Natl Acad Sci USA* 2004; 101: 5553-8.
5. Schuller BW, Binns PJ, Riley KJ, et al. Selective irradiation of the vascular endothelium has no effect on the survival of murine intestinal crypt stem cells. *Proc Natl Acad Sci USA* 2006; 103: 3787-92.
6. Bhanja P, Saha S, Kabarriti R, et al. Protective role of R-spondin1, an intestinal stem cell growth factor, against radiation-induced gastrointestinal syndrome in mice. *PLoS One* 2009; 4: e8014.
7. Qiu W, Carson-Walter EB, Liu H, et al. PUMA regulates intestinal progenitor cell radiosensitivity and gastrointestinal syndrome. *Cell Stem Cell* 2008; 2: 576-83.
8. Komarova EA, Kondratov RV, Wang K, et al. Dual effect of p53 on radiation sensitivity *in vivo*: p53 promotes hematopoietic injury, but protects from gastro-intestinal syndrome in mice. *Oncogene* 2004; 23: 3265-71.
9. Merritt AJ, Allen TD, Potten CS, Hickman JA. Apoptosis in small intestinal epithelial from p53-null mice: evidence for a delayed, p53-independent G2/M-associated cell death after gamma-irradiation. *Oncogene* 1997; 14: 2759-66.
10. Rotolo JA, Maj JG, Feldman R, et al. Bax and Bak do not exhibit functional redundancy in mediating radiation-induced endothelial apoptosis in the intestinal mucosa. *Int J Radiat Oncol Biol Phys* 2008; 70: 804-15.
11. Terry NH, Travis EL. The influence of bone marrow depletion on intestinal radiation damage. *Int J Radiat Oncol Biol Phys* 1989; 17: 569-73.
12. Joubert D, Hollande F, Jay P, Legraverend C. Les cellules souches intestinales : 30 ans d'une histoire exemplaire. *Med Sci (Paris)* 2009; 25: 441-4.
13. François A, Milliat F, Jullien N, et al. Radiothérapie : quelles orientations thérapeutiques contre les séquelles digestives ? *Med Sci (Paris)* 2009; 25: 267-72.

## NOUVELLE

### Le séquençage de l'exome au secours des maladies génétiques

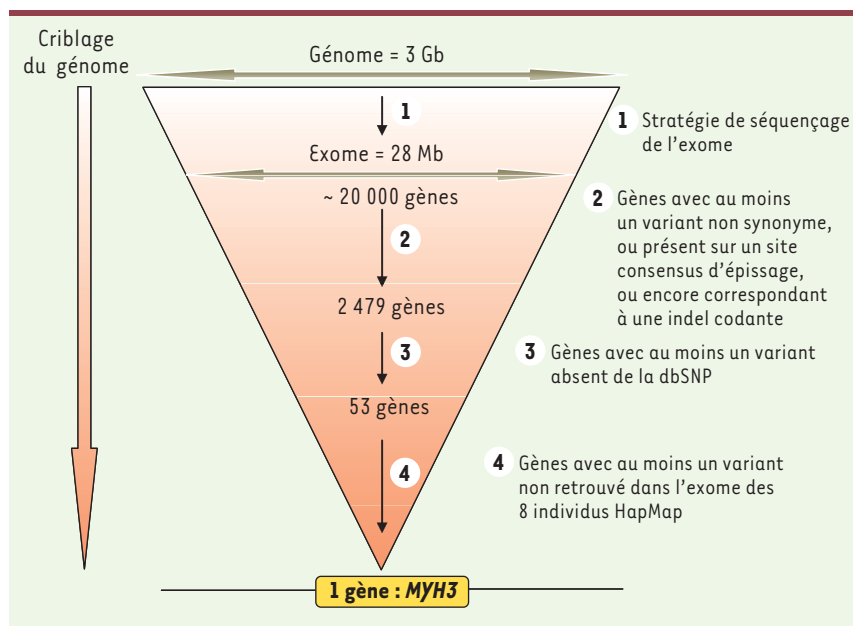
Audrey Criqui, Sylvain Baulande

PartnerChip-CEA, Bâtiment G2,  
2, rue Gaston Crémieux,  
91000 Évry, France.  
[sylvain.baulande@cea.fr](mailto:sylvain.baulande@cea.fr)

> Avec plus de 6 000 pathologies répertoriées, les maladies génétiques constituent un grand problème de santé publique affectant plus de 6 % de la population mondiale dont 25 millions d'Européens. La prise en charge de ces maladies demeure compliquée en raison de la difficulté tant du diagnostic que du développement de thérapies pour ces pathologies qui regroupent peu de malades. Actuellement, le diagnostic génétique est principalement fondé sur la recherche d'anomalies cytogénétiques, la mise en œuvre d'analyses de liaison génétique ou encore le re-séquençage de gènes candidats. Malheureusement, l'identification des gènes en cause s'avère très difficile en raison de la diversité des mutations à mettre en évidence parmi les 3 milliards de bases qui composent le génome humain. Ainsi, beaucoup trop de patients restent sans diagnostic moléculaire bien que ce soit nécessaire à la prise en charge des familles, à la mise en place d'un diagnostic prénatal et que cela constitue une étape indispensable à l'élaboration de stratégies thérapeutiques. L'émergence des technologies de séquençage de nouvelle génération (dites NGS

pour *next generation sequencing*) [1] ouvre de nouvelles perspectives pour le diagnostic des maladies génétiques. Jusqu'à présent, en raison du coût important du séquençage conventionnel, seuls quelques privilégiés comme le prix Nobel James Watson [2] codécouvreur de l'ADN avec Francis Crick, ou encore Craig Venter [3], pionnier du séquençage du génome humain, ont vu leur génome entièrement séquencé. Ces nouvelles technologies, qui permettent une lecture massivement parallèle de plusieurs Gb par semaine, ont permis de réduire considérablement ces coûts et vont donc faciliter le séquençage complet de génomes à haut débit. Cette technologie a donc logiquement été retenue dans le cadre du *1 000 genomes project* qui prévoit le séquençage du génome de 1 000 individus afin d'établir une nouvelle base de données plus complète des polymorphismes du génome humain [4, 5]. Malgré l'avancée significative apportée par le NGS, ces techniques restent encore trop onéreuses pour envisager le séquençage à grande échelle du génome entier d'individus atteints d'une même maladie génétique. De plus, les millions de polymorphismes

ainsi révélés rendent la démarche de recherche de mutations particulièrement délicate et nécessitant donc une méthodologie adaptée. Récemment, un consortium américain baptisé *Exome Project* a été créé afin de développer une approche centrée sur le séquençage exclusif de l'« exome » correspondant à l'ensemble des parties codantes du génome. Bien qu'il n'en représente qu'environ 1 % (soit près de 28 Mb selon la base CCDS [*consensus coding DNA sequence project*] du NCBI [*National center for biotechnology information*]), l'exome contient à lui seul 85 % des mutations mises en cause dans les maladies génétiques [6]. Une équipe de l'université de Washington impliquée dans ce consortium a donc isolé les 180 000 régions génomiques correspondantes par une technique de capture sur puce à ADN. Ces régions ont ensuite été amplifiées puis séquencées à l'aide d'une plateforme Solexa d'Illumina [7]. Dans un premier temps, ces chercheurs ont étudié l'exome de huit personnes originaires de 3 populations différentes (4 Africains Yoruba, 2 Asiatiques et 2 Caucasiens) dont la séquence avait



**Figure 1. Stratégie de criblage du génome de patients atteints du syndrome de Freeman-Sheldon.** La pyramide inversée représente le nombre de gènes variants communs aux 4 patients au cours de l'application de différents filtres. En partant de plus de 20 000 gènes (l'exome), seuls les gènes ayant au moins un cSNP non synonyme, un variant touchant un site responsable de l'épissage, ou encore une indel sont sélectionnés. Un second filtre est appliqué afin d'exclure parmi ces variants ceux présents dans la dbSNP et dans l'exome des 8 individus HapMap. Finalement, l'unique gène restant correspond au gène *MYH3*, responsable du syndrome de Freeman-Sheldon.

déjà fait l'objet d'investigations dans le cadre du projet HapMap [8].

Au total, 6,4 Gb de séquences ont été générées par individu, ce qui a permis d'obtenir une excellente couverture moyenne d'environ 50, correspondant au nombre de fois qu'une position donnée du génome a été séquencée. Ainsi, plus de 96 % des positions génomiques séquencées présentent une couverture supérieure à 8 fois, jugée suffisante pour une bonne fiabilité de lecture. La comparaison de la séquence des 8 exomes d'individus de référence avec les variations identifiées dans le cadre du projet HapMap permet d'observer 99,94 % de concordance pour les génotypes homozygotes et 99,57 % pour les hétérozygotes, reflétant un faible taux d'erreur.

En comparant avec la séquence de référence du génome humain, les auteurs ont retrouvé 56 240 SNP codants (cSNP) dont 13 347 non décrits dans la base de données de SNP du NCBI (dbSNP). Concernant les autres types de polymorphismes, chaque individu présente en moyenne 17 positions variantes situées sur des sites responsables de l'épissage et 166 insertions ou délétions (indels) codantes dont respectivement 30 % et 37 % non annotées dans la dbSNP. Ainsi,

les nouveaux polymorphismes identifiés chez ces individus renforcent de manière significative le catalogue de la dbSNP.

Dans un second temps, les chercheurs ont étudié l'exome de 4 patients non apparentés atteints du syndrome de Freeman-Sheldon (SFS)<sup>1</sup> [7], une maladie héréditaire dominante rare causée par des mutations du gène *MYH3* (codant l'*embryonic myosin heavy chain*) [9]. Ils ont alors mis en place une stratégie de criblage des gènes afin de voir s'il était possible de retrouver le gène *MYH3* muté chez ces patients. Ce criblage repose sur l'examen des différents variants portés par chacun des gènes et identifiés lors du séquençage exomique (Figure 1). La première étape consiste à rechercher les gènes qui présentent au moins un cSNP non synonyme, une indel ou un variant affectant l'épissage chez tous les patients. Ainsi, 2 479 gènes ont été sélectionnés à partir des 20 000 gènes codés par l'exome. Les gènes ne contenant que des polymorphismes déjà décrits dans la dbSNP ont été écartés, ce qui a permis de réduire le nombre de

gènes variants en commun à 53. Ensuite, les gènes ne contenant que des variants identifiés lors du séquençage exomique des 8 individus HapMap ont aussi été rejetés. À l'issue de cette dernière étape, il ne restait alors qu'un seul gène, le gène *MYH3*, présentant chez chaque patient un variant (mutation causale) qui n'était répertorié ni dans la dbSNP, ni parmi les 8 exomes HapMap. L'approche par séquençage d'exome permet donc de retrouver la mutation causale déjà connue pour cette affection.

Plus récemment, cette même équipe a publié des travaux similaires basés sur le séquençage exomique du génome de 4 patients atteints du syndrome de Miller<sup>2</sup> [10] pour lequel aucune mutation n'avait encore été décrite. En appliquant la même stratégie de criblage des gènes communs aux malades, le gène en cause, appelé *DHODH* (codant la *dihydroorotate dehydrogenase*), a pu être identifié et validé [11]. La méthode se montre donc capable de révéler le gène impliqué dans une pathologie en l'absence de toute indication préalable.

<sup>1</sup> Le syndrome de Freeman-Sheldon (SFS), aussi connu comme « syndrome de la face du siffleur », est un syndrome congénital rare caractérisé par une dysmorphie combinant des anomalies osseuses et des contractures articulaires avec un faciès typique (extrait du site Orphanet).

<sup>2</sup> Syndrome de Miller : dysostose acrofaciale (DAF) post-axiale. Les DAF constituent un groupe hétérogène d'affections associant des anomalies des membres et du développement facial/mandibulaire (site Orphanet).

De qualité équivalente au séquençage conventionnel mais à moindre coût, l'approche du séquençage de l'exome s'avère très efficace pour identifier le gène et les mutations responsables de maladies monogéniques. La puissance de cette démarche s'appuie sur la qualité des bases de données des polymorphismes humains comme la dbSNP, la plus prometteuse étant celle en cours d'élaboration par le « 1 000 genomes project » [5]. Sachant que sur les 6 000 maladies génétiques connues seuls 1 600 gènes responsables ont été identifiés, l'émergence de ces nouvelles technologies est porteuse d'un réel espoir pour plus de 75 % des patients atteints de maladies génétiques chez

lesquels aucun diagnostic génétique n'est possible à ce jour. ♦

### Exome sequencing applied to monogenic disorders

#### CONFLIT D'INTÉRÊTS

Les auteurs déclarent n'avoir aucun conflit d'intérêts concernant les données publiées dans cet article.

#### RÉFÉRENCES

1. Jordan B. La génération suivante, déjà... *Med Sci (Paris)* 2009 ; 25 : 649-50.
2. Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 2008 ; 452 : 872-6.
3. Levy S, Sutton G, Ng PC, et al. The diploid genome sequence of an individual human. *PLoS Biol* 2007 ; 5 : e254.

4. Siva N. 1 000 Genomes project. *Nat Biotechnol* 2008 ; 26 : 256.
5. Jordan B. Un, deux, trois... mille génomes ? *Med Sci (Paris)* 2008 ; 24 : 237-8.
6. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA* 2009 ; 27 octobre (online).
7. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 2009 ; 461 : 272-6.
8. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007 ; 449 : 851-61.
9. Toydemir RM, Rutherford A, Whitby FG, et al. Mutations in embryonic myosin heavy chain (MYH3) cause Freeman-Sheldon syndrome and Sheldon-Hall syndrome. *Nat Genet* 2006 ; 38 : 561-5.
10. Miller M, Fineman R, Smith DW. Postaxial acrofacial dysostosis syndrome. *J Pediatr* 1979 ; 95 : 970-5.
11. Ng SB, Buckingham K, Lee C, et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet* 2010 ; 42 : 31-6.

## NOUVELLE

### Les cellules du muscle chantent en chœur une berceuse pour cellules souches

Rana Abou-Khalil, Terry Partridge, Bénédicte Chazaud

R. Abou-Khalil, B. Chazaud :

Inserm U567, Paris, 75014, France ;  
 Institut Cochin, CNRS, UMR8104, Paris, 75014, France ;  
 Université Paris Descartes, Paris, 75014, France.  
 Département génétique et développement,  
 24, rue du Faubourg Saint-Jacques, 75014 Paris, France.  
[benedicte.chazaud@inserm.fr](mailto:benedicte.chazaud@inserm.fr)  
 T. Partridge : Center for Genetic Medicine,  
 Children's National Medical Center,  
 Washington DC 20010, États-Unis.

#### Cellules satellites, précurseurs myogéniques et cellules de réserve dans le muscle

Le muscle strié squelettique adulte peut se régénérer grâce à une population de précurseurs, les cellules satellites. Dans le muscle adulte normal, les cellules satellites sont quiescentes et logées le long des fibres musculaires [1]. Leur capacité à assurer à la fois le renouvellement continu des noyaux des fibres musculaires et la régénération musculaire suggère l'existence de mécanismes qui permettent de maintenir un réservoir de cellules satellites tout au long de la vie [2]. *In vivo*, la preuve de l'autorenouvellement des cellules satellites a

été obtenue après greffe de fibres isolées [3] ou de populations de cellules satellites quiescentes [4]. Des études *ex vivo* faites sur des fibres isolées qui étaient en phase de prolifération ont montré qu'un petit nombre des cellules précurseurs myogéniques (mpc) issues de la division des cellules satellites ne s'engagent pas dans la différenciation terminale mais constituent une réserve des cellules progénitrices sans qu'on puisse, dans ces conditions expérimentales de manipulation de fibres isolées, les nommer cellules satellites [5]. En culture, on observe également cette propriété d'autorenouvellement, puisqu'une sous-population des mpc constitue les « cellules de réserve » (RC) : ces cellules qui ne sont pas engagées dans le cycle cellulaire et sont indifférenciées peuvent donner naissance à des cellu-

les différenciées et à de nouvelles RC dès lors qu'elles sont réensemencées et activées ; elles expriment plusieurs des caractéristiques des cellules satellites<sup>1</sup>. Les cellules mpc en phase de prolifération coexpriment les facteurs de transcription Pax7 et MyoD. Puis elles suivent deux destins : la majorité d'entre elles sort du cycle cellulaire et s'engage dans la différenciation myogénique terminale (que caractérisent la répression de Pax7 et la surexpression de MyoD) pour fusionner en myotubes. Une minorité sort du cycle cellulaire pour entrer en quiescence, réprime l'expression de MyoD et surexprime Pax7, arborant une signature caractéristique des cellules satel-

Le titre de cet article est la traduction du commentaire de S. Duguez et T. Partridge publié dans *Cell Stem Cell* [12] : *Skeletal muscle sings a choral stem cell lullaby*.

<sup>1</sup> On ne peut toutefois pas les qualifier de cellules satellites, cette appellation reposant sur l'existence d'une localisation anatomique précise le long d'une fibre musculaire.