

Tècniques de diagnòstic molecular aplicades a l'estudi del càncer

Sergio Alonso,¹ Andreu Alibés¹ i Manuel Perucho^{1,2}

¹ Institut de Medicina Predictiva i Personalitzada del Càncer

² Institució Catalana de Recerca i Estudis Avançats (ICREA)

Correspondència: Manuel Perucho. Institut de Medicina Predictiva i Personalitzada del Càncer (IMPPC). Carretera de Can Ruti, camí de les Escoles, s/n. 08916 Badalona. Adreça electrònica: mperucho@imppc.org.

DOI: 10.2436/20.1501.02.160

ISSN (ed. impresa): 0212-3037

ISSN (ed. digital): 2013-9802

<http://revistes.iec.cat/index.php/TSCB>

Rebut: 03/11/2014

Acceptat: 20/02/2015

Resum

La transformació neoplàsica succeeix per l'acumulació consecutiva d'alteracions genètiques i epigenètiques que confereixen a les cèl·lules on ocorren avantatges selectius sobre les seves germanes, cosa que condueix al tumor per selecció clonal i evolució. Aquest treball pretén facilitar una visió general de diverses tècniques de diagnòstic d'alteracions genètiques i epigenètiques comunament utilitzades en l'estudi molecular dels càncers humans en l'àmbit de la investigació bàsica, algunes de les quals s'estan aplicant actualment o estan properes a implementar-se en l'entorn clínic. Entre les tècniques de detecció d'alteracions genètiques descriurem aquelles que identifiquen alteracions cromosòmiques i mutacions puntuals. I entre els mètodes de detecció d'alteracions epigenètiques descriurem aquells dedicats als canvis somàtics de metilació en el DNA.

Paraules clau: càncer, diagnòstic molecular, mutació, tècnica, seqüenciació.

Introducció. Alteracions somàtiques en càncer

El càncer comprèn un conjunt molt divers de malalties caracteritzades pel creixement anormal i incontrolat de les cèl·lules d'un determinat teixit, que arriba en última instància a envair i colonitzar altres teixits i òrgans. La transformació d'una cèl·lula normal en una cèl·lula cancerosa (carcinogènesi) és un procés seqüencial d'acumulació d'alteracions genètiques i epigenètiques que confereix a les cèl·lules canceroses la capacitat de divisió cel·lular sense restriccions mitjançant l'activació contínua dels senyals de divisió, la pèrdua de resposta als senyals d'arrest del cicle cel·lular i l'evasió dels mecanismes de control que, en circumstàncies normals, condueixen a la mort cel·lular programada (apoptosi). Algunes mutacions afavoreixen que les cèl·lules canceroses siguin capaces d'induir la formació de nous vasos sanguinis que aporten nutrients al tumor (angiogènesi) i d'envair els teixits circumdants i colonitzar altres òrgans (Hanahan *et al.*, 2000, 2011). Les alteracions en el DNA poden produir-se per l'acció d'agents externs (carcinògens) o interns, per la inestabilitat intrínseca del genoma, o durant la replicació cel·lular que, malgrat els complexos mecanismes de reparació del DNA, és inherentment mutagènica. Fins i tot procedint d'un mateix teixit, els càncers presenten una enorme variabilitat molecular que en reflecteix la complexa etiologia i és determinant en el pronòstic i la resposta al tractament. En les últimes dècades, s'han fet importants esforços per catalogar els tumors ja no solament pel teixit d'origen i les seves característiques clinicopatològiques, que han estat els criteris clàssics durant molts anys, sinó aplicant els més recents descobriments i tecnologies per fer una classificació molecular.

Molecular diagnostic techniques in cancer

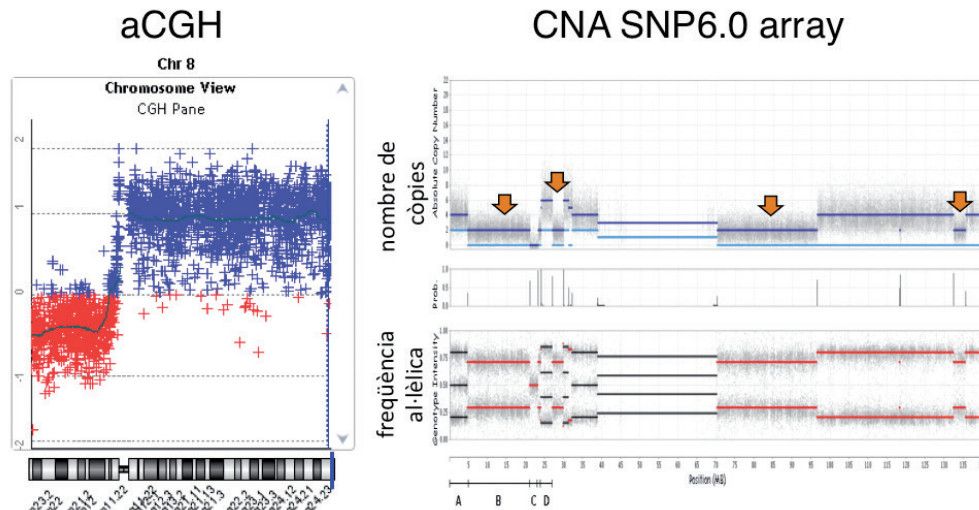
Summary

Neoplastic transformation is caused by the consecutive accumulation of genetic and epigenetic alterations providing selective growth advantages to cancer cells, ultimately leading to tumor development by clonal selection and evolution. This study aims to provide an overview of several epigenetic and genetic alteration diagnostic techniques commonly used in basic research in the molecular study of human cancers, some of which are currently implemented or are close to implementation in the clinical setting. Among the techniques for detecting genetic alterations we describe those that identify chromosomal alterations and point mutations. Among the methods of detecting epigenetic alterations we describe those that deal with somatic changes in DNA methylation.

Keywords: cancer, molecular diagnostics, mutation, techniques, sequencing.

Alteracions cromosòmiques

Les primeres observacions de l'existència d'un nombre anòmal de cromosomes (aneuploidia) en cèl·lules tumorals es remunten a principis del segle passat. Arran d'aquestes observacions, l'aneuploidia es va postular com una de les principals causes de la transformació cel·lular fins al descobriment en els anys setanta i vuitanta dels primers oncogens i supressors tumorals (revisat per Holland *et al.*, 2009). Avui dia, però, no hi ha un consens definitiu sobre si l'aneuploidia és un factor que contribueix d'una manera fonamental a la carcinogènesi o és merament un efecte secundari conseqüència de la pèrdua de control durant la mitosi (Li *et al.*, 2000; Pellman, 2007). Posteriorment, mitjançant tècniques citogenètiques amb més resolució es va descobrir que, a més d'aneuploidia, el genoma de les cèl·lules tumorals pateix reorganitzacions cromosòmiques més complexes, incloent-hi duplicacions, delecions, inversions i translocacions. Totes aquestes alteracions són de gran importància perquè poden alterar els nivells normals d'expressió gènica (Beroukhim *et al.*, 2010). Com a exemples paradigmàtics es troben l'oncogen *MYC*, localitzat en el cromosoma 8q i freqüentment amplificat en els càncers de pulmó, colorectals, gàstrics, de mama i d'úter, el supressor tumoral *APC*, localitzat en una regió del cromosoma 5q freqüentment delecionada en tumors colorectals (Bodmer *et al.*, 1989), o les freqüents translocacions que ocorren en la majoria de les leucèmies, i que tenen utilitat prognòstica i de resposta al tractament (Mitelman *et al.*, 1990; Braekeleer *et al.*, 2014). Per això, la detecció d'alteracions que afecten el nombre i l'estructura dels cromosomes és fonamental en els estudis de diagnòstic i classificació molecular del càncer.



↑ **Figura 1.** Esquerra: anàlisi del nombre de còpies en el cromosoma 8 en un cas de càncer colorectal de la nostra col·lecció mitjançant aCGH. Cada creu representa una sonda. L'escala indica el nombre de còpies en el tumor respecte a la mostra normal, en escala logarítmica en base 2. Així, el 0 indica no-alteració, l'1 la duplicació del nombre de còpies ($4n$), i el -1 la pèrdua d'un al·lel. En vermell, les sondes amb valors més petits que 0, i en blau aquelles amb valors superiors a 0. Cal notar que la pèrdua del 8p no arriba de mitjana, a un valor de -1 , cosa que indica la presència de cèl·lules sense l'alteració (molt probablement cèl·lules no tumorals). En aquest cas, la interpretació és la pèrdua del braç cromosòmic 8p i guany de 8q en la mostra tumoral. Dreta: resultats representatius de l'anàlisi d'alteracions en el nombre de còpies utilitzant xips SNP (figura obtinguda de <http://cancer.sanger.ac.uk/>). Al panell superior es mostren el nombre de còpies (sondes en gris, i el valor mitjà en línia blau fosc) i la freqüència mitjana dels al·lels menors (en blau clar), que s'obté de l'anàlisi de la freqüència al·lèlica (panell inferior). Així, l'absència d'heterozigosi en cap dels locus al llarg d'una regió (línies vermelles en el panell inferior) indica la pèrdua d'un al·lel, que en alguns casos pot anar acompanyada de la duplicació de l'altre al·lel i generar un *copy neutral* LOH (fletxes taronges) indetectable mitjançant aCGH.

Durant molts anys, la detecció d'anomalies cromosòmiques es va fer mitjançant tècniques citogenètiques clàssiques, com per exemple tincions cromosòmiques i observació mitjançant microscòpia. Més endavant, la hibridació *in situ* amb fluorescència (FISH) va millorar la resolució per detectar anomalies complexes (Langer-Safer *et al.*, 1982) i va establir les bases per al desenvolupament de la hibridació genòmica comparativa (CGH), en la qual dues mostres es marquen amb diferents fluoròfors i s'hibriden sobre preparacions de cèl·lules en metafase, cosa que permet identificar variacions en el nombre de còpies entre les dues mostres (Kallioniemi *et al.*, 1992). La tècnica CGH es va adaptar posteriorment a plataformes de microxips (aCGH), cosa que evita les laborioses preparacions metafàsiques i facilita l'automatització del procés (Lucito *et al.*, 2003). Les plataformes aCGH, i més recentment els xips de SNP, són les tecnologies més habituals per a l'anàlisi d'alteracions en el nombre de còpies cromosòmiques.

D'altra banda, poc després de la invenció de la tècnica de reacció en cadena de la polimerasa (PCR) (Mullis *et al.*, 1986), es van publicar dos mètodes, AP-PCR i RAPDS (Welsh *et al.*, 1990; Williams *et al.*, 1990), que mitjançant l'amplificació de seqüències aleatòries generaven «empremtes» genètiques o *fingerprints* característiques de cada individu, població o espècie, per la qual cosa van ser molt populars en estudis filogenètics. A més, la comparació de la intensitat de les bandes dels

fingerprints obtinguts dels tumors i les mostres normals permetien la detecció de guanys i pèrdues en el nombre de còpies (Peinado *et al.*, 1992). Aquestes tècniques van ser molt utilitzades en l'estudi d'alteracions en el nombre de còpies en diversos tipus de càncer (revisat per Samuelsson *et al.*, 2010). L'AP-PCR va ser, a més, crucial en el descobriment del fenotip mutador de microsatèl·lits en posar de manifest l'altíssima freqüència de microdeleccions en seqüències repetides en un subgrup de càncers colorectals (Ionov *et al.*, 1993).

aCGH. Els xips de CGH es basen en

el mateix principi que la tècnica CGH: la comparació de la intensitat relativa entre dos genomes marcats independentment amb diferents fluoròfors (Kallioniemi, *et al.*, 1992). La diferència és que en comptes de fer la hibridació sobre preparacions de cèl·lules en metafase es fa sobre un xip que conté sondes clonades en vectors o oligonucleòtids sintètics. Després de la hibridació, els xips s'escanegen amb càmeres de molt alta resolució i la proporció relativa de la intensitat de senyal d'emissió de cada fluoròfor en cada sonda del xip permet determinar diferències entre les dues mostres en el nombre de còpies en aquest locus (vegeu la figura 1). Les plataformes d'aCGH han assolit resolucions del voltant d'un milió de sondes, cosa que representa una mitjana d'una sonda per cada 2 kb del genoma. Els aCGH que s'utilitzen avui dia als laboratoris clínics permeten la detecció d'alteracions d'unes poques desenes de quilobases (Miller *et al.*, 2010). També es poden utilitzar *custom-arrays*, que interroguen només una fracció del genoma però a molta més resolució, i permeten detectar alteracions més petites, com les que afecten un sol exó (Boone *et al.*, 2010). La robustesa, relativa senzillesa, l'automatització del procés i preu contingut d'aquesta tecnologia, així com la profusió d'eines d'anàlisi tant comercials com de lliure accés (Lin *et al.*, 2009), ha afavorit que l'aCGH sigui avui dia una de les eines habituals en l'àmbit clínic per al diagnòstic, classificació i pronòstic en diferents tipus de càncer (Shinawi *et al.*, 2008).

Xips SNP. Encara que inicialment dissenyats per a la identificació de polimorfismes (*single nucleotide polymorphisms*, SNP), els xips SNP permeten a més detectar alteracions en el nombre de còpies. En generar informació sobre el nombre de còpies de cada al·lel en heterozigosi, els xips SNP permeten detectar un fenomen indetectable per als aCGH: la pèrdua d'heterozigositat (*loss of heterozygosity*, LOH) que no afecta el nombre total de còpies cromosòmiques (*copy neutral* LOH, vegeu la figura 1) (Leeuw *et al.*, 2011). Aquest fenomen, freqüent en càncers sanguinis, es produeix quan la pèrdua d'una regió genòmica en un al·lel es compensa per la duplicació de l'altre al·lel complementari (Langdon *et al.*, 2006), i exposa el fenotip de les variants germinals i mutacions somàtiques amplificades encara que aquestes siguin recessives. Les alteracions *copy neutral* LOH també poden ser detectades mitjançant tècniques més senzilles, com PCR múltiple de seqüències microsatèl·lit, seguides d'una anàlisi informàtica específica (Garcia-Linares *et al.*, 2012). Malgrat l'avantatge que suposa la seva capacitat per detectar *copy neutral* LOH, els xips SNP no són tan efectius com els aCGH per detectar alte-

racions que afecten regions petites del genoma. No obstant això, aquesta limitació es veu compensada pel fet que aquests xips donen informació sobre el nombre de còpies i simultàniament el genotip del pacient, cosa que permet explorar possibles factors de risc (Boyd *et al.*, 2009). Una opció nova i eficaç per combinar les dues aproximacions són els xips mixtos aCGH-SNP, alguns dels quals estan comercialitzats però encara pendents de l'aprovació en diagnòstic clínica (Wiszniewska *et al.*, 2014).

Mutacions puntuals

A més d'alteracions cromosòmiques, el genoma de les cèl·lules canceroses presenta un alt nombre de mutacions puntuals, cadascuna de les quals afecta un o uns pocs nucleòtids. La freqüència mutacional varia dramàticament entre diferents tipus de càncer (Lawrence *et al.*, 2013). Els càncers primordialment infantils, com els tumors rabdoïdes renals i els sarcomes d'Ewing, tenen una freqüència mutacional molt baixa, generalment inferior a una mutació somàtica per megabase, mentre que els càncers adults en teixits exposats als efectes deleteris dels carcinògens, per exemple els melanomes o els càncers de pulmó, en molts casos superen les cent mutacions per megabase. Totes les cèl·lules de l'organisme acumulen mutacions amb una taxa que depèn de la seva freqüència replicativa, de l'eficàcia dels mecanismes de reparació del DNA en aquest tipus cel·lular, i de l'exposició a agents carcinògens. En els càncers de l'edat adulta, la majoria de les mutacions reflecteixen l'acumulació de dany genètic patit per les cèl·lules mare prèviament a l'expansió clonal i no contribueixen necessàriament a la carcinogènesi. La raó per la qual les mutacions anteriors a la carcinogènesi es detecten fàcilment en un tumor i no en el teixit circumdant és la clonalitat de les cèl·lules tumorals (originàries d'una única cèl·lula) (Nowell, 1976). En el teixit normal, però, les mutacions es troben diluïdes a causa de l'heterogeneïtat intratissular, ja que moltes cèl·lules mare (amb diferents espectres mutacionals) contribueixen a la població cel·lular de cada teixit. Els gens mutats poden veure alterada la seva funció, regulació o els seus nivells d'expressió. Tradicionalment, els gens associats al càncer s'han catalogat en oncogenes —aquells en què l'activació o sobreexpressió afavoreix la carcinogènesi— i en gens supressors tumorals, que inhibeixen la tumorigènesi. Mentre que l'activació d'un oncogen freqüentment requereix només un esdeveniment mutacional (activació monoal·lèlica), els supressors tumorals, com a norma general, requereixen una inactivació bial·lèlica perquè la seva funció sigui completament anul·lada (Knudson, 1971). El perfil mutacional dels tumors permet, en certs càncers, determinar el pronòstic i resposta al tractament (Ince *et al.*, 2002; Arteaga, 2006; Hendifar *et al.*, 2014), i en conseqüència s'utilitzen cada vegada més habitualment com a biomarcadors en la classificació diagnòstica dels tumors.

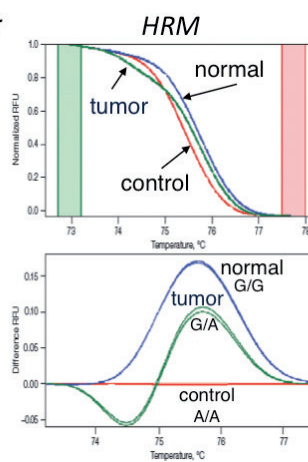
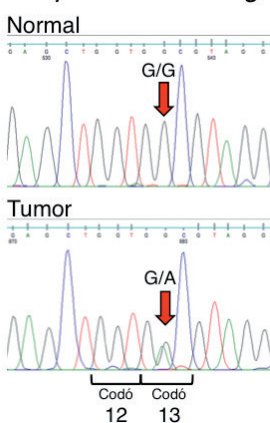
Molts dels mètodes actuals per a la detecció de mutacions puntuals es basen en l'amplificació mitjançant PCR de la seqüència estudiada. Revisions anteriors van resumir aquests mètodes de detecció pre- i immediatament post-PCR (Winter *et al.*, 1985; Almoguera *et al.*, 1989; Perucho, 1994). Fa uns anys, quan la seqüenciació era encara costosa, era freqüent analitzar mutacions mitjançant la digestió del producte de PCR amb enzims de restricció (*restriction fragment length polymorphism*, o RFLP). Si la mutació o mutacions afecten algun lloc de

restricció, es generen fragments de diferent longitud que es poden resoldre mitjançant electroforesi. No obstant això, amb la proliferació i l'abaratiment dels aparells i reactius per a la seqüenciació de DNA, els termocicladors equipats amb detecció de fluorescència en temps real (*real time PCR*) i, més recentment, les plataformes per a seqüenciació massiva en paral·lel, la majoria dels investigadors i laboratoris clínics opten per algun d'aquests mètodes per detectar mutacions en les mostres tumorals.

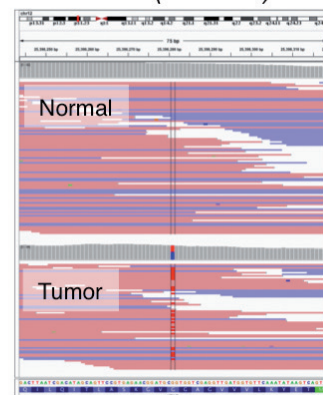
Seqüenciació Sanger. El mètode desenvolupat per F. Sanger, S. Nicklen i A. R. Coulson per seqüenciar DNA mitjançant la incorporació de didesoxinucleòtids terminadors de cadena (Sanger *et al.*, 1977) va suposar una revolució en la biologia molecular. Aquest mètode ha estat el més utilitzat durant més de vint anys, i encara avui s'utilitza de manera rutinària per a la validació de mutacions identificades amb altres mètodes. La modificació d'aquest mètode substituint el marcatge radiactiu per molècules fluorescents (Smith *et al.*, 1986) va permetre combinar l'electroforesi i la detecció en un sol aparell dotat d'un lector de fluorescència (vegeu la figura 2). Els primers seqüenciadors automatitzats comercials (AB370 d'Applied Biosystems, 1987) utilitzaven el mètode Sanger en gels verticals, cosa que limitava la quantitat de mostres que podien analitzar per dia. Els gels verticals van ser posteriorment substituïts per electroforesi en capil·lars (Kasper *et al.*, 1988), que podien instal·lar-se en grups de fins a noranta-sis capil·lars, cosa que facilitava i accelerava tremendament el procés de seqüenciació (Huang *et al.*,

.....
 + **Figura 2.** Detecció de la mutació oncogènica G>A en el codó 13 del gen KRAS en un càncer colorectal mitjançant seqüenciació Sanger (esquerra), *high-resolution melting* (centre) i mitjançant *next generation sequencing* (dreta). En l'al·lel normal, el codó 13 (GGT) codifica un residu de glicina (G), mentre que en l'al·lel mutant (GAC) codifica un aspartat. En el mètode Sanger, la mutació és detectada com la superposició de dos pics (G/A). En el HRM, a causa de la formació de molècules heterodúplex, la corba de desnaturalització dels mutants heterozigots (panell superior, línies verdes, G/A) és fàcilment diferenciable de la de l'al·lel normal (línia blava, G/G) i d'una molècula control amb la mutació (línia vermella, A/A), especialment després de normalitzar la fluorescència i la temperatura i calcular la diferència amb la corba control (panell inferior). En la seqüenciació NGS, la regió és analitzada per més de seixanta seqüències en el normal (panell superior) i en el tumor (panell inferior). En vermell clar, les seqüències en direcció *forward* i en blau aquelles en direcció *reversal*. En el tumor s'identifica la mutació per l'aparició de seqüències amb una timina en comptes d'una citosina, que corresponen a la substitució G>A en la cadena codificant que, en aquest gràfic, es troba a la cadena inferior.

Seqüenciació Sanger



NGS (Illumina)



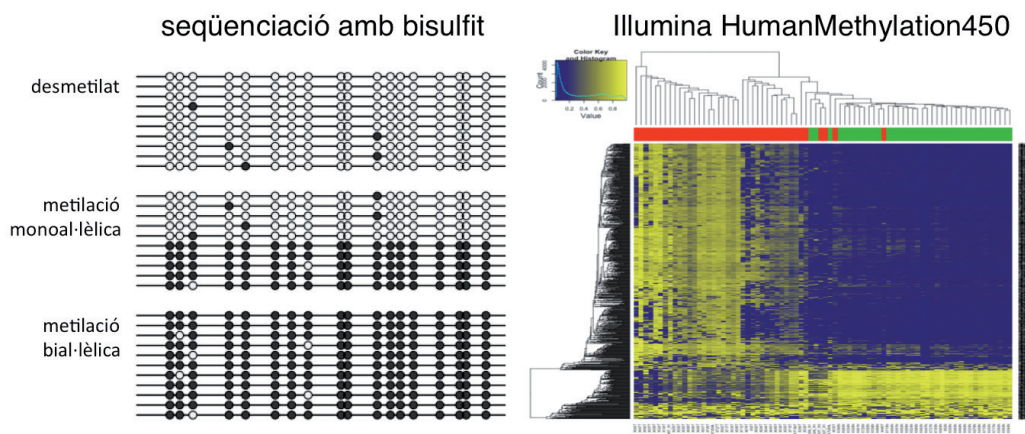
1992). Aquesta generació de seqüenciadors es van emprar en els primers projectes de seqüenciació del genoma humà (Lander *et al.*, 2001; Venter *et al.*, 2001). Encara que avui dia disposem de tècniques amb una capacitat de seqüenciació moltíssim més elevada, els seqüenciadors capilars segueixen sent el mètode més utilitzat per identificar mutacions en l'àmbit acadèmic en projectes d'una escala reduïda, i també per validar les dades obtingudes amb altres plataformes. En l'entorn clínic també són àmpliament utilitzats per la seva robustesa, preu, facilitat d'ús i per la senzillesa de l'anàlisi dels resultats obtinguts.

Determinació de mutacions mitjançant *high-resolution melting* (HRM). Una alternativa a la seqüenciació és la determinació de la temperatura de desnaturalització dels productes de PCR, mitjançant el que s'anomena *high-resolution melting* (HRM) (Wittwer *et al.*, 1997; Wittwer, 2009). La temperatura de desnaturalització (T_m) d'una molècula de DNA depèn principalment de la seva longitud i del seu contingut G+C. Per tant, les mutacions que modifiquen aquests paràmetres poden afectar la T_m . Per estudiar la presència de mutacions mitjançant HRM s'empra una molècula intercalant fluorescent, la intensitat d'emissió de la qual varia en funció de si està intercalada o no. Així, quan el DNA es desnaturalitza, la molècula intercalant s'allibera i s'observa una pèrdua de la intensitat de fluorescència. Inicialment, la resolució d'aquesta tècnica no permetia detectar canvis tan subtils com la substitució d'un sol nucleòtid, però millores posteriors en els fluoròfors intercalants emprats van augmentar molt el poder de discriminació (Wittwer *et al.*, 2003). El procediment habitual de HRM es fa en un termociclador amb un lector de fluorescència (*real-time* PCR). Primer es procedeix a l'am-

plificació mitjançant PCR de la seqüència a estudiar, immediatament seguida d'una desnaturalització amb monitoratge continu de la fluorescència emesa, de la qual es deriven les corbes de desnaturalització. En presència de dos tipus de molècules, com és el cas de les mutacions en heterozigosi en la qual hi ha un allel mutant i un altre de natural (*wild type*), la PCR genera tres productes diferents: el producte mutant, el natural i productes híbrids denominats *heterodúplex* (una cadena de la molècula és mutant i l'altra natural). Aquesta heterogeneïtat produeix una corba de desnaturalització complexa en la qual no es pot determinar amb precisió una T_m única. En aquests casos és habitual comparar les diferències en les corbes de desnaturalització després de normalitzar tant la fluorescència com el rang de temperatures, en comptes de comparar directament les T_m (vegeu la figura 2). Per la seva senzillesa, rapidesa i cost limitat, la tècnica HRM és molt útil en l'entorn clínic per a la detecció de mutacions conegudes (Erali *et al.*, 2008).

Seqüenciació massiva en paral·lel. La seqüenciació massiva en paral·lel, també anomenada *next-generation sequencing* (NGS), comprèn diverses tecnologies que s'han desenvolupat més o menys al mateix temps i que permeten la seqüenciació simultània de centenars de milers, fins i tot milions, de fragments curts de DNA. Hi ha quatre tecnologies principals que s'han disputat el mercat durant els últims anys: Solexa (Illumina), 454 (Roche), Ion Torrent i SOLiD (Applied Biosystems). Hi ha abundant literatura que compara els avantatges i limitacions de cadascuna d'aquestes plataformes, de manera que en aquest treball només citarem unes revisions que considerem molt informatives per proveir una descripció més detallada del seu funcionament (Mardis, 2008; Anson, 2009; Fox *et al.*, 2009). Independentment de la tecnologia emprada, l'anàlisi de l'enorme volum de les dades generades exigeix una alta capacitat computacional. Paradoxalment, s'ha passat d'una època en què el pas limitant de la investigació era la generació de dades a una en què aquest es troba en l'anàlisi (Koboldt *et al.*, 2010). Generalment, l'anàlisi s'inicia amb un pas de filtratge en el qual s'eliminen les seqüències de baixa qualitat i els artefactes generats pels processos d'amplificació i seqüenciació. Posteriorment, aquestes seqüències es mapen sobre una seqüència referència, que pot ser el genoma complet o una fracció d'aquest. Després del mapatge, les variacions trobades es classifiquen en germinals —aquelles presents en la mostra normal i en la tumoral— i mutacions somàtiques, exclusives de la mostra tumoral. Molt recentment, científics de Barcelona han desenvolupat una nova aproximació computacional que no requereix mapar les seqüències sobre una referència, cosa que representa un enorme avanç en facilitar

.....
 ↓ **Figura 3.** Esquerra: metilació d'un locus genòmic mitjançant seqüenciació per bisulfit. Cada línia horitzontal representa la seqüència d'un clon individual. Els cercles blancs indiquen els llocs CpG desmetilats, i els negres els CpG metilats. A la part superior, una mostra completament desmetilada (teixit normal clònic). Al mig, una mostra tumoral amb metilació bial·lèlica, i a la part inferior una mostra amb metilació bial·lèlica (o metilació monoal·lèlica acompanyada de pèrdua de l'allel no metilat). Dreta: resultats representatius de la classificació sense supervisió obtinguda mitjançant l'anàlisi de la metilació de quaranta mostres tumorals (en vermell) i trenta-vuit mostres no tumorals (en verd) de pacients amb càncer colorectal, utilitzant xips Illumina HumanMethylation450. Cada fila representa una sonda que interroga un lloc CpG únic en el genoma. Només es mostren les mil sondes més variables i que interroguen llocs CpG autosòmics. En gradació creixent de blau a groc s'indica el nivell de metilació (de no metilat a completament metilat).



la detecció de mutacions que, per la seva complexitat, són difícils o impossibles de mapar mitjançant les tècniques habituals (Moncunill *et al.*, 2014).

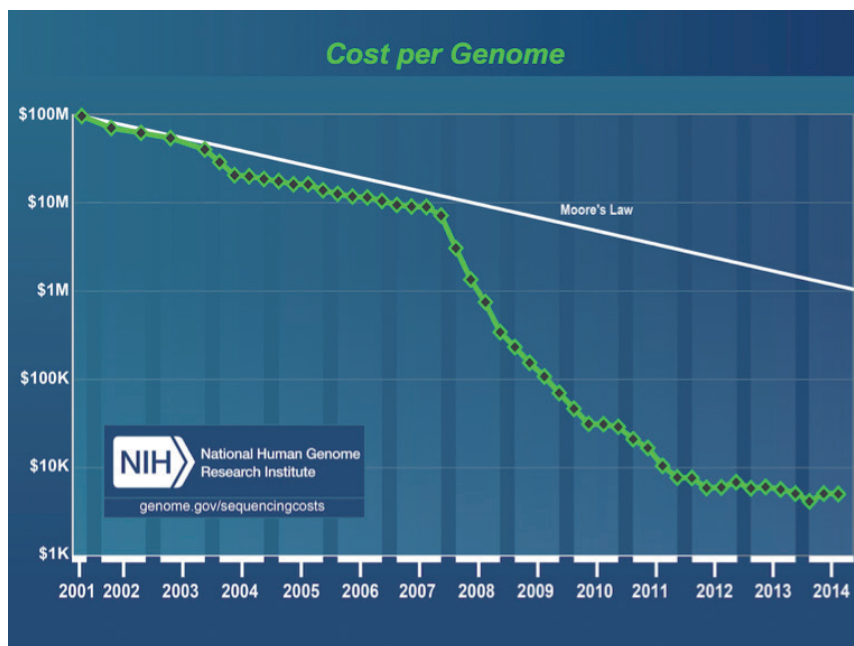
Una de les aplicacions més esteses de la NGS és la detecció de mutacions i variants germinals mitjançant la seqüenciació d'exomes, és a dir, les regions genòmiques corresponents a les seqüències codificants dels gens. Els fragments genòmics corresponents als exomes es capturen mitjançant hibridació

amb seqüències sintètiques complementàries a tots els exons del genoma humà. La captura es fa habitualment mitjançant hibridació sobre xips (Choi *et al.*, 2009) o en solució utilitzant sondes marcades amb biotina que es poden aïllar mitjançant estreptavidina (Mamanova *et al.*, 2010). Aquesta captura redueix la regió a seqüenciar (*coverage*) a unes 30-60 Mb per mostra, en comptes de les 3.200 Mb del genoma complet. Aquesta reducció permet augmentar molt la profunditat de lectura (*depth*), és a dir, el nombre de lectures per cada base, facilita el processament i seqüenciació simultània de diverses mostres i abarateix així els costos tant de la seqüenciació com de l'anàlisi posterior. La seqüenciació d'exomes irremeiablement ignora aquelles mutacions en regions no codificants, però moltes, si no la majoria, de les mutacions causals s'esperen trobar a les regions codificants. A més, és generalment més senzill inferir l'efecte de les mutacions en seqüències codificants que en seqüències no codificants. Pel seu equilibri entre cost i potència, la seqüenciació d'exomes és molt habitual avui dia en l'àmbit acadèmic per definir el perfil mutacional dels càncers i, a poc a poc, s'està obrint pas en l'àmbit clínic.

L'espectacular descens dels costos de seqüenciació (vegeu la figura 4) ha promogut que molts equips d'investigació es decideixin per la seqüenciació de genomes complets, una cosa que fins fa només uns anys era absolutament prohibitiva. L'avantatge fonamental de la seqüenciació del genoma complet és evident: s'obté una visió completa de les alteracions somàtiques, tant les cromosòmiques com les puntuals, al llarg de tot el genoma. A causa de l'abaratiment d'aquesta tecnologia i la possible entrada al mercat de tecnologies fins i tot més poderoses és molt probable que la seqüenciació de genomes complets s'imposarà en un futur proper com la tècnica d'elecció no només per a l'anàlisi mutacional en càncers, sinó fins i tot com a alternativa als xips per a la determinació del genotip dels pacients de manera rutinària. Avui dia, però, l'anàlisi bioinformàtica suposa un important cost en temps i recursos i possiblement es necessitaran noves aproximacions computacionals per millorar i estandarditzar l'anàlisi abans de l'aplicació rutinària en l'àmbit clínic.

Alteracions epigenètiques

Les cèl·lules canceroses exhibeixen perfils epigenètics anormals que afecten els patrons d'expressió gènica i, en alguns casos, l'estabilitat mateixa del genoma (Jones *et al.*, 2002). Entre les modificacions epigenètiques que s'han estudiat amb més profunditat destaquen les alteracions en els patrons de metilcitosa en el DNA, associats a canvis en l'estat de compactació i accessibilitat de la cromatina. En la majoria dels genomes dels vertebrats les citosines poden metilar-se per acció d'enzims nuclears denominats *DNA-metiltransferases* (DNMT), formant 5-metilcitosa. En vertebrats, la metilació de citosines està restringida gairebé exclusivament als llocs CpG (un nucleòtid de citosina seguit de guanina), que estan infrarepresentats en el genoma dels vertebrats a causa de la inestabilitat intrínseca de la metilcitosa que muta a timina fixant una mutació C→T (Russell *et al.*, 1976; Bird, 1980). Hi ha regions en el DNA anomenades *illes CpG* (CGI), d'uns 200-1.000 pb i amb més proporció de CpG (Gardiner-Garden *et al.*, 1987), generalment associades a seqüències repetitives amb alt contingut en G+C, així com



↑ Figura 4. Disminució del cost de la seqüenciació d'un genoma humà en els darrers quinze anys (figura obtinguda del National Human Genome Research Institute, www.genome.gov). Després de la finalització del projecte de seqüenciació del genoma humà (2001), el cost de seqüenciació seguia aproximadament la llei de Moore, inicialment postulada per predir el descens exponencial del cost de computació, i que s'ha emprat per determinar la reducció en el cost d'altres tecnologies. La implementació dels seqüenciadors de nova generació (NGS) el 2007 va suposar una enorme reducció en el cost. Es preveu que, amb la probable irrupció de noves tecnologies, en els propers anys el cost de seqüenciar un genoma complet sigui significativament de menys de mil dòlars.

en la regió promotora d'aproximadament el 60-70 % dels gens humans (Illingworth *et al.*, 2009). En general, la metilació de les CGI en promotors gènics s'associa al seu silenciament transcripcional. A principis dels anys vuitanta es va descobrir que el genoma de les cèl·lules canceroses estava hipometilat (Diala *et al.*, 1982; Feinberg *et al.*, 1983). La hipometilació genòmica reflecteix majoritàriament la desmetilació de seqüències repetitives, que representen més del 50 % del genoma humà, i en menys mesura la desmetilació de les regions no promotores dels gens (*gene bodies*). Encara que s'ha descrit que la hipometilació d'alguns oncogens està associada a la seva sobreexpressió (Feinberg *et al.*, 1983; Vachtenheim *et al.*, 1994), avui dia es creu que l'efecte més important de la hipometilació del DNA en càncer és la reactivació de l'expressió de retrotransposons (Bestor *et al.*, 1996) i l'augment d'inestabilitat genòmica (Ehrlich, 2002). Posteriorment, es va descobrir que alguns tumors presentaven, a més, hipermetilació localitzada en les regions promotores de certs gens implicats en la tumorigènesi i el desenvolupament tumoral (Greger *et al.*, 1989, 1994; Herman *et al.*, 1994, 1995; Gonzalez-Zulueta *et al.*, 1995). Aquests descobriments van iniciar una veritable revolució en el camp de la genètica del càncer, en demostrar que molts gens supressors tumorals resultaven inactivats no solament per mutacions genètiques, sinó per alteracions en la metilació del DNA (revisat per Esteller, 2002). Un exemple paradigmàtic d'aquest fenomen és, sens dubte, la hipermetilació del gen *MLH1*, que és el principal causant dels càncers colorectals esporàdics amb el fenotip mutador (Kane *et al.*, 1997). L'epigenètica ha obert un nou front en la batalla contra el

càncer, basat en el disseny i aplicació de medicaments dirigits a revertir la metilació dels supressors tumorals o, en alguns casos, a explotar les conseqüències dels canvis epigenètics (Yoo *et al.*, 2006).

Els canvis en la metilació del DNA són possiblement les alteracions epigenètiques en càncer més estudiades i millor conegudes. Durant els últims trenta anys s'han desenvolupat i perfeccionat diversos mètodes per identificar aquest tipus d'alteracions. Molts dels mètodes inicials es van basar en l'ús d'endonucleases de restricció sensibles a la metilació. No obstant això, els mètodes actuals més habituals es basen en la transformació química del DNA mitjançant tractament amb bisulfit, que modifica les citosines no metilades i deixa intactes les citosines metilades (Frommer *et al.*, 1992; Clark *et al.*, 1994).

Detecció d'alteracions mitjançant enzims de restricció sensibles a la metilació. Una de les primeres aproximacions experimentals per detectar canvis en els patrons de metilació del DNA va ser l'ús d'endonucleases de restricció sensibles a la metilació. La gran majoria de les endonucleases de restricció utilitzades en biologia molecular provenen dels sistemes bacterians de metilació-restricció de tipus II, que combinen una parella d'enzims, una metilasa i una endonucleasa que reconeixen la mateixa seqüència del DNA. La metilasa actua sobre el DNA propi després de la síntesi, i protegeix de l'acció de l'endonucleasa. El DNA forà, però, no té el mateix patró de metilació i, per tant, queda exposat a digestió de l'endonucleasa (Wilson, 1991). Així, l'estudi de la mida dels fragments del DNA tractat amb enzims de restricció que contenen un o més llocs CpG en la seva seqüència de reconeixement permet identificar llocs CpG metilats en un locus determinat. Posteriorment es van desenvolupar diverses tècniques que permetien analitzar canvis en la metilació d'un major nombre de locus simultàniament mitjançant electroforesi unidimensional generant *fingerprints*, com la *methylation-sensitive arbitrarily primed PCR* (MS-AP-PCR), *restriction fingerprinting* (MSRF), *amplification of inter methylated sites* (AIMS) i *methylation sensitive amplification length polymorphism* (MS-AFLP), o mitjançant electroforesi bidimensional, com la *restriction landmark genomic scanning for methylation* (RLGS-M) (revisat per Samuelsson *et al.*, 2010). En general, aquestes tècniques han estat desplaçades pels mètodes basats en la transformació del DNA amb bisulfit, molt especialment els xips de metilació d'Illumina. No obstant això, la tècnica MS-AFLP ha estat adaptada a plataformes de xips, utilitzant com a sondes els fragments de MS-AFLP clonats (Yamamoto *et al.*, 2004) o oligonucleòtids sintètics (Koizumi *et al.*, 2012). El nostre grup ha utilitzat aquests xips de metilació per identificar alteracions epigenètiques en la mucosa colònica de pacients amb colitis ulcerosa (Koizumi *et al.*, 2012) i més recentment per investigar la metilació d'una regió interna del gen *AXIN2* com a biomarcador associat a la ruta de transformació dels adenomes serrats en carcinomes (Muto *et al.*, 2014).

Captura de DNA metilat. Una altra de les tècniques habituals és l'enriquiment de DNA metilat mitjançant anticossos que reconeixen directament la metilcitosina (MeDIP) o dominis d'unió a citosina metilada (MBD) (Rauch *et al.*, 2006; Jacinto *et al.*, 2008). Conceptualment, les dues tècniques són molt similars. Es basen en la captura de la fracció del DNA que es troba metilat i l'anàlisi posterior mitjançant hibridació en *tiling-microarrays* o mitjançant NGS (Down *et al.*, 2008) per detectar regions amb metilació diferencial (DMR) entre dues o més mostres. Avui dia hi ha múltiples productes comercials basats en aquestes tecnologies. No obstant això, cal tenir en compte que els anticossos i

les proteïnes amb dominis MBD presenten una afinitat diferent per la 5-metilcitosina, cosa que resulta en diferents biaixos en funció de la tècnica emprada (Nair *et al.*, 2011).

Transformació del DNA amb bisulfit. La transformació del DNA mitjançant tractament amb bisulfit ha facilitat enormement l'estudi de la metilació del DNA. Aquest tractament provoca la desaminació de les citosines, i les converteix en uracils. La 5-metilcitosina, però, és molt més resistent a la transformació i roman majoritàriament inalterada. Després del tractament, el DNA és amplificat mitjançant PCR. La Taq-polimerasa confon els residus d'uracil amb residus de timina, per la qual cosa introdueix una adenina com a base complementària. D'aquesta manera, les citosines originalment desmetilades, però no les metilades, són substituïdes per timines en el cicle d'amplificació següent (Frommer *et al.*, 1992; Clark *et al.*, 1994). Encara que és possible, en general es desaconsella seqüenciar directament el producte d'amplificació, i és preferible clonar i seqüenciar clons individuals i obtenir així l'estat de metilació de molècules individuals (vegeu la figura 3). S'han desenvolupat molts mètodes basats en el tractament del DNA amb bisulfit (revisat per Jordà *et al.*, 2010). Una opció molt utilitzada pel seu moderat cost és la digestió del producte d'amplificació amb enzims de restricció, la seqüència de reconeixement dels quals inclou dinucleòtids CpG, que són transformats a TpG a les molècules originalment desmetilades i generen, per tant, fragments de diferent longitud (*combined bisulfite and restriction analysis*, COBRA) (Xiong *et al.*, 1997). Una altra alternativa econòmica que no requereix seqüenciar és la PCR sensible a metilació (*methylation-sensitive PCR*, o MSP), potser un dels mètodes més emprats per investigar l'estat de metilació de locus individuals (Herman *et al.*, 1996). Per estudiar una seqüència mitjançant MSP es dissenyen dues parelles d'encebadors (*primers*): una específica per al producte de transformació de la seqüència desmetilada (considerant totes les citosines transformades a uracil), i l'altra parella reconeix el producte de transformació de la seqüència metilada (totes les citosines transformades a uracil, excepte aquelles en llocs CpG). Tot i que la MSP és molt sensible, és també propensa a generar falsos positius. Una versió quantitativa de la MSP és el mètode MethyLight, que empra sondes fluorescents per quantificar la proporció de molècules metilades (o desmetilades) en una mostra (Eads *et al.*, 2000). Altres opcions menys habituals són la determinació de l'estat de metilació mitjana del producte amplificat mitjançant MS-SSCA (*methylation-sensitive single stranded conformation analysis*) (Bianco *et al.*, 1999) o HRM (Worm *et al.*, 2001), la determinació de l'estat de metilació d'uns pocs llocs CpG interns mitjançant PCR d'inhibició *headloop* (Rand *et al.*, 2005), i la quantificació de l'estat de metilació d'un únic lloc CpG mitjançant Ms-SNuPE (Gonzalvo *et al.*, 1997), o de tots els llocs CpG del producte d'amplificació mitjançant piroseqüenciament (Tost *et al.*, 2003). Totes aquestes tècniques van ser dissenyades per estudiar la metilació d'un sol locus i, excepte per a estudis molt específics o per confirmar les observacions fetes mitjançant altres mètodes, s'han vist relegades per tecnologies més recents que permeten analitzar diversos locus simultàniament.

Microxips basats en el tractament amb bisulfit. Encara que hi ha diverses plataformes comercials per estudiar la metilació del DNA, possiblement la més utilitzada avui dia són els xips desenvolupats per Illumina. La primera generació d'aquests xips disposava d'unes 1.500 sondes basades en la tecnologia GoldenGate (Bibikova *et*

al., 2006). Com a exemple pràctic del poder d'aquesta plataforma per classificar càncers, el 2011 es va utilitzar per generar mapes epigenòmics en més de 1.600 mostres tumorals (Sandoval *et al.*, 2011). Posteriorment es va desenvolupar una segona generació de xips basats en la tecnologia Infinium. El primer d'aquests xips, denominat HumanMethylation27, interrogava uns 27.000 llocs CpG localitzats principalment en promotors gènics (Bibikova *et al.*, 2009). Aquests xips contenen sondes anomenades de tipus I, en el qual cada CpG és interrogat per dues sondes independents, una complementària a la seqüència desmetilada i una altra a la seqüència metilada. L'assaig analitza l'extensió de la base que es troba a continuació del lloc CpG interrogat. Així, la intensitat relativa entre els dos tipus de sondes per a cada lloc CpG permet estimar la proporció de molècules metilades en aquesta posició. Poc després es va comercialitzar el HumanMethylation450, amb molta més resolució, en interrogar uns 480.000 llocs CpG (aproximadament l'1,7 % de tots els llocs CpG del genoma), localitzats tant en regions promotores com en regions intra i intergèniques (vegeu la figura 3) (Bibikova *et al.*, 2011; Sandoval *et al.*, 2011). Aquesta generació de xips inclou sondes anomenades de tipus II, en què es dissenya una sola sonda per a cada lloc CpG i s'analitza la incorporació d'una citosina o una timina en el lloc estudiat. Pel seu preu limitat i alta resolució, els xips HumanMethylation450 són molt emprats per moltíssims grups de recerca, i han estat la plataforma triada per al projecte de l'International Cancer Genome Consortium (ICGC, <https://icgc.org/>) per analitzar la metilació dels tumors.

Seqüenciació de genomes transformats amb bisulfit. Actualment, la tècnica de més resolució per determinar la metilació del DNA és la seqüenciació completa del genoma després de la transformació amb bisulfit (*whole genome bisulfite sequencing*, o WGBS) (Lister *et al.*, 2009; revisat per Stirzaker *et al.*, 2014). Aquesta tècnica permet la identificació de l'estat de metilació de la gran majoria dels llocs CpG genòmics, excloent aquells en seqüències repetitives que sempre suposen un problema per mapar-los. No obstant això, aquesta aproximació comporta considerables obstacles tècnics. Comparat amb la seqüenciació de genomes sense transformar, per a la WGBS es requereix més quantitat de material inicial a l'hora de generar les biblioteques, ja que una gran part del DNA és degradat durant el tractament amb bisulfit. Després de la seqüenciació, totes les citosines fora de llocs CpG hauran estat substituïdes, cosa que redueix la informació útil (de quatre bases a tres) per mapar les seqüències obtingudes. D'altra banda, els llocs CpG poden estar metilats o no, cosa que genera una variabilitat per a la qual no estan dissenyats la majoria dels programes de mapatge de seqüències. L'estratègia més estesa és substituir *in silico* les citosines per timines (o les guanines per adenines en funció de l'orientació) en totes les seqüències obtingudes i, a continuació, mapar-les sobre dues versions del genoma humà, cadascuna generada a partir d'una de les cadenes substituint les citosines per timines (i llavors les cadenes perden la seva complementarietat). Després del mapatge, es reposa la informació referent a l'estat de metilació dels llocs CpG per a cadascuna de les seqüències. A causa de l'elevat cost i la complexitat de l'anàlisi, sobretot si es compara amb

les plataformes de xips actuals, la WGBS no s'ha estès encara en l'àmbit acadèmic excepte per a estudis en què s'analitzen un nombre petit de mostres. S'han desenvolupat alternatives per reduir la complexitat i el cost, com seqüenciar només una part del genoma (*reduced representation bisulfite sequencing*, RRBS) (Meissner *et al.*, 2005) o millorar l'eficiència de la construcció de les biblioteques mitjançant la fragmentació del DNA amb la transposasa Tn5 (*tagmentation-based* WGBS, o T-WGBS) (Wang *et al.*, 2013).

Recapitulació i direccions futures

Avui dia disposem de plataformes molt potents per a la classificació i caracterització dels càncers en funció de les seves alteracions genòmiques i epigenòmiques. Algunes d'aquestes plataformes estan encara restringides a l'àmbit acadèmic però a poc a poc van introduint-se en la pràctica clínica per a la diagnòstic molecular. La PCR, la seqüenciació Sanger i els xips són probablement les tecnologies més comunes avui dia per la seva fiabilitat, relativa senzillesa i cost limitat. No obstant això, la seqüenciació massiva en paral·lel ha suposat una tremenda revolució en el camp, i ha facilitat l'anàlisi de pràcticament tots els tipus d'alteracions esmentades en aquest treball. Els seus inconvenients actuals són l'elevat cost dels equips i els reactius emprats per a la seqüenciació, de manera que l'ús d'aquestes tecnologies en l'àmbit acadèmic està restringit a unitats especialitzades de seqüenciació, públiques o privades. En qualsevol cas, els costos han disminuït de manera espectacular en els darrers anys (vegeu la figura 4), i es preveu que amb el desenvolupament definitiu de l'anomenada *seqüenciació de tercera generació*, amb nous mètodes capaços de generar desenes de milers de seqüències de més de 20 kb cadascuna i partir de molècules individuals, els preus baixaran per sota dels mil dòlars per genoma (cent mil vegades menys del que va costar el primer genoma humà el 2001). D'altra banda, en tractar-se de tecnologies molt joves que generen una enorme quantitat de dades, els mètodes d'anàlisi computacional són complexos i estan en continu desenvolupament, cosa que en dificulta la implementació immediata en l'àmbit clínic, en què la fiabilitat i l'estandardització dels protocols de diagnòstic són absolutament crítics. Malgrat aquestes dificultats, la implementació clínica de les tecnologies d'anàlisi molecular, tant les actuals com les que estan en fase de desenvolupament, és fonamental per a la futura aplicació d'una medicina de precisió per al càncer, en la qual se seleccionarà la teràpia més apropiada per a cada pacient en funció del seu genotip i de la informació genètica/epigenètica del càncer que l'afecta (Esplin *et al.*, 2014). No és improbable, a més, que el desenvolupament de noves tecnologies —avui completament impredecibles— facilitarà encara més l'aplicació de la informació genètica i epigenètica en el diagnòstic dels pacients, tal com ja ha passat tantes i tantes vegades en el passat amb la seqüenciació del DNA, la PCR, els microxips, la NGS, i un llarguíssim etcètera. Parafraçant Sydney Brenner: «Progress in science depends on new techniques, new discoveries and new ideas, probably in that order» (Brenner, 2002).

Treball finançat per Spanish Ministry of Health Grants FIS PI09/2444 i FIS PI12/00511.

Bibliografia

- ALMOGUERA, C. [et al.] (1989). «Application of the Polymerase Chain Reaction for the detection of single-base substitutions by the RNase A mismatch cleavage method». A: H. A. Erlich [et al.] (ed.). *Current Communications in Molecular Biology, Polymerase Chain Reaction*. Cold Spring Harbor Laboratory, p. 37-41.
- ANSORGE, W. J. (2009). «Next-generation DNA sequencing techniques». *N. Biotechnol.*, 25 (4): 195-203.
- ARTEAGA, C. L. (2006). «EGF receptor mutations in lung cancer: from humans to mice and maybe back to humans». *Cancer Cell*, 9 (6): 421-423.
- BEROUKHIM, R. [et al.] (2010). «The landscape of somatic copy-number alteration across human cancers». *Nature*, 463 (7283): 899-905.
- BESTOR, T. H. [et al.] (1996). «Creation of genomic methylation patterns». *Nat. Genet.*, 12 (4): 363-367.
- BIANCO, T. [et al.] (1999). «Methylation-sensitive, single-strand conformation analysis (MS-SSCA): A rapid method to screen for and analyze methylation». *Hum. Mutat.*, 14 (4): 289-293.
- BIBIKOVA, M. [et al.] (2006). «High-throughput DNA methylation profiling using universal bead arrays». *Genome Res.*, 16 (3): 383-393.
- (2009). «Genome-wide DNA methylation profiling using Infinium(R) assay». *Epi-genomics*, 1 (1): 177-200.
- (2011). «High density DNA methylation array with single CpG site resolution». *Genomics*, 98 (4): 288-295.
- BIRD, A. P. (1980). «DNA methylation and the frequency of CpG in animal DNA». *Nucleic Acids Res.*, 8 (7): 1499-1504.
- BODMER, W. F. [et al.] (1989). «Genetic analysis of colorectal cancer». *Princess Takamatsu Symp.*, 20: 49-59.
- BOONE, P. M. [et al.] (2010). «Detection of clinically relevant exonic copy-number changes by array CGH». *Hum. Mutat.*, 31 (12): 1326-1342.
- BOYD, L. K. [et al.] (2009). «Use of SNPs in cancer predisposition analysis, diagnosis and prognosis: tools and prospects». *Expert Opin. Med. Diagn.*, 3 (3): 313-326.
- BRAEKELEER, E. de [et al.] (2014). «Genetic diagnosis in malignant hemopathies: from cytogenetics to next-generation sequencing». *Expert Rev. Mol. Diagn.*, 14 (2): 127-129.
- BRENNER, S. (2002). «Life sentences: Detective Rummage investigates». *Genome Biology*, 3 (9): comment1013.1-comment1013.2.
- CHOI, M. [et al.] (2009). «Genetic diagnosis by whole exome capture and massively parallel DNA sequencing». *Proc. Natl. Acad. Sci. USA*, 106 (45): 19096-19101.
- CLARK, S. J. [et al.] (1994). «High sensitivity mapping of methylated cytosines». *Nucleic Acids Res.*, 22 (15): 2990-2997.
- DIALA, E. S. [et al.] (1982). «Hypomethylation of HeLa cell DNA and the absence of 5-methylcytosine in SV40 and adenovirus (type 2) DNA: analysis by HPLC». *Biochem. Biophys. Res. Commun.*, 107 (1): 19-26.
- DOWN, T. A. [et al.] (2008). «A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis». *Nat. Biotechnol.*, 26 (7): 779-785.
- EADS, C. A. [et al.] (2000). «MethylLight: a high-throughput assay to measure DNA methylation». *Nucleic Acids Res.*, 28 (8): E32.
- EHRlich, M. (2002). «DNA methylation in cancer: too much, but also too little». *Oncogene*, 21 (35): 5400-5413.
- ERALL, M. [et al.] (2008). «High resolution melting applications for clinical laboratory medicine». *Exp. Mol. Pathol.*, 85 (1): 50-58.
- ESTELLER, M. (2002). «CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future». *Oncogene*, 21 (35): 5427-5440.
- FEINBERG, A. P. [et al.] (1983a). «Hypomethylation distinguishes genes of some human cancers from their normal counterparts». *Nature*, 301 (5895): 89-92.
- (1983b). «Hypomethylation of ras oncogenes in primary human cancers». *Biochem. Biophys. Res. Commun.*, 111 (1): 47-54.
- FOX, S. [et al.] (2009). «Applications of ultra-high-throughput sequencing». *Methods Mol. Biol.*, 553: 79-108.
- FROMMER, M. [et al.] (1992). «A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands». *Proc. Natl. Acad. Sci. USA*, 89 (5): 1827-1831.
- GARCIA-LINARES, C. [et al.] (2012). «Applying microsatellite multiplex PCR analysis (MMPA) for determining allele copy-number status and percentage of normal cells within tumors». *PLoS One*, 7 (8): e42682.
- GARDINER-GARDEN, M. [et al.] (1987). «CpG islands in vertebrate genomes». *J. Mol. Biol.*, 196 (2): 261-282.
- GONZALEZ-ZULUETA, M. [et al.] (1995). «Methylation of the 5' CpG island of the p16/CDKN2 tumor suppressor gene in normal and transformed human tissues correlates with gene silencing». *Cancer Res.*, 55 (20): 4531-4535.
- GONZALGO, M. L. [et al.] (1997). «Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (MSP-PCR)». *Nucleic Acids Res.*, 25 (12): 2529-2531.
- GREGER, V. [et al.] (1989). «Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma». *Hum. Genet.*, 83 (2): 155-158.
- (1994). «Frequency and parental origin of hypermethylated RB1 alleles in retinoblastoma». *Hum. Genet.*, 94 (5): 491-496.
- HANAHAN, D. [et al.] (2000). «The hallmarks of cancer». *Cell*, 100 (1): 57-70.
- (2011). «Hallmarks of cancer: the next generation». *Cell*, 144 (5): 646-674.
- HENDIFAR, A. [et al.] (2014). «Biomarker-driven EGFR therapy improves outcomes in patients with metastatic colorectal cancer». *Expert Rev. Anticancer Ther.*, 14 (9): 1051-1061.
- HERMAN, J. G. [et al.] (1994). «Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma». *Proc. Natl. Acad. Sci. USA*, 91 (21): 9700-9704.
- (1995). «Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers». *Cancer Res.*, 55 (20): 4525-4530.
- (1996). «Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands». *Proc. Natl. Acad. Sci. USA*, 93 (18): 9821-9826.
- HOLLAND, A. J. [et al.] (2009). «Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis». *Nat. Rev. Mol. Cell Biol.*, 10 (7): 478-487.
- HUANG, X. C. [et al.] (1992). «DNA sequencing using capillary array electrophoresis». *Anal. Chem.*, 64 (18): 2149-2154.
- ILLINGWORTH, R. S. [et al.] (2009). «CpG islands—a rough guide». *FEBS Lett*, 583 (11): 1713-1720.
- INCE, T. A. [et al.] (2002). «Functional genomics and the breast cancer problem». *Cancer Cell*, 1 (1): 15-17.
- IONOV, Y. [et al.] (1993). «Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis». *Nature*, 363 (6429): 558-561.
- JACINTO, F. V. [et al.] (2008). «Methyl-DNA immunoprecipitation (MeDIP): hunting down the DNA methylome». *Biotechniques*, 44 (1): 35, 37, 39 passim.
- JONES, P. A. [et al.] (2002). «The fundamental role of epigenetic events in cancer». *Nat. Rev. Genet.*, 3 (6): 415-428.
- JORDA, M. [et al.] (2010). «Methods for DNA methylation analysis and applications in colon cancer». *Mutat. Res.*, 693 (1-2): 84-93.
- KALLIONIEMI, A. [et al.] (1992). «Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors». *Science*, 258 (5083): 818-821.
- KANE, M. F. [et al.] (1997). «Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines». *Cancer Res.*, 57 (5): 808-811.
- KASPER, T. J. [et al.] (1988). «Separation and detection of DNA by capillary electrophoresis». *J. Chromatogr.*, 458: 303-312.
- KNUDSON, A. G., Jr. (1971). «Mutation and cancer: statistical study of retinoblastoma». *Proc. Natl. Acad. Sci. USA*, 68 (4): 820-823.
- KOBOLDT, D. C. [et al.] (2010). «Challenges of sequencing human genomes». *Brief Bioinform.*, 11 (5): 484-498.
- KOIZUMI, K. [et al.] (2012). «Array-based identification of common DNA methylation alterations in ulcerative colitis». *Int. J. Oncol.*, 40 (4): 983-994.
- LANDER, E. S. [et al.] (2001). «Initial sequencing and analysis of the human genome». *Nature*, 409 (6822): 860-921.
- LANGDON, J. A. [et al.] (2006). «Combined genome-wide allelotyping and copy number analysis identify frequent genetic losses without copy number reduction in medulloblastoma». *Genes Chromosomes Cancer*, 45 (1): 47-60.
- LANGER-SAFER, P. R. [et al.] (1982). «Immunological method for mapping genes on *Drosophila polytene* chromosomes». *Proc. Natl. Acad. Sci. USA*, 79 (14): 4381-4385.
- LAWRENCE, M. S. [et al.] (2013). «Mutational heterogeneity in cancer and the search for new cancer-associated genes». *Nature*, 499 (7457): 214-218.
- LEEUEW, N. de [et al.] (2011). «SNP array analysis in constitutional and cancer genome diagnostics—copy number variants, genotyping and quality control». *Cytogenet. Genome Res.*, 135 (3-4): 212-221.
- LI, R. [et al.] (2000). «Aneuploidy vs. gene mutation hypothesis of cancer: recent study claims mutation but is found to support aneuploidy». *Proc. Natl. Acad. Sci. USA*, 97 (7): 3236-3241.
- LIN, S. [et al.] (2009). «Using free and open-source bioconductor packages to analyze array comparative genomics hybridization (aCGH) Data». *Curr. Genomics*, 10 (1): 60-63.
- LISTER, R. [et al.] (2009). «Human DNA methylomes at base resolution show widespread epigenomic differences». *Nature*, 462 (7271): 315-322.
- LUCITO, R. [et al.] (2003). «Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variations». *Genome Res.*, 13 (10): 2291-2305.
- MAMANOVA, L. [et al.] (2010). «Target-enrichment strategies for next-generation sequencing». *Nat. Methods*, 7 (2): 111-118.
- MARDIS, E. R. (2008). «Next-generation DNA sequencing methods». *Annu. Rev. Genomics Hum. Genet.*, 9: 387-402.
- MEISSNER, A. [et al.] (2005). «Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis». *Nucleic Acids Res.*, 33 (18): 5868-5877.
- MILLER, D. T. [et al.] (2010). «Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies». *Am. J. Hum. Genet.*, 86 (5): 749-764.
- MITELMAN, F. [et al.] (1990). «Chromosome abnormalities in cancer». *Cancer Detect. Prev.*, 14 (5): 527-537.
- MONCUNILL, V. [et al.] (2014). «Comprehensive characterization of complex structural variations in cancer by directly comparing genome sequence reads». *Nat. Biotechnol.*, 32 (11): 1106-1112.
- MULLIS, K. [et al.] (1986). «Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction». *Cold Spring Harb. Symp. Quant. Biol.*, 51 Pt 1: 263-73.
- MUTO, Y. [et al.] (2014). «DNA methylation alterations of AXIN2 in serrated adenomas and colon carcinomas with microsatellite instability». *BMC Cancer*, 14: 466.
- NAIR, S. S. [et al.] (2011). «Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias». *Epigenetics*, 6 (1): 34-44.
- NOWELL, P. C. (1976). «The clonal evolution of tumor cell populations». *Science*, 194 (4260): 23-28.
- PEINADO, M. A. [et al.] (1992). «Isolation and characterization of allelic losses and gains

- in colorectal tumors by arbitrarily primed polymerase chain reaction». *Proc. Natl. Acad. Sci. USA*, 89 (21): 10065-10069.
- PELLMAN, D. (2007). «Cell biology: aneuploidy and cancer». *Nature*, 446 (7131): 38-39.
- PERUCHO, M. (1994). «PCR and cancer diagnostics: detection and characterization of single point mutations in oncogenes and antioncogenes». A: K. B. Mullis [et al.] (ed.). *The polymerase chain reaction*. Birkhauser: Springer-Verlag, p. 369-394.
- RAND, K. N. [et al.] (2005). «Headloop suppression PCR and its application to selective amplification of methylated DNA sequences». *Nucleic Acids Res.*, 33 (14): e127.
- RAUCH, T. [et al.] (2006). «MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells». *Cancer Res.*, 66 (16): 7939-7947.
- RUSSELL, G. J. [et al.] (1976). «Doublet frequency analysis of fractionated vertebrate nuclear DNA». *J. Mol. Biol.*, 108 (1): 1-23.
- SAMUELSSON, J. K. [et al.] (2010). «DNA fingerprinting techniques for the analysis of genetic and epigenetic alterations in colorectal cancer». *Mutat. Res.*, 693 (1-2): 61-76.
- SANDOVAL, J. [et al.] (2011). «Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome». *Epigenetics*, 6 (6): 692-702.
- SANGER, F. [et al.] (1977). «DNA sequencing with chain-terminating inhibitors». *Proc. Natl. Acad. Sci. USA*, 74 (12): 5463-5467.
- SHINAWI, M. [et al.] (2008). «The array CGH and its clinical applications». *Drug Discov. Today*, 13 (17-18): 760-770.
- SMITH, L. M. [et al.] (1986). «Fluorescence detection in automated DNA sequence analysis». *Nature*, 321 (6071): 674-679.
- STIRZAKER, C. [et al.] (2014). «Mining cancer methylomes: prospects and challenges». *Trends Genet.*, 30 (2): 75-84.
- TOST, J. [et al.] (2003). «Analysis and quantification of multiple methylation variable positions in CpG islands by Pyrosequencing». *Biotechniques*, 35 (1): 152-156.
- VACHTENHEIM, J. [et al.] (1994). «Hypomethylation of CCGG sites in the 3' region of H-ras protooncogene is frequent and is associated with H-ras allele loss in non-small cell lung cancer». *Cancer Res.*, 54 (5): 1145-1148.
- VENTER, J. C. [et al.] (2001). «The sequence of the human genome». *Science*, 291 (5507): 1304-1351.
- WANG, Q. [et al.] (2013). «Tagmentation-based whole-genome bisulfite sequencing». *Nat. Protoc.*, 8 (10): 2022-2032.
- WELSH, J. [et al.] (1990). «Fingerprinting genomes using PCR with arbitrary primers». *Nucleic Acids Res.*, 18 (24): 7213-7218.
- WILLIAMS, J. G. [et al.] (1990). «DNA polymorphisms amplified by arbitrary primers are useful as genetic markers». *Nucleic Acids Res.*, 18 (22): 6531-6535.
- WILSON, G. G. (1991). «Organization of restriction-modification systems». *Nucleic Acids Res.*, 19 (10): 2539-2566.
- WINTER, E. [et al.] (1985). «A method to detect and characterize point mutations in transcribed genes: amplification and overexpression of the mutant c-Ki-ras allele in human tumor cells». *Proc. Natl. Acad. Sci. USA*, 82 (22): 7575-7579.
- WISZNIEWSKA, J. [et al.] (2014). «Combined array CGH plus SNP genome analyses in a single assay for optimized clinical testing». *Eur. J. Hum. Genet.*, 22 (1): 79-87.
- WITTWER, C. T. (2009). «High-resolution DNA melting analysis: advancements and limitations». *Hum. Mutat.*, 30 (6): 857-859.
- WITTWER, C. T. [et al.] (1997). «Continuous fluorescence monitoring of rapid cycle DNA amplification». *Biotechniques*, 22 (1): 130-131, 134-138.
- (2003). «High-resolution genotyping by amplicon melting analysis using LCGreen». *Clin. Chem.*, 49 (6 Pt 1): 853-860.
- WORM, J. [et al.] (2001). «In-tube DNA methylation profiling by fluorescence melting curve analysis». *Clin. Chem.*, 47 (7): 1183-1189.
- XIONG, Z. [et al.] (1997). «COBRA: a sensitive and quantitative DNA methylation assay». *Nucleic Acids Res.*, 25 (12): 2532-2534.
- YAMAMOTO, F. [et al.] (2004). «A DNA microarray-based methylation-sensitive (MS)-AFLP hybridization method for genetic and epigenetic analyses». *Mol. Genet. Genomics*, 271 (6): 678-686.
- YOO, C. B. [et al.] (2006). «Epigenetic therapy of cancer: past, present and future». *Nat. Rev. Drug Discov.*, 5 (1): 37-50.