

## Com es pot mesurar La variabilitat humana?

Estudiar la variabilitat de la nostra espècie ha estat, i és, un dels pilars de la nostra recerca bàsica i aplicada, ja sigui associada a l'estudi de l'evolució de la nostra espècie o a la comprensió que només un catàleg complet de la variabilitat genètica humana ens permetrà associar les variants genètiques a malalties simples i complexes per a les quals encara no tenim una base genètica.

S'han anat succeint diverses etapes històriques per recuperar el catàleg més ampli de les variants genètiques existents en les poblacions humanes. Des dels estudis inicials en grups sanguinis fins a la seqüenciació actual de genomes complets, hem anat passant per visions més o menys parcials de la nostra espècie.

Actualment, hi ha dues opinions generalitzades i una mica contraposades. D'una banda, tenim projectes com el dels 1.000 Genomes (<http://www.1000Genomes.org>) que intenten obtenir informació del màxim nombre d'individus. Altres projectes se centren a caracteritzar un nombre molt limitat d'individus, tot i que n'obtenen una informació més extensiva de cadascun d'ells. El doctor David Comas, expert en genètica de poblacions, defensarà l'opció dels 1.000 Genomes, mentre que el doctor Tomàs Marquès i Bonet defensarà la necessitat de tenir menys genomes però millors.

Tomàs Marquès i Bonet  
Institut de Biologia Evolutiva  
(CSIC-UPF)

Han passat deu anys des de la publicació del primer esborrany del genoma humà. En aquells moments, semblava que resoldria bona part de les preguntes científiques més importants, des de l'evolució de la nostra espècie fins al descobriment de causants de malalties que fins aleshores semblaven infranquejables. Els resultats esperats, però, no acabaven d'assolir les perspectives plantejades i, amb això, s'anaven publicant actualitzacions del genoma humà —fins a nou en deu anys! Malgrat això, la versió que en tenim en ple segle XXI i comptant l'última actualització del genoma humà, conté més de sis milions de nucleòtids en vuitanta-tres llocs diferents del genoma en regions eucromàtiques per a les quals

no tenim seqüència, i hi ha desenes d'errors d' anotació en l'estructura (com, per exemple, errors de milions de seqüència en orientació). Tot això basat en un sol haplotip de representació, que, a més, és una barreja de diferents persones amb diferents haplotips! Complicat, oi? Cal recordar que la tecnologia emprada en l'acoblament del genoma era molt més costosa i laboriosa que les que utilitzem avui (Next generation, Illumina, SOLiD o 454).

En aquest context, neix el projecte dels 1.000 Genomes, amb la promesa de recollir i crear un mapa global de la variabilitat en poblacions humanes, i d'aquesta manera poder desxifrar la base genètica de malalties genètiques complexes. Inicialment, es va decidir apostar per seqüenciar molts individus amb aquesta tecnologia nova —i més barata— que ens donava, en principi, més informació, però no més del 80-90 % del genoma. Però quina informació n'obtenim? Doncs, bàsicament SNP (*Single Nucleotide Polymorphisms; polimorfisme d'un sol nucleòtid*), i ni tan sols de tot el genoma! Atesa la poca



seqüència generada (només de 2X/4X, que vol dir que, de mitjana, cada regió del genoma està coberta per només de 2 a 4 *reads* de seqüència!), moltes regions prou complicades i importants, així com la majoria de reestructuracions genòmiques, en queden totalment a fora.

Quina és l'alternativa? Al meu entendre —i en el d'una part de la comunitat científica—, seria millor agafar cinquanta genomes i mirar-los amb prou detall per entendre'n la dinàmica. Hi ha massa coses, i massa assumpcions simplistes adquirides, per pretendre que amb una informació superficial sobre els genomes serem capaços de comprendre la base genètica de les malalties complexes o de la variació natural. Hi ha massa processos biològics bàsics que encara no entenem del tot (recombinació, expressió gènica, metilació, etc.). Estic convençut que només amb el més fi escrutini de genomes acabats podrem guanyar aquest pas de coneixement bàsic que encara ens falta abans de fer el gran pas cap al MAPA (amb majúscules!) de variabilitat global humana. En fi, com deien a casa meva: «Menjar poc i païr bé.» |

David Comas  
Institut de Biologia Evolutiva  
(CSIC-UPF)

L'estudi de la diversitat genètica de la nostra espècie ha tingut des de l'inici aplicacions varies. A una part de la comunitat científica li ha interessat aquesta diversitat per diferenciar el que és *normal* respecte del que és *anormal* o patològic: és a dir, tractar d'entendre quines variants estan associades a la malaltia. D'altres, han centrat la seva recerca en l'estudi de la diversitat genètica per entendre el que ens fa diferents els uns dels altres com a població o com a individus. Aquests són els interessos principals de la genètica de poblacions humanes i de la genètica forense.

Fins i tot hi ha hagut recentment un interès creixent per la diversitat genètica en el camp de la genètica recreativa, un intent d'apropar els estudis genòmics i genealògics al terreny quotidià de la individualitat.

D'un quant temps ençà s'han engegat alguns projectes internacionals amb la finalitat d'estudiar la diversitat genètica dels humans, entre els quals cal destacar el projecte HapMap, el Genographic i l'ambiciós projecte dels 1.000 Genomes. Aquest darrer té per objectiu la seqüenciació completa de mil individus de la nostra espècie representatius de totes les poblacions humanes. En aquests moments, i fora d'aquest projecte, disposem de menys d'una desena de genomes complets, però aquesta informació limitada ens permet adonar-nos que allò que sabem de la diversitat del nostre genoma és encara insuficient.

Si bé és cert que seqüenciar molts genomes a baixa resolució mantindrà algunes regions del nostre genoma poc definides, la informació de la diversitat general de la nostra espècie serà extremament refinada. Segons quin sigui el nostre interès científic, aquest projecte pot ser insuficient (que no vol dir «inadequat») o extremament valuós (que no vol dir «complet»). Per a les persones interessades en l'estudi de regions molt concretes i específiques del genoma, aquells a qui no els és necessari disposar de tota la informació de la diversitat sobre l'espècie, el projecte dels 1.000 Genomes els semblarà coix; mentre que per als qui tenen per objectiu una visió en conjunt del genoma, una visió global de la diversitat de l'espècie, el catàleg exhaustiu de mil genomes és una oportunitat única per entendre la diversitat dels humans. Es tracta de dues visions oposades: o bé tenir un genoma totalment seqüenciat sense cap ambigüïtat o bé tenir-ne molts de complets que ens permetin una bona visió de l'espècie. Els meus interessos científics basats en la comprensió global del genoma fan que consideri el projecte dels 1.000 Genomes com una eina sense precedents. Parafraçant el company Marquès, tal com deien a casa meva: «Preu per preu, sabates grosses.» |