

# Problèmes d'interprétation du diagnostic génétique par le séquençage haut débit

**Anne BARLIER**

**AP-HM** Hôpital de la Conception,  
**Aix-Marseille Université** UMR 7286 CNRS

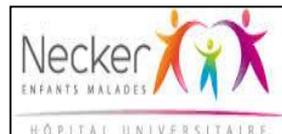
[anne.barlier@univ-amu.fr](mailto:anne.barlier@univ-amu.fr)



**Jean-Paul BONNEFONT**

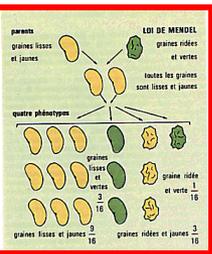
**AP-HP** GH Necker Enfants malades  
**Université Paris Descartes**, IHU IMAGINE UMR 1153

[jean-paul.bonnefont@inserm.fr](mailto:jean-paul.bonnefont@inserm.fr)

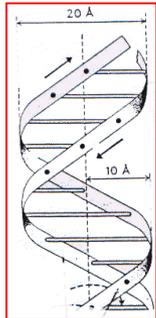


# 150 ans de génétique...

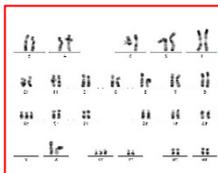
**1865**  
Lois de Mendel



**1953**  
Structure de l'ADN



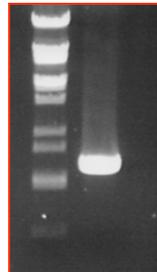
**1959**  
Trisomie 21



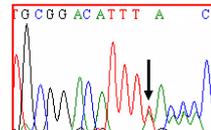
**1975**  
Séquençage « manuel »  
(Sanger)



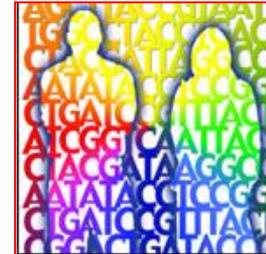
**1985**  
PCR



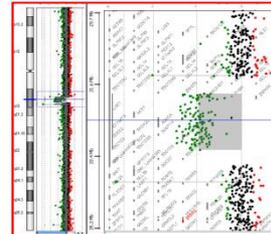
**1993**  
Séquenceurs « automatiques »  
(Sanger)



**2004**  
Séquençage génome humain



**2007**  
Analyses pan-génomiques



**1865 – 1975**  
Ere pré-moléculaire

**1975 – 2004**  
Ere moléculaire

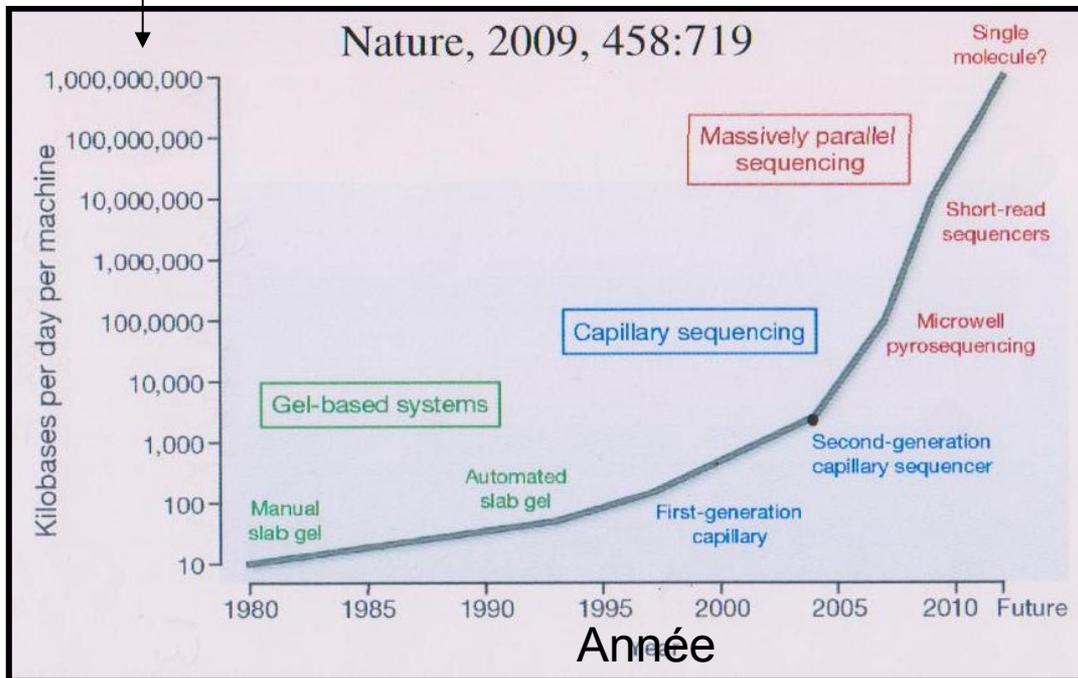
**2004-**  
Ere génomique

# Un séquençage de plus en plus rapide et de moins en moins onéreux...

Génome humain: 3 milliards de nucléotides

Kilobase / jour/  
machine

Séquençage x 10



- 1985 Gel Based Systems :  
**12 ans pour 1000 machines**

100 millions \$

- 2005 Capillary sequencing :  
**1 an pour 120 machines**

1.5 million \$

- 2012 Short Read Sequencers:

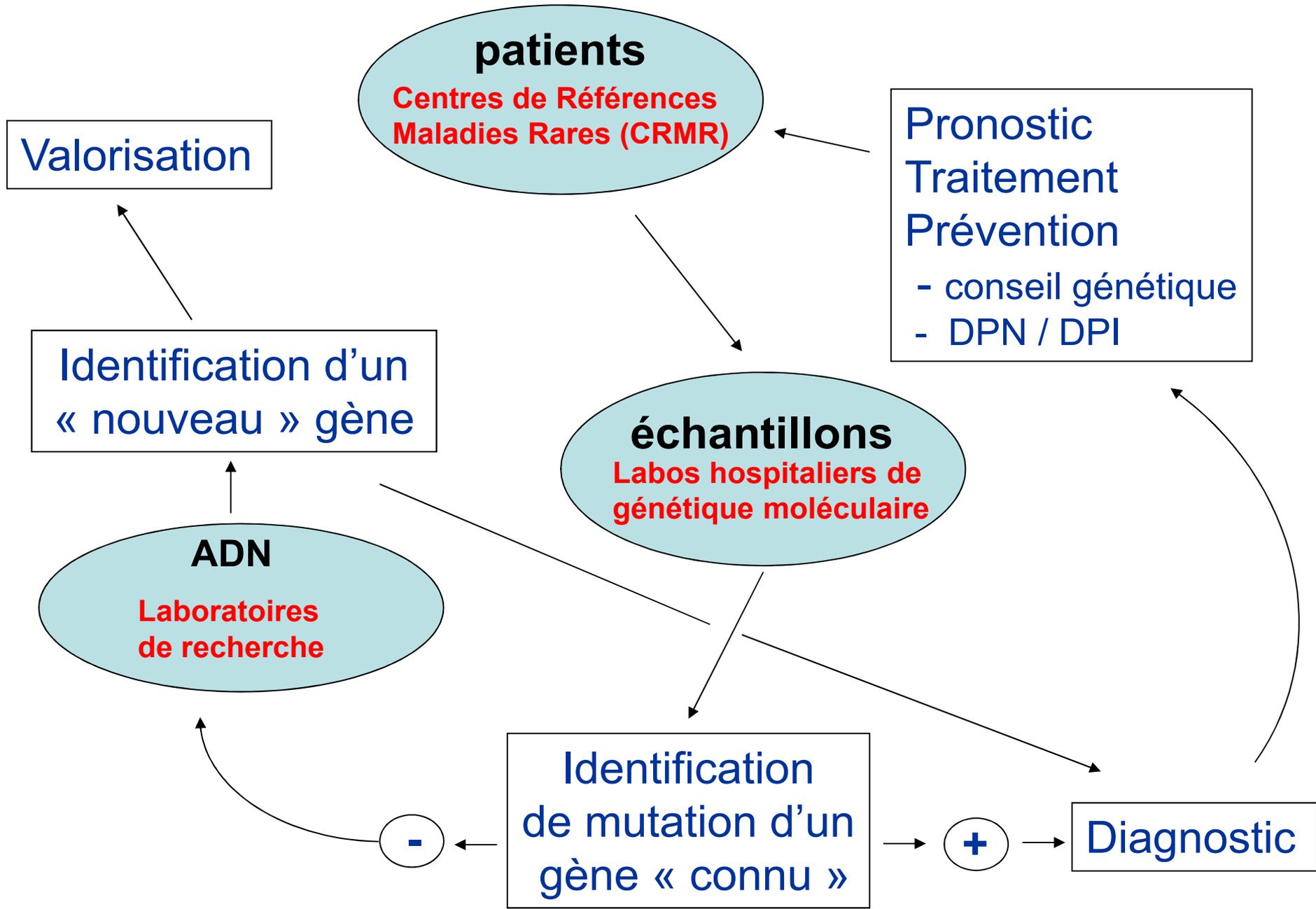
(Hiseq2000)

**1 semaine pour 1 machine**

1000 \$

- **Le séquençage à visée diagnostique  
en 2015**

**« mettre en œuvre la médecine d'aujourd'hui  
et préparer la médecine de demain »**



# • Pourquoi le séquençage haut débit

(« HTS » = High Throughput Sequencing)

(« NGS » = Next Generation Sequencing)

## est-il en train de supplanter les techniques « conventionnelles » de génétique moléculaire dans les laboratoires de diagnostic ?

NEWS IN FOCUS

288 | NATURE | VOL 482 | 16 FEBRUARY 2012

### Sequencing set to alter clinical landscape

*Access to whole genomes shifts potential for diagnosis, but poses challenges for doctors and regulators.*

BY ERIKA CHECK HAYDEN

Sequencing a patient's complete genome can cost as little as tests that target specific genes, and in a handful of cases it has led to a life-changing diagnosis. No wonder the technology is moving fast from bench to bedside. But as the trend accelerates, researchers are left grappling with complex questions about the medical value of patient genomes, and how sequencing in the clinic should be regulated.

Two clinical sequencing programmes that launched in January, one at Baylor College of Medicine in Houston, Texas, and the other at the University of California, Los Angeles (UCLA), suggest that genomics is about to enter the clinic on a large scale. Already, a growing list of companies and hospitals offer sequencing of whole genomes or the exome (the gene-coding part of the genome). Several other efforts are set to begin this year, and many are in the spotlight this week at the Advances in Genome Biology and Technology meeting in Marco Island, Florida.

"We're all amazed by the speed at which this is happening," says Michael Watson, executive director of the American College of Medical Genetics (ACMG) in Bethesda, Maryland.

of UCLA's clinical sequencing programme, says that, so far, about half of the 10 patients who have come through his programme have received a diagnosis. Geneticists caution, however, that these samples are small and highly selective, and that the true rate of



#### PUT TO THE TEST

A growing number of successes have spurred the use of whole-genome sequencing as a diagnostic tool.

METHODS

Human Mutation

OFFICIAL JOURNAL

HGVS

HUMAN GENOME VARIATION SOCIETY

www.hgvs.org

2013

### A Post-Hoc Comparison of the Utility of Sanger Sequencing and Exome Sequencing for the Diagnosis of Heterogeneous Diseases

Kornelia Neveling,<sup>1,2†</sup> Ilse Feenstra,<sup>1†</sup> Christian Gilissen,<sup>1,2,3,4†</sup> Lies H. Hoefsloot,<sup>1,5</sup> Erik-Jan Kamsteeg,<sup>1</sup>

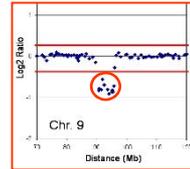
# NGS: vers une technique unique pour la détection de différents types de mutations

## • Grands réarrangements

(délétions/ insertions/duplications > 250 kb)

- **CGH array** → pangénomique

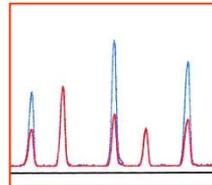
- **FISH** → ciblé



## • Réarrangements intra-géniques

(Délétions/ duplications ≥ 1 exon)

**MLPA, QM-PSF, qPCR...**



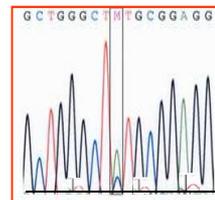
## • Micro-réarrangements

(mut ponctuelles, ins/del quelques pb)

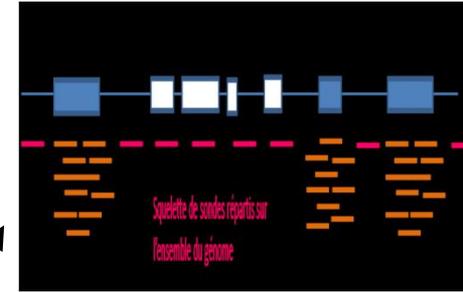
- homogènes

- en mosaïque

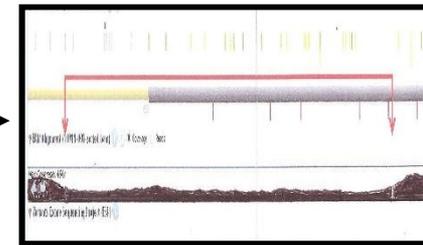
**Sanger**



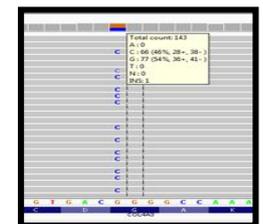
**NGS**



Recherche Del/Dup pangénomique



Délétion Intra-génique



mosaïque

# Un diagnostic plus rapide pour plus de patients...

Phénotypage des patients

gènes « candidats »

⊕

Laboratoire hospitalier

Séquençage de Sanger

≤ 10 patients

gène 1

Si négatif → gène 2

Si négatif → gène 3 ...

Des mois, voire des années de travail

De + en + de - techniciens / biologistes

- séquenceurs / moyens financiers

*La plupart des patients demeuraient sans diagnostic*

**Hier** : transfert des connaissances  
génomique → patients  
= bloqué par un obstacle méthodologique

Phénotypage des patients

Laboratoire hospitalier

gènes « candidats »

⊕

Séquençage ciblé  
« moyen débit »

panels ≤ 30 gènes  
50 patients (1 manip)

1 semaine

⊕ diagnostic ⊖

⊖

Séquençage « haut débit »

panels 30 à 1200 gènes

80 patients

OU

exome entier (20 000 gènes)

(1 manip)

1-2 mois

(bioinformatique)

⊕ diagnostic ⊖ → recherche

*La plupart des patients ont un diagnostic*

**Aujourd'hui** un diagnostic

- dans des délais « raisonnables »

- à un prix « acceptable »

pour de + en + de patients

- Exome vs Panels ciblés:

# un débat tactique au sein des laboratoires de diagnostic moléculaire en 2015

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

## Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders

Yaping Yang, Ph.D., Donna M. Muzny, M.Sc., Jeffrey G. Reid, Ph.D., Matthew N. Bainbridge, Ph.D., Alecia Willis, Ph.D., Patricia A. Ward, M.S., Alicia Braxton, M.S., Joke Beuten, Ph.D., Fan Xia, Ph.D., Zhiyv Niu, Ph.D., Matthew Hardison, Ph.D., Richard Person, Ph.D., Mir Reza Bekheirnia, M.D., Magalie S. Leduc, Ph.D., Amelia Kirby, M.D., Peter Pham, M.Sc., Jennifer Scull, Ph.D., Min Wang, Ph.D., Yan Ding, M.D., Sharon E. Plon, M.D., Ph.D., James R. Lupski, M.D., Ph.D., Arthur L. Beaudet, M.D., Richard A. Gibbs, Ph.D., and Christine M. Eng, M.D.

**Octobre 2013**

Saudi Mendeliome Group *Genome Biology* (2015) 16:134  
DOI 10.1186/s13059-015-0693-2



**RESEARCH** **Open Access**



### Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases

Saudi Mendeliome Group

**Février 2015**

# Un débat tactique au sein des laboratoires de diagnostic moléculaire en 2015

## • Exome pangénomique (les 25 000 gènes du génome)

- avantages:
  - pas d'étape de confection de « librairies » (gain de temps)
  - possibilité d'identification de nouveaux gènes de la pathologie
  - possibilité de ré-analyse bioinfo « élargie » ultérieure si nécessaire
- inconvénients:
  - Equipement + manips plus onéreux
  - analyse bioinfo plus lourdes → délai de résultats plus long
  - couverture encore insuffisante de certaines régions d'intérêt
  - risque de découverte fortuite de variants dans des gènes hors champ du diagnostic (« Incidental findings ») (pbme éthique)
  - comment stocker un gros volume de données bioinfo
  - confusion entre diagnostic (gènes « connus ») et recherche (nouveaux gènes)

## • Panels ciblés (1 à > 1000 gènes)

- avantages:
  - réponse à une question « clinique » (« diabetes », « retards staturaux »...)
  - équipement et manips moins onéreux
  - qualité de la séquence (couverture et profondeur)
  - plus simple techniquement

## - Inconvénients      nécessité

- de confection de librairies ➔ risque de contamination / d'erreur d'échantillons
- d'actualiser les librairies à mesure de l'identification de nouveaux gènes (validations...)

# • Problèmes d'interprétation du diagnostic génétique par NGS

© American College of Medical Genetics and Genomics

**ACMG PRACTICE GUIDELINES** | **Genetics  
in Medicine**

25 July 2013

**ACMG clinical laboratory standards for next-generation  
sequencing**

Heidi L. Rehm, PhD<sup>1,2</sup>, Sherri J. Bale, PhD<sup>3</sup>, Pinar Bayrak-Toydemir, MD, PhD<sup>4</sup>, Jonathan S. Berg, MD<sup>5</sup>,  
Kerry K. Brown, PhD<sup>6</sup>, Joshua L. Deignan, PhD<sup>7</sup>, Michael J. Friez, PhD<sup>8</sup>, Birgit H. Funke, PhD<sup>1,2</sup>,  
Madhuri R. Hegde, PhD<sup>9</sup> and Elaine Lyon, PhD<sup>4</sup>; for the Working Group of the American College of  
Medical Genetics and Genomics Laboratory Quality Assurance Committee

**EuroGentest**

**Guidelines for diagnostic  
next generation sequencing**

17 October 2014



**Reward bioinformaticians**

Biological data will continue to pile up unless those who analyse them are recognized as creative collaborators in need of career paths, says **Jeffrey Chang**

9 APRIL 2015 | VOL 520 | NATURE | 151

# Des impératifs de qualité de la séquence

↙ Régions « NON COUVERTES » ↘

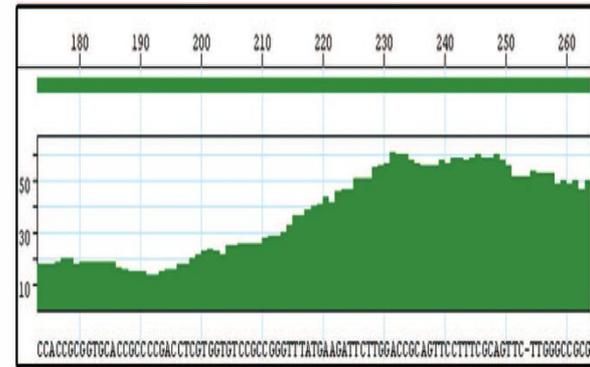
Séquence de référence

```
1929 1938 1948 1958 1968 1978 1988 1998
GCCAAAGAGCCGTTTAAATCTCGGGCGCTTAATTCGCCTCGTGAAAGAATATCATCTGCTGAACCCGGTCATTGTTGAC
GCCAAAGAGCCGTTTAAATCTCGGGCGCTTAATTCGCCTCGTGAAAGAATATCATCTGCTGAACCCGGTCATTGTTGAC
GCCGTTTAAATCTCGGGCGCTTAATTCGCCTCGTGAA
GCCGTTTAAATCTCGGGCGCTTAATTCGCCTCGTGAA
CCGTTTAAATCTCGGGCGCTTAATTCGCCTCGTGAA
CCGTTTAAATCTCGGGCGCTTAATTCGCCTCGTGAA
GTTTAAATCTCGGGCGCTTAATTCGCCTCGTGAAAG
GTTTAAATCTCGGGCGCTTAATTCGCCTCGTGAAAG
TTAATCTCGGGCGCTTAATTCGCCTCGTGAAAGAA
TTAATCTCGGGCGCTTAATTCGCCTCGTGAAAGAA
TTAATCTCGGGCGCTTAATTCGCCTCGTGAAAGAA
TAATCTCGGGCGCTTAATTCGCCTCGTGAAAGAA
TAATCTCGGGCGCTTAATTCGCCTCGTGAAAGAA
AATCTCGGGCGCTTAATTCGCCTCGTGAAAGAA
ATCTCGGGCGCTTAATTCGCCTCGTGAAAGAAAT
CTCGGGCGCTTAATTCGCCTCGTGAAAGAAATATCA
CGGGCGCTTAATTCGCCTCGTGAAAGAAATATCATC
CGGGCGCTTAATTCGCCTCGTGAAAGAAATATCATC
GGCGCTTAATTCGCCTCGTGAAAGAAATATCATCTG
GGCGCTTAATTCGCCTCGTGAAAGAAATATCATCTG
GCGCTTAATTCGCCTCGTGAAAGAAATATCATCTG
GCGCTTAATTCGCCTCGTGAAAGAAATATCATCTG
CGCTTAATTCGCCTCGTGAAAGAAATATCATCTGCT
CGCTTAATTCGCCTCGTGAAAGAAATATCATCTGCT
CGCTTAATTCGCCTCGTGAAAGAAATATCATCTGCT
```

1 read  
= 1 fragment lu →  
= 50 à 200pb

Couverture  
(taille de la région  
réellement séquencée)

Profondeur  
(moyenne du  
nombre de  
lectures pour  
chaque base ;  
exprimée en X)



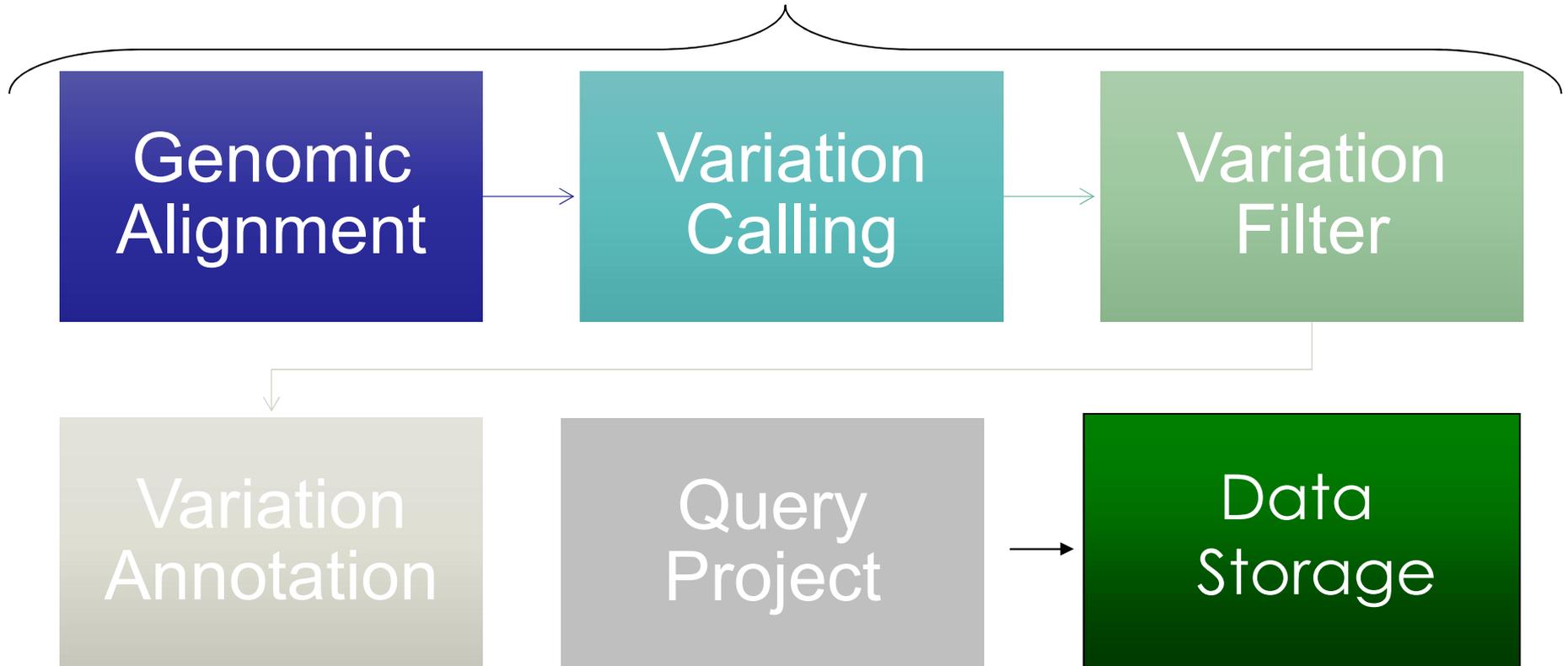
Etude de la qualité de chaque base  
(400 10<sup>6</sup> erreurs de séquences  
potentielles / run de 600 Gb)

## Qualité de séquence insuffisante

- risques:
  - méconnaître des variants (« faux-négatifs »)
  - filtre insuffisant des artefacts (« faux positifs »)
  - assignement incorrect de la « zygotie » (homo, hétéro, mos)
- exigences:
  - couverture 30x sur 90-95 % de la séquence totale
  - couverture moyenne 100x pour le patient
  - couverture ≥ 50x des régions des variants candidats

# Des impératifs de qualité de l'analyse informatique

« Pipeline »





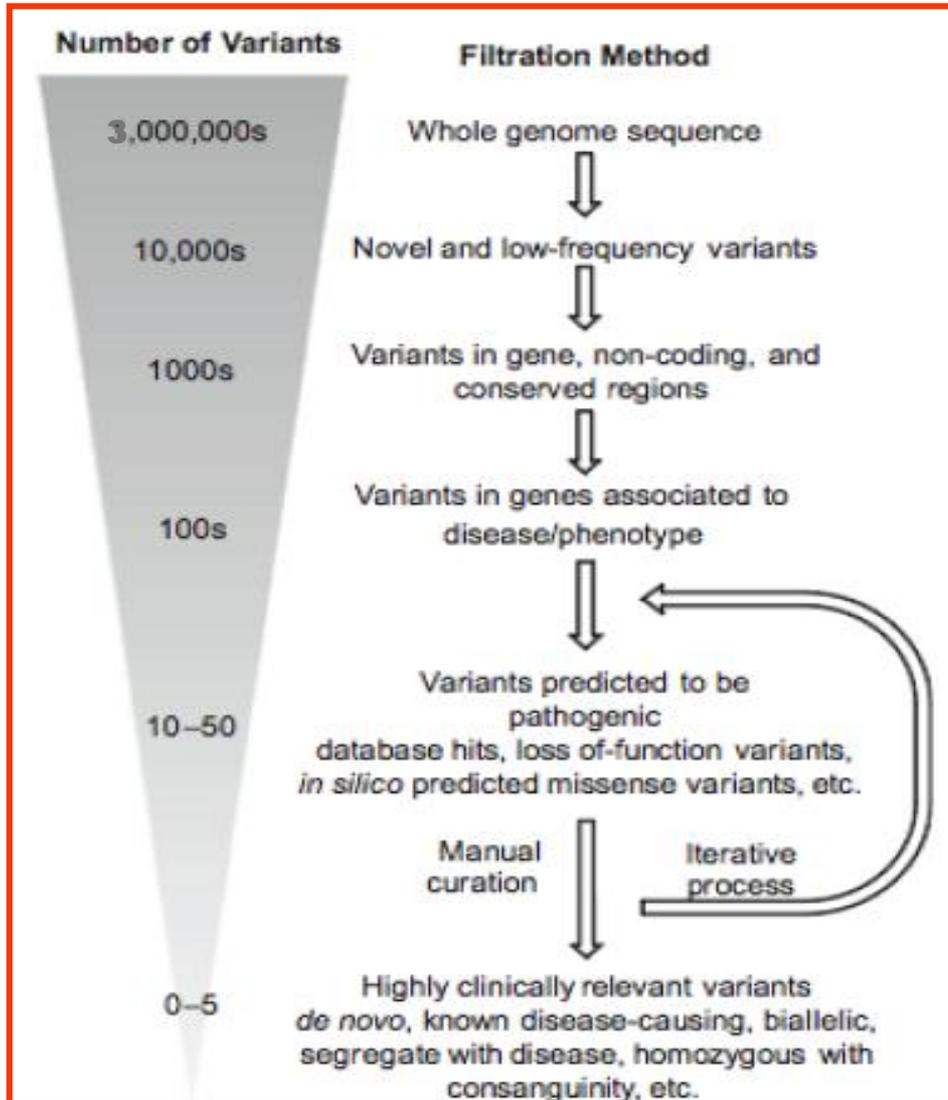
# Pourquoi un filtre bioinformatique est-il nécessaire ?

Nombre approximatif de variants de séquence mononucléotidique (« SNV »)

- Génome: 3,5. 000 000
- Exome : 15 000
- Seq exons de 80 gènes : 150

# NGS : filtre bio informatique des variants

Le NGS détecte un **grand nombre de variants**, d'autant plus élevé que le nombre de gènes testés est important, qu'il faudra « filtrer » pour sélectionner le ou les variants candidats (« mutations »)



## Les risques

critères de filtre

1/ trop stringents  
risque de faux-négatifs

2/ trop permissifs  
trop de variants à évaluer  
perte de temps

un équilibre à trouver...

+++

- expertise clinico-biologique
- concertation biologiste-bioinformaticien

Annotation des variants:  
sélectionner parmi les variants candidats  
le ou les variants responsables de la maladie

1/ le variant a t-il un impact sur la fonction de la protéine ?  
(« pathogénicité »)

2/ Si la réponse à 1/ est  
« OUI » ou « PROBABLEMENT » ou « POSSIBLEMENT »



quels sont les arguments en faveur de  
sa responsabilité dans le phénotype de l'individu testé ?

# 1/ Quels arguments en faveur ou défaveur du caractère délétère du variant (« damaging » or « deleterious » variant) ?

## # Le variant est « à l'évidence » ou « probablement » non délétère

Variant ponctuel (SNV) fréquent (> 1%?) dans la pop. générale → bases de données

(1000 génomes, dbSNP, ExAC..)

## # Le variant est « à l'évidence » délétère

### Données du séquençage

- « grand » réarrangement (affectant  $\geq 1$  exon du gène)
- micro réarrangement intra génique (ins/del) décalant le cadre de lecture de l'ARNm
- mutation ponctuelle tronquant la protéine (« non-sens », « épissage »...) **MAIS...**

## # Le variant est « possiblement » ou « probablement » délétère

(mutation ponctuelle exonique prédite « faux-sens », mutation intronique...)

### Logiciels de prédiction d'impact du variant sur

- l'ARNm  
(*HSF*)



- création/suppression d'un site consensus d'épissage,
- création d'un site « cryptique » d'épissage (activation d'un « enhancer » ou inactivation d'un « silencer » d'épissage)

- La protéine  
(*Sift, Polyphen, Swissprot..*)

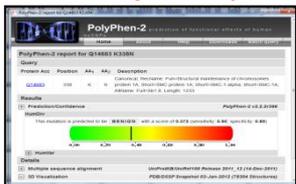
- disparition, modification de la structure tridimensionnelle
- modification de la fonction (conservation inter espèces de l'ac. aminé site actif, interaction protéine/protéine)

Alamut

### Etudes de « validation fonctionnelle »

Démonstration expérimentale de l'impact du variant sur l'ARNm et/ou la protéine

*Etudes in vitro (cellules patients, modèles cellulaires), modèles animaux...*

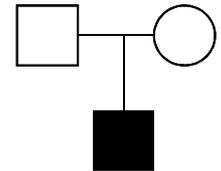


# Les logiciels de prédiction: des résultats parfois discordants... imposant le recours aux études de validation fonctionnelle

NM\_001093.3:c.952G>A  
Chr12(GRCh37):g.109609636G>A  
p.Val318Ile

enfant et père HETEROZYGOTES

**Gène ACACB**  
Protéine mitochondriale



Nucléotide très conservé (phyloP: 6.02 [-14.1;6.4])  
Acide aminé très conservé jusqu'à Levure de boulanger (sur 11 espèces)  
Ecart physico-chimique entre Val et Ile peu important (Dist. de Grantham: 29 [0-215])  
Variation située dans les domaines protéiques :

Carbamoyl-phosphate synthase, large subunit, N-terminal  
Biotin carboxylation domain

Align GVGD: C0 (GV: 65.28 - GD: 28.68)

SIFT: **Tolerated** (score: 0.08, median: 3.33)

MutationTaster: **disease causing** (p-value: 1)

Polyphen **possibly damaging** with a score of 0.647 (sensitivity: 0.80; specificity: 0.84)

Homme  
Chimpanzé  
Rat  
Souris  
Chien  
Poulet  
Xénope  
Tétraodon  
Mouche drosophile  
C. elegans

E	Y	I	K	M	A	D	H	Y	V	P	V	P	G	G	P	N	N	N	N
E	Y	I	K	M	A	D	H	Y	V	P	V	P	G	G	P	N	N	N	N
E	Y	I	K	M	A	D	Q	Y	V	P	V	P	G	G	P	N	N	N	N
E	Y	I	K	M	A	D	Q	Y	V	P	V	P	G	G	P	N	N	N	N
E	Y	I	R	M	A	D	H	Y	V	P	V	P	G	G	A	N	N	N	N
E	Y	I	K	M	A	D	H	Y	V	P	V	P	G	G	P	N	N	N	N
E	Y	I	K	M	A	D	H	Y	V	P	V	P	G	G	P	N	N	N	N
E	Y	I	K	M	A	D	H	Y	V	P	V	P	G	G	S	N	N	N	N
E	Y	I	K	M	A	D	H	Y	V	P	V	P	G	G	S	N	N	N	N
E	S	L	R	M	P	N	T	L	A	E	S	P	S	G	T	N	K	N	N

NM\_001093.3:c.3893G>A  
Chr12(GRCh37):g.109665186G>A  
p.Arg1298Gln

enfants et mère HETEROZYGOTE

Homme  
Chimpanzé  
Rat  
Souris  
Chien  
Poulet  
Xénope  
Tétraodon  
Mouche drosophile  
C. elegans  
Levure de boulanger

V	Y	V	R	R	G	Y	I	A	Y	E	L	N	S
V	Y	V	R	R	G	Y	I	A	Y	E	L	N	S
V	Y	V	R	R	G	Y	I	A	Y	E	L	N	S
V	Y	V	R	R	G	Y	I	A	Y	E	L	N	S
V	Y	V	R	R	G	Y	I	A	Y	E	L	N	S
V	Y	V	R	R	A	Y	I	A	Y	E	L	N	S
V	Y	V	R	R	A	Y	I	A	Y	E	L	N	S
R	Y	V	A	M	A	H	I	S	A	E	E	G	S
I	Y	A	R	Y	A	Y	G	N	Y	Q	L	K	S

Nucléotide très conservé (phyloP: 5.94 [-14.1;6.4])  
Acide aminé très conservé jusqu'à Levure de boulanger (sur 11 espèces)  
Ecart physico-chimique entre Arg et Gln peu important (Dist. de Grantham: 43 [0-215])  
Cette variation est située dans le domaine protéique: Acetyl-CoA carboxylase, central domain

Align GVGD: C0 (GV: 111.26 - GD: 16.48)

SIFT: **Tolerated** (score: 0.07, median: 3.33)

MutationTaster: **disease causing** (p-value: 1)

Polyphen: **probably damaging** with a score of 0.999 (sensitivity: 0.09; specificity: 0.99)

# INTERPRETATION DES VARIANTES DANS LES DEFICITS HYPOPHYSAIRES COMBINES MULTIPLES



Service d'Endocrinologie  
Conception

Dr C Rochette  
Pr T Brue  
Dr F Castinetti

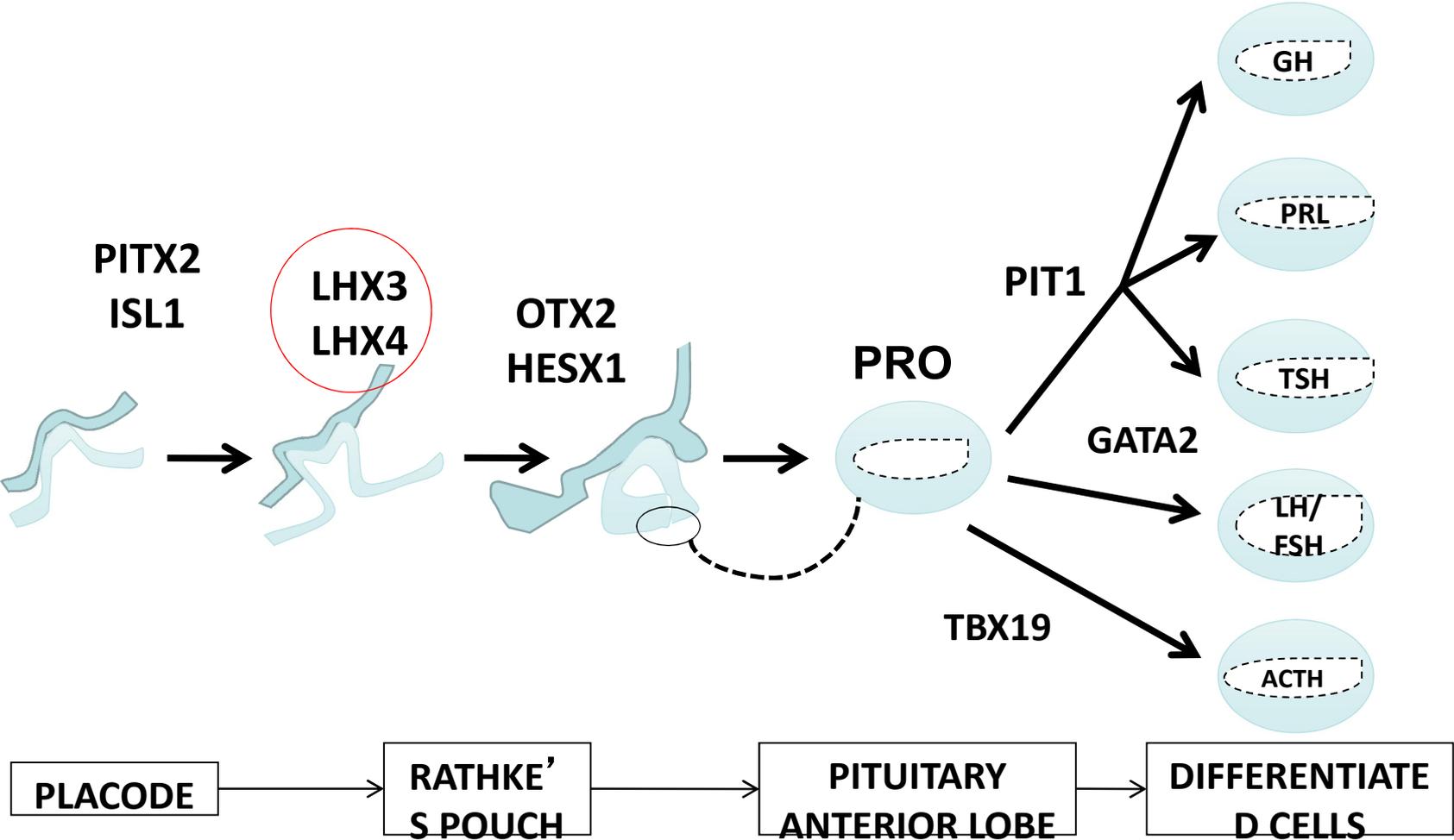
Biologie Moléculaire  
Conception

Pr A Barlier,  
Dr A Saveanu

# L'EXEMPLE DES VARIANTS DE LHX3 ET LHX4

## Phénotype classique:

- LHX4: déficit hypophysaire multiple, section de tige. Variants à l'état hétérozygote.
- LHX3: déficit hypophysaire multiple, section de tige, surdité. Variants à l'état homozygote.



## EXEMPLE 1: LE VARIANT T90M DE LHX4

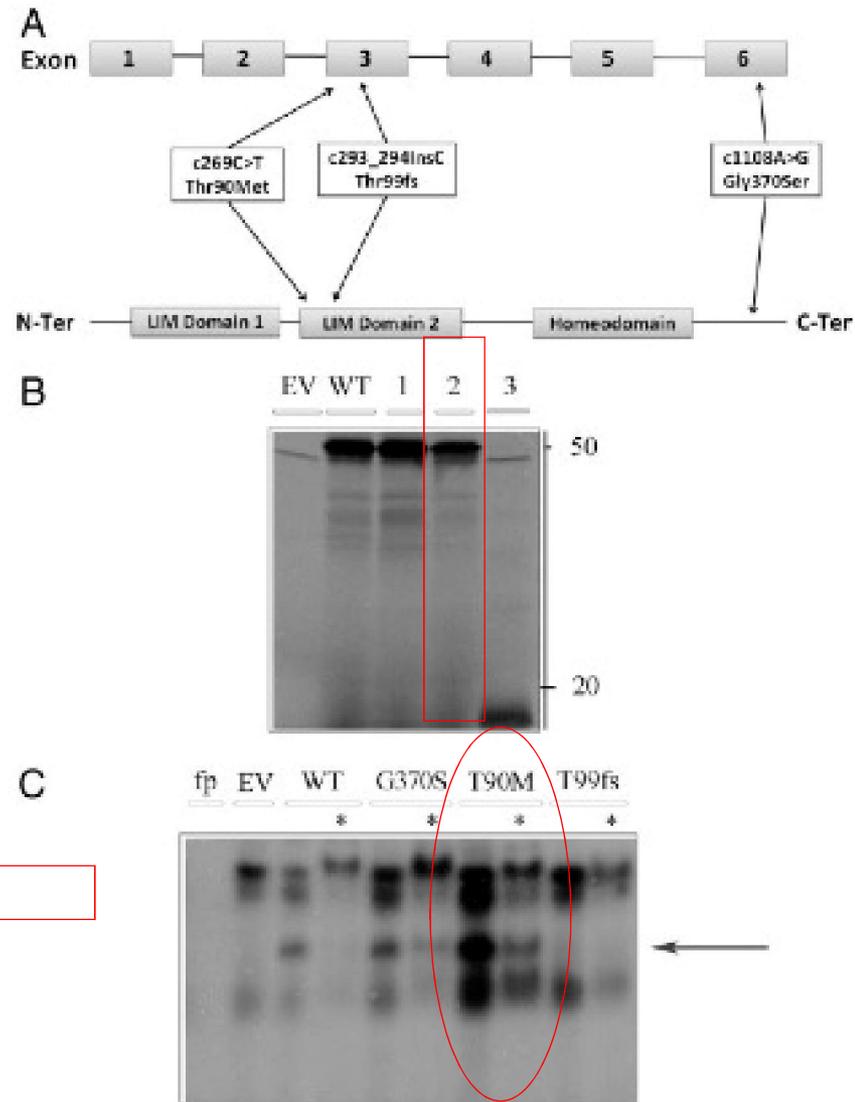
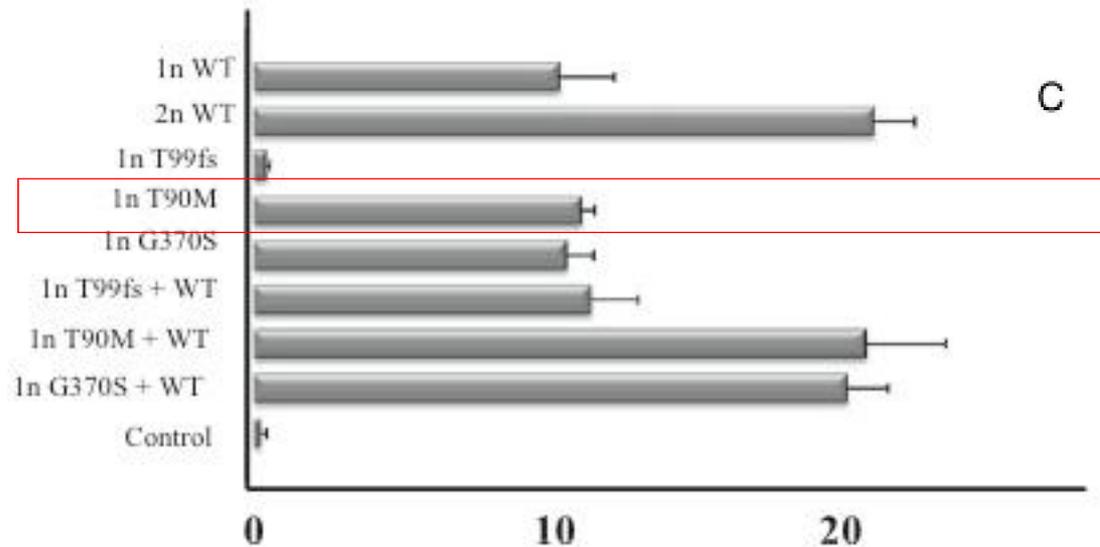
Variant T90M: 2<sup>e</sup> domaine LIM. Hétérozygote.

Phénotype compatible: déficit hypophysaire multiple néonatal, section de tige

In silico: probablement pathogène

EXAC: <1/10.000

Expression protéique conservée, liaison ADN conservée, absence de modification en analyse fonctionnelle (transactivation de différents promoteurs cibles).



## EXEMPLE 2: LE VARIANT Q346R DE LHX4

Variant Q346R: pas dans un domaine fonctionnel.  
Hétérozygote.

Phénotype compatible: déficit hypophysaire multiple néonatal, section de tige **RETROUVE CHEZ 2 PATIENTS AVEC LE MEME PHENOTYPE, NON APPARENTES, D'ORIGINE GEOGRAPHIQUE DIFFERENTE**

In silico: non pathogène

EXAC: <1/100.000

Expression protéique conservée, liaison ADN conservée, absence de modification en analyse fonctionnelle (transactivation de différents promoteurs cibles).

Rochette et al., 2015 PLOSone

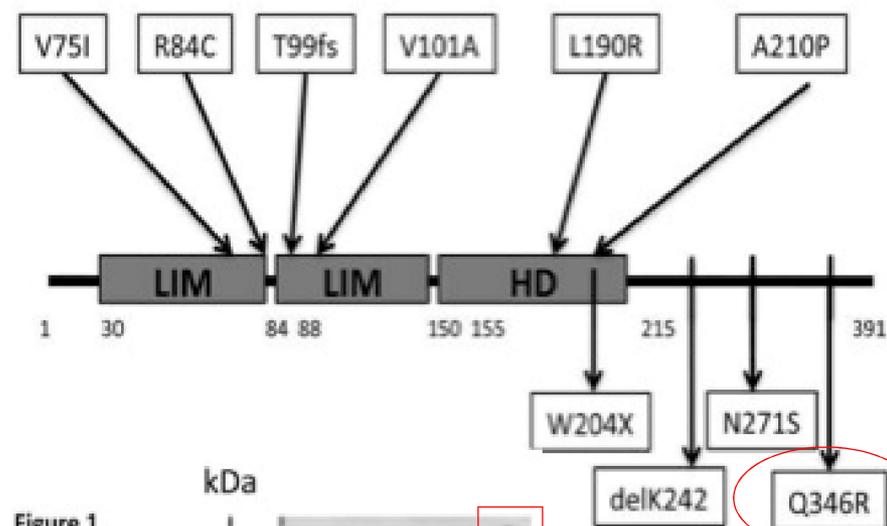
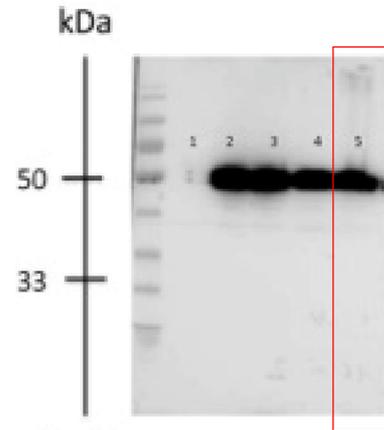
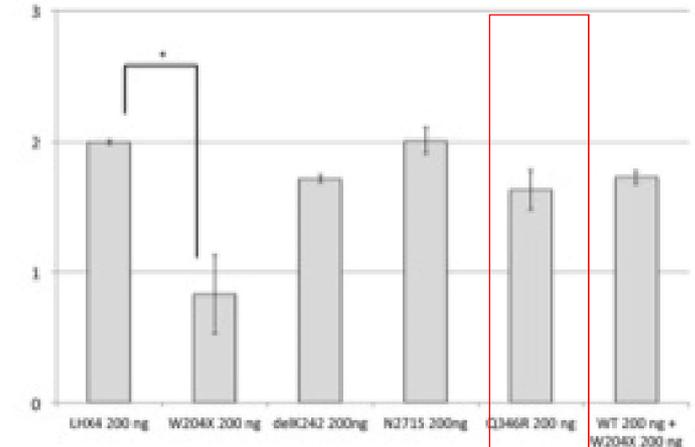


Figure 1



Relative luciferase units  
(fold versus EV)



### EXEMPLE 3: LE VARIANT R208G DE LHX3

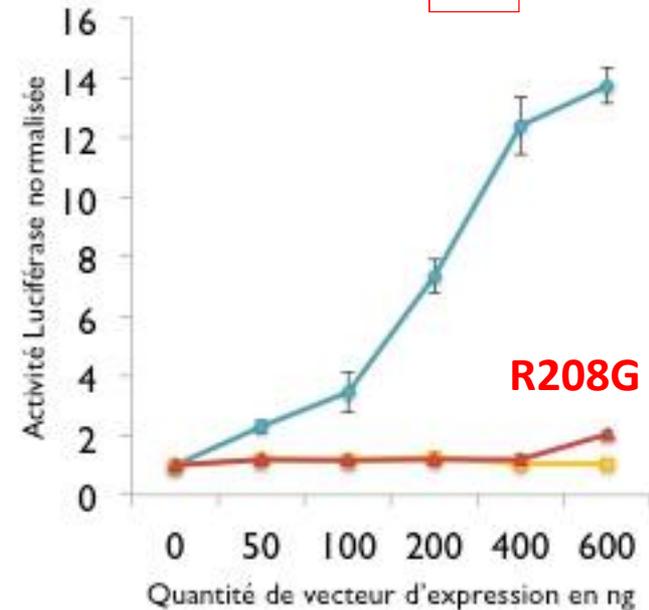
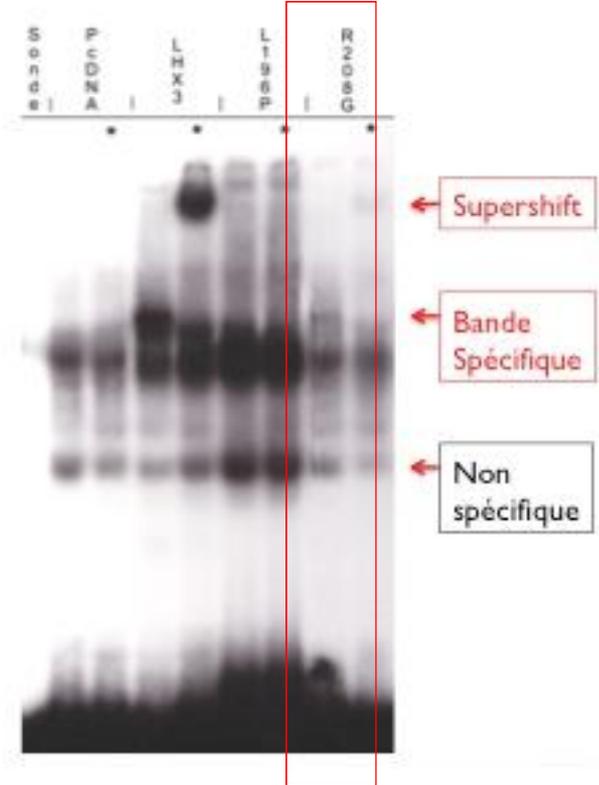
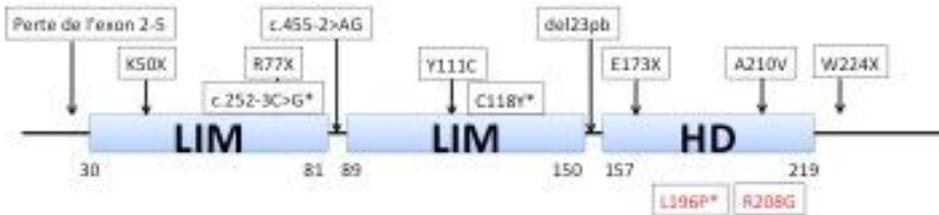
Variant R208G: homéodomaine. Homozygote.

Phénotype compatible: déficit hypophysaire multiple néonatal, agénésie du corps calleux.  
**2 parents hétérozygotes sains.**

In silico: non pathogène

EXAC: Non décrit

Expression protéique conservée, liaison ADN conservée mais absence d'activation des promoteurs cibles in vitro.



## EXEMPLE 4: LE VARIANT L196P DE LHX3

Variant L196P: homéodomaine

Phénotype compatible: déficit hypophysaire multiple néonatal, surdité (retrouvée dans >50% des patients porteurs d'anomalie de LHX3)

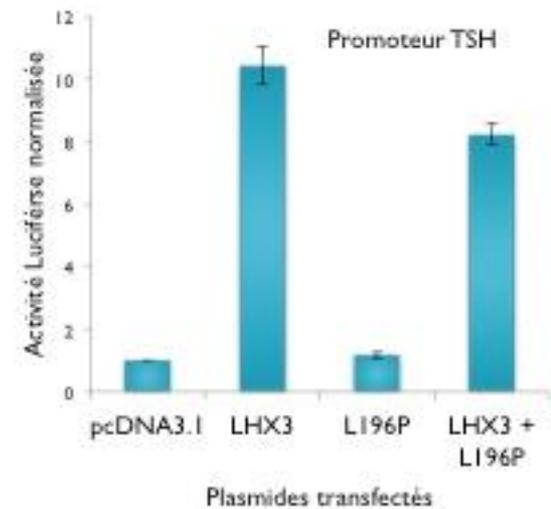
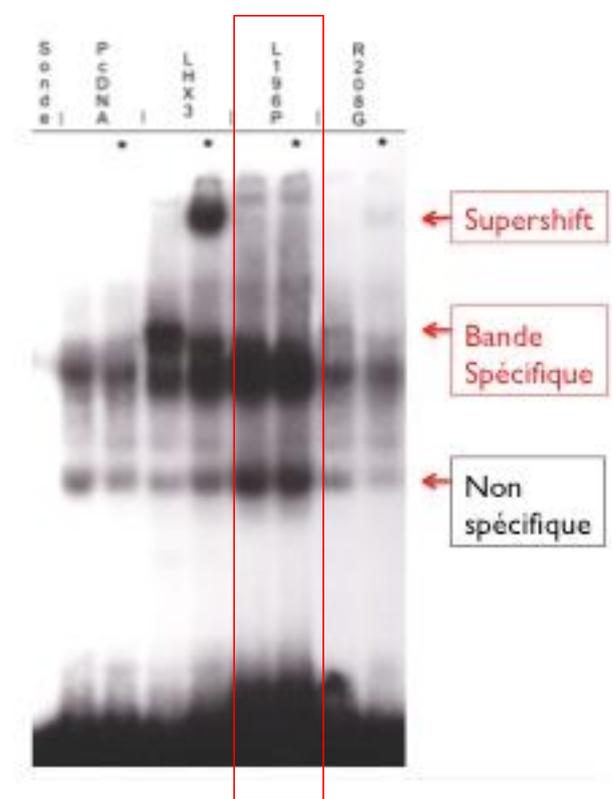
In silico: pathogène

EXAC: non décrit

Expression protéique conservée, liaison ADN NON conservée, absence d'activation des promoteurs cibles. Pas d'effet dominant négatif évident (haploinsuffisance ?)

**Problème:** TOUS LES VARIANTS PATHOGENES DE LHX3 SONT HOMOZYGOTES.

**Or Patient porteur d'un variant à l'état HETEROZYGOTE.**



## 2/ Quels arguments en faveur de la responsabilité d'un variant, dont la pathogénicité est avérée / probable / possible, vis-à-vis du phénotype de l'individu testé ?

- Arguments « inter-familiaux »

- 1/ gène connu comme associé à une maladie génétique (panels, exome entier)

- le variant identifié a été antérieurement rapporté (> 1 fois) par des équipes indépendantes et « sérieuses », en association avec le phénotype
- un variant différent mais ayant le même impact sur la séquence primaire de la protéine p. ex: c.34G>C (p.Val12Leu) et c.34G>T (p.Val12Leu)

a été antérieurement rapporté (> 1 fois) par des équipes indépendantes et « sérieuses » en association avec le phénotype

Bases de donnée: « ClinVar », « HGMD », « OMIM », « LOVD »...

- 2/ gène non connu comme associé au phénotype (exome entier)

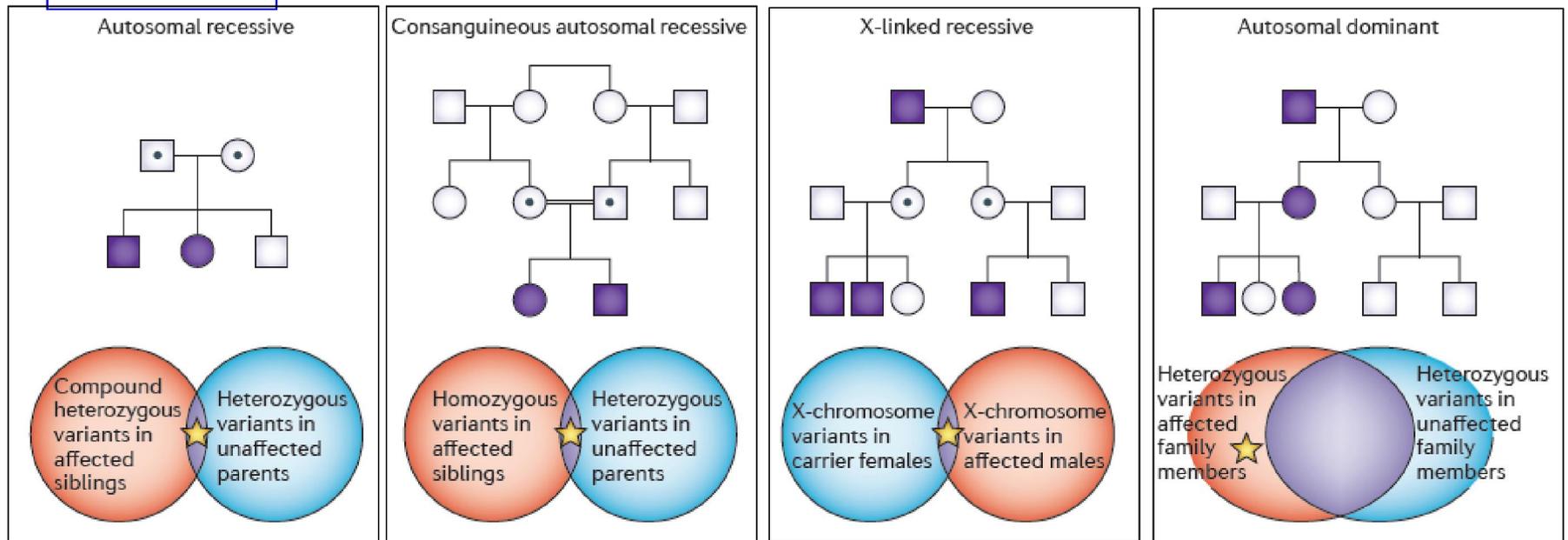
- le variant identifié affecte une protéine dont la distribution tissulaire et la fonction paraissent « compatibles » avec le phénotype observé (canal potassium exprimé dans le pancréas dans un hyperinsulinisme...)

- Argument « intra-familial »

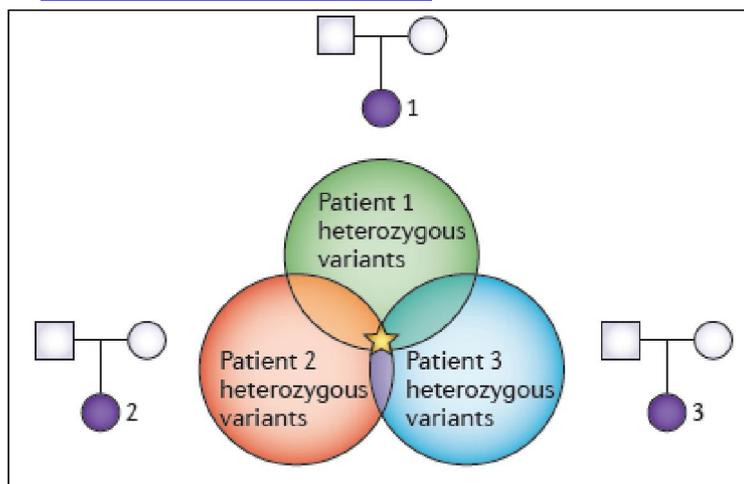
La ségrégation intra-familiale du variant est en accord avec le mode de transmission avéré ou présumé de la maladie

# Une aide à la sélection des variants candidats: l'étude simultanée de plusieurs apparentés MAIS le coût augmente en proportion !

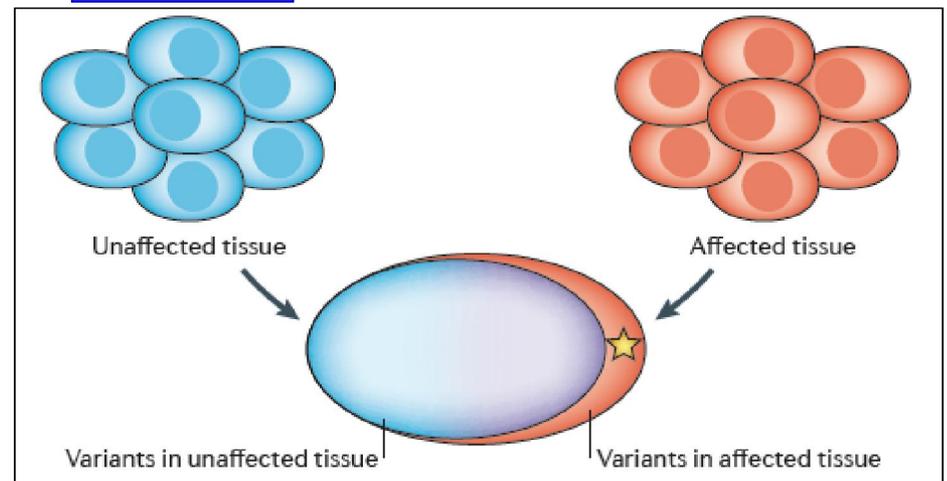
## Inherited mutations



## De novo dominant mutations



## Mosaic mutations



MAIS 50 - 100 mut de novo/ génome dans pop générale !

Au terme de cette confrontation clinique, biologique et bioinformatique, la responsabilité du ou des variants identifié(s) demeure encore parfois incertaine

**PERSPECTIVE** OPEN  
doi:10.1038/nature13127

24 APRIL 2014 | VOL 508 | NATURE | 469

## Guidelines for investigating causality of sequence variants in human disease

D. G. MacArthur<sup>1,2</sup>, T. A. Manolio<sup>3</sup>, D. P. Dimmock<sup>4</sup>, H. L. Rehm<sup>5,6</sup>, J. Shendure<sup>7</sup>, G. R. Abecasis<sup>8</sup>, D. R. Adams<sup>9,10</sup>, R. B. Altman<sup>11</sup>, S. E. Antonarakis<sup>12,13</sup>, E. A. Ashley<sup>14</sup>, J. C. Barrett<sup>15</sup>, L. G. Biesecker<sup>16</sup>, D. F. Conrad<sup>17</sup>, G. M. Cooper<sup>18</sup>, N. J. Cox<sup>19</sup>, M. J. Daly<sup>1,2</sup>, M. B. Gerstein<sup>20,21</sup>, D. B. Goldstein<sup>22</sup>, J. N. Hirschhorn<sup>2,23</sup>, S. M. Leal<sup>24</sup>, L. A. Pennacchio<sup>25,26</sup>, J. A. Stamatoyannopoulos<sup>27</sup>, S. R. Sunyaev<sup>28,29</sup>, D. Valle<sup>30</sup>, B. F. Voight<sup>31</sup>, W. Winckler<sup>2†</sup> & C. Gunter<sup>18†</sup>

© American College of Medical Genetics and Genomics **ACMG STANDARDS AND GUIDELINES** **Genetics in Medicine**

*Genet Med* 2015; 17: 405-24

### Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology

Sue Richards, PhD<sup>1</sup>, Nazneen Aziz, PhD<sup>2,16</sup>, Sherri Bale, PhD<sup>3</sup>, David Bick, MD<sup>4</sup>, Soma Das, PhD<sup>5</sup>, Julie Gastier-Foster, PhD<sup>6,7,8</sup>, Wayne W. Grody, MD, PhD<sup>9,10,11</sup>, Madhuri Hegde, PhD<sup>12</sup>, Elaine Lyon, PhD<sup>13</sup>, Elaine Spector, PhD<sup>14</sup>, Karl Voelkerding, MD<sup>13</sup> and Heidi L. Rehm, PhD<sup>15</sup>; on behalf of the ACMG Laboratory Quality Assurance Committee

**Table 5** Rules for combining criteria to classify sequence variants

<b>Pathogenic</b>	<ul style="list-style-type: none"> <li>(i) 1 Very strong (PVS1) AND               <ul style="list-style-type: none"> <li>(a) <math>\geq 1</math> Strong (PS1–PS4) OR</li> <li>(b) <math>\geq 2</math> Moderate (PM1–PM6) OR</li> <li>(c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) OR</li> <li>(d) <math>\geq 2</math> Supporting (PP1–PP5)</li> </ul> </li> <li>(ii) <math>\geq 2</math> Strong (PS1–PS4) OR</li> <li>(iii) 1 Strong (PS1–PS4) AND               <ul style="list-style-type: none"> <li>(a) <math>\geq 3</math> Moderate (PM1–PM6) OR</li> <li>(b) 2 Moderate (PM1–PM6) AND <math>\geq 2</math> Supporting (PP1–PP5) OR</li> <li>(c) 1 Moderate (PM1–PM6) AND <math>\geq 4</math> supporting (PP1–PP5)</li> </ul> </li> </ul>
<b>Likely pathogenic</b>	<ul style="list-style-type: none"> <li>(i) 1 Very strong (PVS1) AND 1 moderate (PM1–PM6) OR</li> <li>(ii) 1 Strong (PS1–PS4) AND 1–2 moderate (PM1–PM6) OR</li> <li>(iii) 1 Strong (PS1–PS4) AND <math>\geq 2</math> supporting (PP1–PP5) OR</li> <li>(iv) <math>\geq 3</math> Moderate (PM1–PM6) OR</li> <li>(v) 2 Moderate (PM1–PM6) AND <math>\geq 2</math> supporting (PP1–PP5) OR</li> <li>(vi) 1 Moderate (PM1–PM6) AND <math>\geq 4</math> supporting (PP1–PP5)</li> </ul>
<b>Benign</b>	<ul style="list-style-type: none"> <li>(i) 1 Stand-alone (BA1) OR</li> <li>(ii) <math>\geq 2</math> Strong (BS1–BS4)</li> </ul>
<b>Likely benign</b>	<ul style="list-style-type: none"> <li>(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) OR</li> <li>(ii) <math>\geq 2</math> Supporting (BP1–BP7)</li> </ul>
<b>Uncertain significance</b>	<ul style="list-style-type: none"> <li>(i) Other criteria shown above are not met OR</li> <li>(ii) the criteria for benign and pathogenic are contradictory</li> </ul>

Variants of Unknown Significance (« VUS »)

# **L'union fait la force: projet national d'homogénéisation de l'interprétation de variants de séquence**

- ✓ **Bases de données de variants partagées ?**
- ✓ **Variants rares difficiles à interpréter** : groupes de discussion (*mailing lists*)
- ✓ **Variants pour lesquels les laboratoires experts n'ont pu aboutir à une réponse consensuelle** : réunions de confrontation clinico-biologiques :
  - Acteurs du diagnostic + équipes de recherche
  - Réunions physiques ou visioconférences
  - ( GEM NGS ? )
- ✓ **Contrôles qualité:**
  - ✓ Echanges de fichiers de variants pour interprétation en aveugle
  - ✓ Echanges d'ADN entre laboratoires



# Comment faire face à la complexité de l'interprétation des données du NGS ?

- Des centres de génétique experts pour une thématique donnée
- Confrontation clinicien/biologiste/bioinformaticien/chercheur +++
- Des compte-rendus d'analyse intelligibles !
- **résultat** « positif » ou « négatif » ou « non concluant »
- **outils** utilisés pour 1- la production de séquence  
(séquenceur, confection des librairies, CQI, vérification par 2<sup>ème</sup> méthode...)
- 2- l'analyse bioinformatique  
(logiciels utilisés, type d'interface bioinformaticien-biologiste...)
- liste exhaustive des **gènes testés** (panels)
- liste des **régions des gènes insuffisamment couvertes**
- variants décrits selon la nomenclature internationale
- politique du laboratoire en matière de découverte fortuite de **variants de prédisposition** à une maladie grave sans rapport avec l'affection ayant motivé la prescription du test (exome)
- nécessité ou non de vérification sur un 2<sup>ème</sup> prélèvement (identité)

