

CÈL·LULES MÍNIMES: EVOLUCIÓ I DISSENY

ROSARIO GIL

*Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Departament de Genètica,
Universitat de València.*

Adreça per a la correspondència: Rosario Gil. Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València. Apartat oficial 22085. 46071 València.
Adreça electrònica: rosario.gil@uv.es.

RESUM

Tots els organismes vius coneguts estan formats per cèl·lules, i cadascuna d'aquestes cèl·lules emmagatzema la informació necessària per al funcionament en el seu genoma. La possibilitat d'estudiar i comparar els genomes complets de diversos organismes ens permet definir un conjunt d'informació genètica essencial per al manteniment d'una cèl·lula viva en unes condicions definides. En aquest article repassarem com l'estudi de genomes reduïts de manera natural ens permet apropar-nos a conèixer quins són els elements essencials per a la vida i com, a partir de la definició d'una cèl·lula mínima teòrica, sorgeix la possibilitat de dissenyar i fabricar noves cèl·lules semiartificials mitjançant enginyeria biològica. A més de possibles aplicacions biotecnològiques o biosanitàries, l'estudi d'aquestes noves cèl·lules mínimes ens pot ajudar a comprendre l'essència de la vida.

Paraules clau: genoma mínim, cèl·lula mínima, vida, reducció genòmica, cèl·lula semi-sintètica.

MINIMAL CELLS: EVOLUTION AND DESIGN

SUMMARY

All known living organisms are made of cells, and each cell stores all necessary information for its proper performance in its genome. The possibility to study and compare complete genomes from many organisms allows the characterization of a genetic core that is essential to keep a cell alive in a defined environment. In this article we analyze how the study of naturally reduced genomes has deepened our knowledge on the elements that are essential for life, and how the definition of a theoretical minimal cell can help us to design and engineer novel semiartificial cells. In additions to their potential for biotechnological

or biomedical applications, the study of such minimal cells can help us to understand the essence of life.

Key words: minimal genome, minimal cell, life, genome reduction, semisynthetic cell.

INTRODUCCIÓ: LA VIDA, LA CÈL·LULA, EL GENOMA

La biologia és la ciència que estudia la vida. Pot semblar, doncs, sorprenent, que els biòlegs s'hagin interessat més per descriure i catalogar els éssers vius que per aprofundir en els fonaments de la vida, buscar el seu origen i definir-la. No parlarem de l'origen de la vida, però sí de la seva essència. No sembla que aquest tema preocupés excessivament Darwin, convençut que no era necessari conèixer l'essència de la vida per a l'avenç de la biologia, de la mateixa manera que no era necessari conèixer l'essència de la gravitació universal per a l'avenç de la física (Peretó, 2008). Tot i això, a ningú no se li escapava que els organismes biològics són molt complexos en realitat. Podríem dir que la cèl·lula és l'estructura de dimensions micromètriques més complexa coneguda, i encara avui dia desconexim en gran mesura la manera com funciona: com s'estructura cadascuna de les seves parts per exercir la seva funció particular i com aquestes parts s'organitzen de manera harmonitzada per donar lloc a un organisme viu. Ja al segle xx, amb el desenvolupament de la biologia molecular, es va passar de la descripció cel·lular a la descripció molecular de la vida, i el desenvolupament de la genòmica en els darrers temps no ha fet més que reforçar aquesta visió. Els avenços tecnològics han permès l'acumulació de quantitats impressionants de dades sobre genomes complets, però el nostre coneixement sobre la vida quasi no avança. No sembla que la descripció molecular de la vida sigui prou per comprendre-la. En l'era postgenòmica, ens trobem limitats per la capacitat d'inter-

pretar aquesta gran quantitat de dades per arribar a conclusions rellevants que ens ajudin a comprendre què és la vida.

Sabem que tots els organismes vius coneguts estan formats per cèl·lules. Si acceptem que la informació bàsica per donar lloc a una cèl·lula viva es troba emmagatzemada en el seu genoma, podem tractar d'apropar-nos al coneixement íntim d'una cèl·lula a través de l'estudi del seu genoma, tractant d'esbrinar quins són els elements essencials presents en tots els genomes que permeten a la cèl·lula estar viva. Però fins i tot si analitzem només la informació genètica representada pels gens codificants de proteïnes, al voltant d'un terç dels gens presents en qualsevol genoma té una funció desconeguda (i el nombre s'incrementaria molt més si tenim en consideració també aquells gens per als quals només disposem d'informació poc definida sobre la seva funció), de manera que és molt difícil predir amb seguretat quines són les capacitats reals d'un organisme concret. El que sí que és cert, és que tal com acumulem dades sobre la composició dels genomes dels éssers vius, resulta més evident que totes les formes de vida sobre el nostre planeta tenen unes característiques comunes que ens poden apropar a una definició de la vida. Aquest nucli central d'informació bàsica per a la vida és el que podem anomenar *genoma mínim*.

Però abans d'endinsar-nos en la definició d'un genoma mínim, hem de tractar de definir què és la vida, trobar una definició que ens faciliti la comprensió del fenomen de la vida per apropar-nos al seu coneixement. Podem afirmar que la vida sorgeix de l'encaixament d'elements no vius de manera apropiada en l'espai i el temps, per donar

lloc a un sistema que presenta tres propietats fonamentals: homeòstasi, autoreproducció i evolució (Luisi *et al.*, 2002). En els darrers anys, juntament amb la preocupació per la definició de la vida, s'ha despertat també la febre per sintetitzar-la. Aquest és un dels principals objectius de l'anomenada *biologia sintètica*, que pretén d'aquesta manera comprendre millor la vida i conèixer com s'han d'integrar les diferents parts que conformen una cèl·lula perquè funcioni com una entitat organitzada. Per tant, si les instruccions es troben en el DNA, necessitem primer definir aquest genoma mínim que conté totes les instruccions necessàries per posar una cèl·lula a funcionar. Una cèl·lula amb tots els seus components coneguts serà més fàcil d'estudiar, de manipular i, fins i tot, podrà servir com a recipient on es puguin afegir altres funcions addicionals que ens interessin, amb vista a aconseguir cèl·lules modificades amb utilitats biotecnològiques, biosanitàries o bioenergètiques. Sembla ciència-ficció, però no està tan lluny com podríem pensar. Per comprendre millor la dimensió d'aquest paradigma científic, cal conèixer què és una cèl·lula mínima —natural, teòrica o artificial— abans d'estudiar per a què ens pot servir.

CÈL·LULES MÍNIMES NATURALS

En els nostres dies, disposem de sistemes molt fiables per estimar les dimensions dels genomes, ja sigui de manera aproximada a través de tècniques electroforètiques (Riethman *et al.*, 1997) o mitjançant la seqüenciació completa del genoma en estudi. En general, els genomes procarïotes són més petits que els dels organismes eucariotes. Aquesta circumstància, sumada a l'interès que presenten molts bacteris per la seva implicació en malalties o pel seu interès biotecnològic, ha fet que en poc més

de deu anys s'hagin seqüenciat completament 686 genomes bacterians i 53 genomes d'arqueus (data obtinguda el 17 de juliol de 2008 a la base de dades GOLD, <http://www.genomesonline.org/gold.cgi>). Les anàlisis de les dades disponibles indiquen que, normalment, les dimensions dels genomes bacterians es correlacionen perfectament amb el nombre de gens que contenen (Casjens, 1998), de manera que genomes més grans corresponen a organismes generalistes, capaçs fins i tot de certes formes de desenvolupament o amb capacitats metabòliques complexes (com és el cas de *Myxococcus xanthus*, amb 9,14 Mb, que presenta capacitat de formar agregats amb «motilitat social»; o *Burkholderia xenovorans* LB400 que, amb un genoma de 9,73 Mb de longitud, és capaç de degradar amb eficàcia els bifenils policlorats, compostos tòxics antigament emprats com a refrigerants i lubricants); els genomes més petits corresponen a organismes especialistes de dos tipus: simbiotes obligats (ja siguin paràsits o mutualistes) i organismes de vida lliure adaptats a entorns oligotròfics. Els genomes microbians més petits seqüenciats es troben indicats a la taula 1. Hi trobem tant organismes autotròfics com heterotròfics, i s'observa que tots aquests genomes estan altament compactats, sense que presentin elements repetitius ni redundància gènica. L'evolució reductiva d'aquests genomes està relacionada amb la seva adaptació als nínxols ecològics que ocupen aquests bacteris i la seva dinàmica poblacional. Tots els genomes presentats a la taula 1 comparteixen una sèrie de característiques que, en conjunt, s'anomenen *síndrome del genoma resident* (perquè inicialment es van descriure en bacteris que vivien a l'interior de cèl·lules eucariotes). La reducció genòmica està acompanyada d'un increment en el contingut en A + T, pèrdua de mecanismes de reparació i recombinació (excepte en el cas de l'arqueu *Nanoarchaeum*

TAULA 1. Genomes minimitzats naturals completament seqüenciats, ordenats d'acord amb la seva longitud. Dades obtingudes de GOLD

Organisme	Tipus	Estil de vida	Entorn	Metabolisme	Genoma, kb	Gens codificants	Contingut en G + C (%)
<i>Ca. Carsonella ruddii</i> ¹	γ -proteobacteri	Mutualista	<i>Pachypsylla venusta</i>	Heteròtrof	159	182	16,0
<i>Ca. Sulcia muelleri</i> ²	Bacteroidetes	Mutualista	<i>Homalodisca coagulata</i>	Heteròtrof	245	227	22,4
<i>Buchnera aphidicola</i> BCc ³	γ -proteobacteri	Mutualista	<i>Cinara cedri</i>	Heteròtrof	420	362	20,1
<i>Nanoarchaeum equitans</i>	Arqueu ⁴	Simbiònt ⁵	<i>Ignicoccus</i> sp.	Heteròtrof	490	536	31,6
<i>Mycoplasma genitalium</i>	Mollicutes	Paràsit	<i>Homo sapiens</i>	Heteròtrof	580	477	31,7
<i>Ca. Vesicomys okutanii</i>	γ -proteobacteri	Mutualista	<i>Calyptogenia okutanii</i>	Autòtrof	1.022	976	34,0
<i>Pelagibacter ubique</i>	α -proteobacteri	Lliure	Aigua marina	Autòtrof	1.308	937	31,6
<i>Prochlorococcus marinus</i>	Cianobacteri	Lliure	Aigua marina	Autòtrof	1.657	1.717	30,8

¹ El seu contingut gènic sembla incompatible amb la vida; potser s'apropa més a un nou tipus d'òrganul (Tammes *et al.*, 2007).

² Coexisteix amb *Baumannia cicadellincola* (γ -proteobacteri).

³ Coexisteix amb *Serratia symbiotica* (γ -proteobacteri).

⁴ Tot i que es va proposar que pertanyia a un nou fílum, Nanoarchaeota (Huber *et al.*, 2002), recentment s'ha considerat que representa un llinatge relacionat amb arqueus termococals amb acceleració evolutiva (Brochier *et al.*, 2005).

⁵ No està totalment aclarit si la relació amb l'hoste és mutualista, comensal o parasítica.

equitans, per al qual, a causa de la necessitat de sobreviure en un ambient hipertermofílic, aquests gens resulten essencials) i un increment en la taxa evolutiva dels gens preservats que, en alguns casos, es comporten com pseudogens (restes de gens que no donen lloc a una proteïna activa).

Els estudis filogenètics demostren que els simbionts obligats deriven d'avantpassats de vida lliure amb genomes més grans (Baumann, 2005; Brochier *et al.*, 2005; Kuwahara *et al.*, 2007). L'adaptació a la vida a l'interior de cèl·lules hostes fa disminuir la pressió selectiva sobre gens que no són necessaris en un ambient protegit o resulten redundants, ja que l'hoste pot proveir els dos membres de la simbiosi. L'anàlisi

comparativa dels genomes d'endosimbionts ha posat de manifest que existeix un nucli de gens essencial per al manteniment de la maquinària necessària per sobreviure en simbiosi obligada, i que les diferències observades deriven de l'adequació als requeriments específics de cada tipus de simbiosi, ja sigui parasítica o mutualista i, en aquest segon cas, de les necessitats nutricionals de l'hoste i la ubicació del bacteri en les cèl·lules especialitzades (Gil *et al.*, 2003).

Els genomes més reduïts estudiats pertanyen a endosimbionts d'insectes, i els casos més extrems són els representats per *Candidatus Carsonella ruddii* (Nakabachi *et al.*, 2006), *Candidatus Sulcia muelleri* (McCutcheon i Moran, 2007), i *Buchnera aph-*

dicola BCc (Perez-Brocal *et al.*, 2006). Molts insectes tenen dietes restringides, nutricionalment deficientes, i depenen d'aquests bacteris endosimbionts per a la síntesi d'aminoàcids essencials o vitamines per a la supervivència i l'adequada reproducció. *Candidatus Carsonella ruddii* (endosimbiont del psílid *Pachypsylla venusta*, que s'alimenta només de floema de lledoner) té un genoma de només 160 kb, molt compactat, amb pautes de lectura oberta superposades. Només cent vuitanta-dos gens codificants de proteïnes han estat proposats per a aquest genoma (Nakabachi *et al.*, 2006). Tot i ser el genoma més petit seqüenciat, no queda clar si *Ca. Carsonella ruddii* és realment un bacteri o ha transferit part dels seus gens al nucli del hoste, de manera que es trobaria a mitjan camí entre una cèl·lula i un orgànul. De fet, no tots els gens essencials implicats en els processos de replicació, transcripció i traducció han pogut ser identificats en aquest genoma, i alguns dels gens implicats en la síntesi d'aminoàcids essencials (fins i tot vies metabòliques completes) estan absents (Tamames *et al.*, 2007).

Candidatus Sulcia muelleri (McCutcheon i Moran 2007), endosimbiont de l'insecte *Homalodisca coagulata*, que s'alimenta de xilema de diferents plantes, presenta un genoma de 245 kb amb només 227 gens codificants de proteïnes. En aquest cas, hem de considerar que l'associació simbiòtica és en realitat un consorci, del qual també participa un segon bacteri, *Candidatus Baumannia cicadellinicola*, amb 595 ORF en un genoma de 686 kb (Wu *et al.*, 2006). L'estudi funcional dels gens presents en aquests dos genomes ha permès posar de manifest que *Ca. Sulcia muelleri* manté una relació més antiga amb l'hoste, i que la relació amb *Ca. Baumannia cicadellinicola* es va produir en el moment del canvi d'alimentació de l'avantpassat de l'insecte, d'una dieta a base de floema a una altra composta per xile-

ma (Takiya *et al.*, 2006). L'evolució conjunta d'aquest consorci ha dut a una situació en la qual existeix una perfecta complementarietat en el subministrament d'aminoàcids i vitamines per part dels dos endosimbionts.

En el cas de *B. aphidicola* BCc (Perez-Brocal *et al.*, 2006), endosimbiont del pugó del cedre, també s'ha establert un consorci amb un altre bacteri, *Candidatus Serratia symbiotica*, per a la provisió de nutrients essencials a l'hostatger. Tot i que aquest genoma és més gran que els anteriorment descrits, cal destacar que hi ha fins a quatre genomes de diferents soques de *B. aphidicola* (procedents de diferents espècies de pugons) completament seqüenciats, i que la soca BCc presenta una reducció de fins a 200 kb respecte als altres tres genomes. A més, *Ca. Serratia symbiotica* està considerada com un simbiònt facultatiu en les altres espècies de pugons, però sembla un endosimbiont obligat de *Cinara cedri*, el pugó del cedre. El genoma de *B. aphidicola* BCc, amb 420 kb i 362 gens codificants de proteïnes, en ser comparat amb el dels altres genomes de *B. aphidicola* ens ha permès aprofundir en l'anàlisi de la reducció dels genomes d'endosimbionts en fases avançades del procés i en les seves conseqüències evolutives i ecològiques.

També els genomes dels paràsits obligats pateixen el síndrome del genoma resident. El cas més notable conegut és el de *Mycoplasma genitalium* (Fraser *et al.*, 1995), paràsit epicel·lular humà causant d'infeccions de transmissió sexual. El seu genoma de 580 kb va ser el segon genoma bacterià seqüenciat i encara és el genoma més petit conegut d'un organisme que pot ser cultivat en el laboratori.

Altres casos estudiats amb genomes no tan reduïts (tot i que encara pel voltant d'1 Mb) corresponen a bacteris autòtrofs mutualistes obligats, capaços d'aportar quasi tots els nutrients necessaris al seu hoste. En aquesta circumstància trobem

Candidatus Ruthia magnifica (Newton *et al.*, 2007) i *Candidatus Vesicomysocius okutanii* (Kuwahara *et al.*, 2007), endosimbionts de molluscs del gènere *Calyptogena*, que habiten les profunditats marines. Aquests bacteris són capaços de fixar CO₂ emprant electrons procedents del H₂S i tenen el potencial de sintetitzar pràcticament totes les biomolècules essencials per als seus hosts. A canvi, les cèl·lules especialitzades del mollusc que contenen els bacteris (bacteriòcits) estan situades entre les cèl·lules epitelials branquials i el sistema circulatori, i eviten que el H₂S i el nitrogen o oxigen reaccionin espontàniament, per tal de poder-los emprar com a donants i receptors d'electrons, respectivament, una situació que no podria donar-se de manera simultània en una cèl·lula de vida lliure.

També en organismes de vida lliure pot donar-se un fenomen de minimització genòmica similar al descrit en endosimbionts. És el cas de bacteris que viuen de manera independent formant grans poblacions en ambients oligotròfics, en què el procés selectiu d'adaptació a un ambient pobre en nutrients afavoreix l'aparició d'organismes amb una maquinària cel·lular reduïda i un augment de la relació superfície-volum per tal d'economitzar recursos. Bacteris fotòtrofs marins com *Prochlorococcus marinus* (cianobacteri) i *Pelagibacter ubique* (α -proteobacteri) es troben entre els organismes més abundants al nostre planeta i presenten genomes del voltant d'1 Mb, similars, per tant, als de bacteris endosimbionts autòtrofs. Diversos genomes de diferents ecotipus de *Prochlorococcus* han estat seqüenciats (Dufresne *et al.*, 2003; Hess, 2004; Rocop *et al.*, 2003), i mostren una gran variabilitat, fet que indica una gran capacitat per modificar-se en funció de l'ambient. L'ecotip MED4 de *P. marinus*, adaptat a l'elevada intensitat lumínica de la superfície dels oceans, presenta el genoma més petit descrit

per a aquesta espècie autòtrofa (1,66 Mb). Però el genoma més petit conegut per a un ésser viu capaç de viure independentment correspon a *P. ubique*, un bacteri fotoheterotròfic pertanyent a l'abundant clade SAR11 (Morris *et al.*, 2002), amb només 1,308 Mb i un conjunt de 1.354 gens codificants de proteïnes (Giovannoni *et al.*, 2005).

L'existència d'aquests genomes reduïts en la natura ens planteja una qüestió: si els organismes que els posseeixen han pogut prescindir de part dels genomes ancestrals, és que aquestes parts no eren essencials per a la vida; com podria ser, doncs, una cèl·lula mínima? Quin seria el genoma mínim, és a dir, el conjunt mínim de gens capaç de sustentar una forma de vida en condicions controlades (absència d'estress i medi nutricionalment ric)? (Koonin, 2000). La resposta a aquesta pregunta que, en definitiva, ens apropa un poc més a conèixer l'essència de la vida, s'ha abordat de diferents maneres, que passem a analitzar a continuació.

CÈL·LULES MÍNIMES TEÒRIQUES

Els espectaculars avenços de la genòmica en els darrers temps estan permetent que cada dia coneguem més i més genomes complets d'organismes. Podem aprofundir en el coneixement de quins gens són essencials per al sosteniment d'una cèl·lula viva per comparació del genomes coneguts, ja que les funcions essencials han d'estar representades en tots els organismes. Aquesta aproximació comparada és especialment útil en bacteris, no sols a causa de l'elevat nombre de genomes disponibles, sinó també al fet que la majoria de les proteïnes bacterianes s'han mantingut suficientment conservades al llarg de l'evolució per permetre predir la funció d'una proteïna desconeguda codificada en un genoma d'un bacteri poc caracteritzat d'acord amb la similitud

amb proteïnes perfectament caracteritzades procedents d'organismes model ben estudiats genèticament i bioquímicament. Tot i això, una vegada identificada la funció teòrica de totes les proteïnes codificades en un genoma, s'haurien de procedir a identificar no sols els gens comuns als diferents genomes comparats, sinó també quines són les funcions comunes, ja que és possible que proteïnes diferents, amb orígens diferents (proteïnes no ortòlogues), tinguin una mateixa funció, i són realment les funcions, i no les proteïnes concretes, les que són imprescindibles per al funcionament correcte de la cèl·lula. Finalment, l'única forma d'esbrinar si realment un gen (i, per tant, la funció que exerceix el seu producte proteic) és essencial, seria comprovant-ho experimentalment: l'eliminació, interrupció o silenciament d'un gen essencial serà incompatible amb la vida. D'acord amb aquests principis, hi ha dues maneres bàsiques d'aproximar-nos al concepte de cèl·lula mínima (poseïdora d'un genoma mínim): de manera computacional i experimental. La combinació d'ambdues ha estat molt útil en la definició d'un hipotètic genoma mínim teòric.

La primera aproximació al genoma mínim amb genòmica comparada va ser realitzada quan van ser seqüenciats els dos primers genomes bacterians (Mushegian i Koonin, 1996). Els genomes corresponien a *Haemophilus influenzae* i *Mycoplasma genitalium*, dos patògens humans amb genomes naturalment reduïts, a causa de la seva forma de vida parasítica, i molt allunyats filogenèticament. La hipòtesi de partida era que els gens conservats entre dues formes de vida que han patit una evolució reductiva independent haurien de ser essencials i totes les funcions essencials devien estar compartides. El resultat va ser la proposta d'un conjunt mínim format per 256 gens, la majoria dels quals estan implicats en l'emmagatzemament i processament de la informació

genètica, el processament i manteniment de proteïnes funcionals i unes capacitats metabòliques molt limitades. Moltes altres anàlisis comparatives s'han fet des d'aquell moment (revisades per Carbone, 2006), però el sistema té algunes limitacions: resulta complicat reconèixer els veritables casos de desplaçaments no ortòlegs (proteïnes no ortòlogues amb la mateixa funció) i, com que no disposem d'informació de quasi un terç dels gens representats en un genoma, fins i tot el conjunt mínim proposat conté gens de funció desconeguda. D'altra banda, la majoria dels conjunts mínims proposats d'acord amb estudis comparatius està principalment format per gens implicats en funcions per a les quals no hi ha alternativa en la natura (un bon percentatge dels quals està implicat en la complexa maquinària de traducció, incloent-hi els ribosomes), però contenen xarxes metabòliques deficientes que no serien compatibles amb el manteniment d'una cèl·lula viva. Això és degut al fet que hi ha molts metabolismes possibles, depenent de les disponibilitats d'energia i nutrients de l'entorn, de manera que cap de les possibles formes de metabolisme està completament conservada.

Una aproximació alternativa amb base genètica consisteix a identificar quins són els gens essencials per al manteniment d'una cèl·lula viva de manera experimental. Estudis d'inactivació de gens a gran escala s'han realitzat amb diferents organismes model, emprant diferents estratègies que van des de la inactivació sistemàtica de gens, l'ús massiu d'elements transposables o la utilització de RNA antisentit per inhibir l'expressió gènica (revisats a Feher *et al.*, 2007; Reznikoff i Winterberg, 2008). Però, una vegada més, existeixen limitacions. A banda de falsos positius (gens considerats essencials que només retarden el creixement sense impedir-lo) i negatius (gens essencials que poden permetre interrupcions en zones no

vitals per a la proteïna codificada o poden funcionar amb una expressió residual), una de les principals limitacions és que la inactivació d'un gen individual no permet identificar funcions essencials codificades per gens redundants, i alguns gens que poden ser eliminats individualment poden no ser simultàniament dispensables. De manera que, en general, el conjunt de gens essencials determinats d'aquesta manera sol ser només un subconjunt dels gens que hauria de contenir un genoma mínim.

És possible apropar-nos un poc més al contingut gènic d'un hipotètic genoma mínim si combinem totes les estratègies anteriors i tractem de donar un sentit coherent al metabolisme mínim que s'inferiria de l'essentat genoma. Aquesta va ser l'aproximació emprada pel nostre grup (Gil *et al.*, 2004) per proposar un genoma mínim teòric compost per dos-cents vuit gens, seixanta-dos dels quals estan implicats en el metabolisme celular formant una xarxa metabòlica amb unes propietats topològiques i estequiòmètriques coherents (Gabaldon *et al.*, 2007). En qualsevol cas, sembla segur que es podrien proposar diferents genomes mínims basats en les funcions definides com a essencials per a una cèl·lula moderna, en la qual dos terços del genoma estarien dedicats a processos informàtics (replicació, transcripció, traducció i manteniment de proteïnes funcionals) i serien constants, amb un terç de gens implicats en el metabolisme, variables en funció de les condicions de l'entorn.

Una altra via possible per definir el contingut gènic d'un hipotètic genoma mínim seria partir de la definició prèvia de les funcions essencials, buscant el mínim completament bioquímic capaç de realitzar-les, encara que el conjunt definit així no s'hagi trobat mai en una cèl·lula moderna coneguda (Forster i Church, 2006). Aquest sistema «modular» per determinar el contingut d'un genoma mínim té l'avantatge que cap

dels gens considerats com a integrants de l'hipotètic genoma mínim tindria una funció desconeguda, de manera que totes les capacitats del sistema serien conegudes *a priori*. D'altra banda, cadascun dels mòduls funcionals pot haver estat analitzat prèviament per comprovar el seu funcionament independent, de manera que només faltaria comprovar que el conjunt és capaç de combinar-se per treballar com una entitat organitzada. L'aproximació bioquímica de Foster i Church (2006) els portà a definir un genoma mínim compost per només cent cinquanta-un gens. Cal considerar que els autors assumeixen que tots els nutrients i l'energia han de ser aportats per l'entorn (i, per tant, no cal incloure gens implicats en el metabolisme), i que el sistema pot estar contingut en una vesícula lipídica generada espontàniament en un entorn amb molècules amfílites. Tot i que aquestes assumpcions estan lluny de ser demostrades, i que queden altres reptes per resoldre abans que siguem capaços de sintetitzar un sistema d'aquestes característiques, el pla per dur el projecte a la pràctica (Forster i Church, 2007) ens deixa molt a prop de la síntesi d'una veritable cèl·lula viva artificial.

CÈL·LULES MÍNIMES (QUASI) SINTÈTIQUES

Sembla, per tant, que la síntesi d'un sistema dotat de les propietats que li confereixen la propietat d'estar viu no és un objectiu impossible. A més, un sistema dotat d'un genoma mínim en el qual tots els components tinguin una funció coneguda i essencial podria ser la millor manera d'apropar-nos al coneixement de les característiques d'aquest sistema viu, capaç d'automantenir-se, reproduir-se i evolucionar.

A banda del sistema dissenyat per Foster i Church (2006), que ens portaria a una

forma de vida completament diferent de les que ja coneixem (tot i que inspirada en aquesta), hi ha dues possibles aproximacions a un genoma mínim mitjançant la manipulació de cèl·lules ja existents: eliminant les parts no essencials de genomes més grans o introduint un genoma sintètic en una cèl·lula a la qual prèviament se li ha eliminat el genoma propi.

La minimització de genomes naturals ha estat una estratègia molt emprada en els darrers anys. Els genomes d'organismes models com *Escherichia coli*, *Bacillus subtilis* o *M. genitalium* (que ja abans de la reducció addicional presenta el genoma més petit conegut per a un bacteri amb capacitat de vida autònoma) o el d'organismes d'interès biotecnològic com *Corynebacterium glutamicum*, han estat sotmesos a diferents experiments orientats a eliminar aquelles parts del genoma que no semblen essencials (revisats a Feher *et al.*, 2007). Per aconseguir aquest objectiu, és convenient determinar primer quines són les regions que poden ser prescindibles (com ara, aquelles que són producte de transferència horitzontal, bacteriòfags, plasmidis conjugatius o elements transposables). Una vegada identificades, existeixen diferents mètodes que ens permeten l'eliminació de les regions seleccionades, mitjançant l'ús de plasmidis o DNA lineal, combinats amb la utilització de recombinases específiques de lloc o transposons. Seguint aquests procediments, ha estat possible reduir el genoma d'*E. coli* entre un 15,3 % i un 21 % per aconseguir noves soques que presenten, fins i tot, algunes propietats millorades (Mizoguchi *et al.*, 2007; Posfai *et al.*, 2006). En situacions més extremes, quan el genoma d'*E. coli* es va reduir en quasi un 30 % (Hashimoto *et al.*, 2005), van aparèixer efectes negatius, com són una morfologia cel·lular aberrant o la reducció de la taxa de creixement, tot i que la nova soca amb el genoma

així reduït s'havia obtingut per combinació de delecions en sèrie de regions prèviament eliminades sense efectes adversos. És a dir, superat un cert límit (que no està ben definit) l'eliminació de regions genòmiques teòricament prescindibles resulta nociva per a la cèl·lula. Potser l'arquitectura del cromosoma, l'existència d'una certa forma de connexió encara no ben coneguda entre les diferents parts del genoma d'una cèl·lula real, estigui implicada en aquesta dificultat per eliminar de manera seriada fragments cada vegada més llargs de genoma.

Una altra possibilitat per arribar a aconseguir un genoma mínim seria sintetitzar-lo *de novo* i introduir-lo després a l'interior d'un compartiment cel·lular ja existent. El que fa uns anys hagués resultat impensable, és avui dia una tasca abordable. La tecnologia disponible actualment per a la síntesi de llargs fragments de DNA és suficient per permetre la síntesi de genomes petits amb suficient fiabilitat i amb un preu cada vegada més assequible. De fet, s'ha passat en pocs anys de la capacitat de sintetitzar un genoma víric viable (Cello *et al.*, 2002) a la síntesi de mòduls genètics complets (Kodumal *et al.*, 2004) i, a principis de 2008, a la síntesi completa del genoma de *M. genitalium* (Gibson *et al.*, 2008). Introduir un genoma de DNA nu dins una cèl·lula també ha estat una tasca aconseguida recentment pel grup de C. J. Venter (Lartigue *et al.*, 2007), una vegada més aprofitant les peculiars característiques dels micoplasmes, amb genomes reduïts i sense paret cel·lular. Tot i això, no queda clar com el genoma original de la cèl·lula emprada com a receptora va deixar pas al nou genoma trasplantat, de manera que no és possible saber si l'experiment funcionaria en altres circumstàncies, amb un altre tipus de bacteri receptor o amb un DNA sintètic.

COMENTARI FINAL: ENGINYERIA BIOLÒGICA AMB CÈL·LULES MÍNIMES

La possibilitat de sintetitzar cèl·lules mínimes dotades de vida no sembla un objectiu llunyà. Aquestes cèl·lules ens ajudaran sens dubte a conèixer amb més profunditat l'essència de la vida, a través de l'estudi del fenomen d'una manera simplificada i controlada, per adquirir un millor coneixement de com funcionen els sistemes cellulars integrant cadascun dels seus components. Però podem (i sens dubte és un dels reptes de la biologia del futur) anar un pas més enllà. Podem emprar aquestes noves cèl·lules simplifiades com a contenidors per fer enginyeria biològica i dotar el sistema de les propietats que ens interessin (Andrianantoandro *et al.*, 2006; Drubin *et al.*, 2007; Isalan *et al.*, 2008; Serrano, 2007). Podem fabricar cèl·lules que tinguin aplicacions en el camp de la bioremediació, la biomedicina o la fabricació de biocombustibles més ecològics i sense dependre de fonts esgotables, per exemple. Podem crear noves formes de vida. Fins i tot podem modificar el codi genètic per poder introduir nous aminoàcids en les proteïnes, i dotar-les de noves propietats (Drubin *et al.*, 2007). Però potser tot això forma part d'una altra història que encara estem començant a escriure. Una història que, sens dubte, hauria semblat de ciència-ficció al nostre benvolgut Charles Darwin.

AGRAÏMENTS

Vull donar les gràcies a A. Navarro per donar-me l'oportunitat de participar en aquest projecte d'homenatge a Darwin; a J. Peretó, A. Latorre i A. Moya per ajudar-me a posar en ordre les meves idees sobre el tema i per la fantàstica atmosfera de treball i discussió científica que es respira al

Grup de Genètica Evolutiva de l'ICBiBE. El treball ha estat finançat amb projectes del MEC (BFU2005-03477/BMC) i la Generalitat Valenciana (GV/2007/050).

BIBLIOGRAFIA

- ANDRIANANTOANDRO, E.; BASU, S.; KARIG, D. K.; WEISS, R. (2006). «Synthetic biology: New engineering rules for an emerging discipline». *Mol. Syst. Biol.*, 2: 2006-2008.
- BAUMANN, P. (2005). «Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects». *Annu. Rev. Microbiol.*, 59: 155-189.
- BROCHIER, C.; GRIBALDO, S.; ZIVANOVIC, Y.; CONFALONIERI, F.; FORTERRE, P. (2005). «Nanoarchaea: Representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to thermococcales?». *Genome Biol.*, 6: R42.
- CARBONE, A., (2006). «Computational prediction of genomic functional cores specific to different microbes». *J. Mol. Evol.*, 63: 733-746.
- CASJENS, S., (1998). «The diverse and dynamic structure of bacterial genomes». *Annu. Rev. Genet.*, 32: 339-377.
- CELLO, J.; PAUL, A. V.; WIMMER, E. (2002). «Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template». *Science*, 297: 1016-1018.
- DRUBIN, D. A.; WAY, J. C.; SILVER, P. A. (2007). «Designing biological systems». *Genes Dev.*, 21: 242-254.
- DUFRESNE, A.; SALANOUBAT, M.; PARTENSKY, F. [et al.] (2003). «Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome». *Proc. Natl. Acad. Sci. USA*, 100: 10020-10025.
- FEHER, T.; PAPP, B.; PAL, C.; POSFAL, G. (2007). «Systematic genome reductions: Theoretical and experimental approaches». *Chem. Rev.*, 107: 3498-3513.
- FORSTER, A. C.; CHURCH, G. M. (2006). «Towards synthesis of a minimal cell». *Mol. Syst. Biol.*, 2: 45.
- (2007). «Synthetic biology projects in vitro». *Genome Res.*, 17: 1-6.
- FRASER, C. M.; GOCAYNE, J. D.; WHITE, O. [et al.] (1995). «The minimal gene complement of *Mycoplasma genitalium*». *Science*, 270: 397-403.
- GABALDON, T.; PERETO, J.; MONTERO, F. [et al.] (2007). «Structural analyses of a hypothetical minimal metabolism». *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 362: 1751-1762.
- GIBSON, D. G.; BENDERS, G. A.; ANDREWS-PFANNKUCH,

- C. [et al.] (2008). «Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome». *Science*, 319: 1215-1220.
- GIL, R.; SILVA, F. J.; PERETO, J.; MOYA, A. (2004). «Determination of the core of a minimal bacterial gene set». *Microbiol. Mol. Biol. Rev.*, 68: 518-537.
- GIL, R.; SILVA, F. J.; ZIENTZ, E. [et al.] (2003). «The genome sequence of *Blochmannia floridanus*: Comparative analysis of reduced genomes». *Proc. Natl. Acad. Sci. USA*, 100: 9388-9393.
- GIOVANNONI, S. J.; TRIPP, H. J.; GIVAN, S. [et al.] (2005). «Genome streamlining in a cosmopolitan oceanic bacterium». *Science*, 309: 1242-1245.
- HASHIMOTO, M.; ICHIMURA, T.; MIZOGUCHI, H. [et al.] (2005). «Cell size and nucleoid organization of engineered *Escherichia coli* cells with a reduced genome». *Mol. Microbiol.*, 55: 137-149.
- HESS, W. R. (2004). «Genome analysis of marine photosynthetic microbes and their global role». *Curr. Opin. Biotechnol.*, 15: 191-198.
- HUBER, H.; HOHN, M. J.; RACHEL, R. [et al.] (2002). «A new phylum of archaea represented by a nano-sized hyperthermophilic symbiont». *Nature*, 417: 63-67.
- ISALAN, M.; LEMERLE, C.; MICHALODIMITRAKIS, K. [et al.] (2008). «Evolvability and hierarchy in rewired bacterial gene networks». *Nature*, 452: 840-845.
- KODUMAL, S. J.; PATEL, K. G.; REID, R. [et al.] (2004). «Total synthesis of long DNA sequences: Synthesis of a contiguous 32-kb polyketide synthase gene cluster». *Proc. Natl. Acad. Sci. USA*, 101: 15573-15578.
- KOONIN, E. V. (2000). «How many genes can make a cell: The minimal-gene-set concept». *Annu. Rev. Genomics Hum. Genet.*, 1: 99-116.
- KUWAHARA, H.; YOSHIDA, T.; TAKAKI, Y. [et al.] (2007). «Reduced genome of the thioautotrophic intracellular symbiont in a deep-sea clam, *Calyptogena okutanii*». *Curr. Biol.*, 17: 881-886.
- LARTIGUE, C.; GLASS, J. I.; ALPEROVICH, N. [et al.] (2007). «Genome transplantation in bacteria: Changing one species to another». *Science*, 317: 632-638.
- LUISI, P. L.; OBERHOLZER, T.; LAZCANO, A. (2002). «The notion of a DNA minimal cell: a general discourse and some guidelines for an experimental approach». *Helv. Chim. Acta*, 85: 1759-1777.
- MCCUTCHEON, J. P.; MORAN, N. A. (2007). «Parallel genomic evolution and metabolic interdependence in an ancient symbiosis». *Proc. Natl. Acad. Sci. USA*, 104: 19392-19397.
- MIZOGUCHI, H.; MORI, H.; FUJIO, T. (2007). «*Escherichia coli* minimum genome factory». *Biotechnol. Appl. Biochem.*, 46: 157-167.
- MORRIS, R. M.; RAPPE, M. S.; CONNON, S. A. [et al.] (2002). «SAR11 clade dominates ocean surface bacterioplankton communities». *Nature*, 420: 806-810.
- MUSHEGIAN, A. R.; KOONIN, E. V. (1996). «A minimal gene set for cellular life derived by comparison of complete bacterial genomes». *Proc. Natl. Acad. Sci. USA*, 93: 10268-10273.
- NAKABACHI, A.; YAMASHITA, A.; TOH, H. [et al.] (2006). «The 160-kilobase genome of the bacterial endosymbiont *Carsonella*». *Science*, 314: 267.
- NEWTON, I. L.; WOYKE, T.; AUCHTUNG, T. A. [et al.] (2007). «The *Calyptogena magnifica* chemoautotrophic symbiont genome». *Science*, 315: 998-1000.
- PERETÓ, J. (2008). «Sobre la naturalesa y la fabricación de seres vivos» *Apuntes de Cyt*, 27: 30-38.
- PEREZ-BROCAL, V.; GIL, R.; RAMOS, S. [et al.] (2006). «A small microbial genome: The end of a long symbiotic relationship?». *Science*, 314: 312-313.
- POSFAL, G.; PLUNKETT, G. III; FEHER, T. [et al.] (2006). «Emergent properties of reduced-genome *Escherichia coli*». *Science*, 312: 1044-1046.
- REZNIKOFF, W. S.; WINTERBERG, K. M., (2008) «Transposon-based strategies for the identification of essential bacterial genes». A: OSTERMAN, A. L.; GERDES, S. Y. [ed.]. *Microbial gene essentiality: Protocols and bioinformatics*. Totowa: Humana Press, 13-26.
- RIETHMAN, H.; BIRREN, B.; GNIKKE, A. (1997). «Preparation, manipulation, and mapping of high molecular weight DNA». A: BIRREN, B.; GREEN, E. D.; KLAPHOLZ, S.; MYERS, R. M.; ROSKAM, J. [ed.]. *Genome analysis: A laboratory manual. Vol. 1. Analyzing DNA*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 83-248.
- ROCAP, G.; LARIMER, F. W.; LAMERDIN, J. [et al.] (2003). «Genome divergence in two prochlorococcus ecotypes reflects oceanic niche differentiation». *Nature*, 424: 1042-1047.
- SERRANO, L. (2007). «Synthetic biology: Promises and challenges». *Mol. Syst. Biol.*, 3: 158.
- TAKIYA, D. M.; TRAN, P. L.; DIETRICH, C. H.; MORAN, N. A. (2006). «Co-cladogenesis spanning three phyla: Leafhoppers (insecta, Hemiptera, Cicadellidae) and their dual bacterial symbionts». *Mol. Ecol.*, 15: 4175-4191.
- TAMAMES, J.; GIL, R.; LATORRE, A. [et al.] (2007). «The frontier between cell and organelle: Genome analysis of *Candidatus Carsonella ruddii*». *BMC Evol. Biol.*, 7: 181.
- WU, D.; DAUGHERTY, S. C.; AKEN, S. E. VAN [et al.] (2006). «Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters». *PLoS Biol.*, 4: e18.