

Estudiant els virus de RNA: equacions dins el tub d'assaig?

Els sistemes biològics presenten una enorme complexitat. Diferents tipus d'interaccions i de processos en governen les dinàmiques, des dels ecosistemes fins als processos de regulació cel·lular o la infecció vírica d'una fulla. Malgrat la complexitat observada a diferents escales en els sistemes biològics, que presumiblement ens farien pensar en la impossibilitat d'entendre'ls i de caracteritzar-los, l'estudi de la complexitat biològica es pot abordar des de diferents disciplines científiques. D'una manera general, la idea que, per estudiar una dinàmica simple, necessitem un model simple és òbvia. Però, per estudiar una dinàmica complexa, sempre necessitem un model complex? Un model amb una multitud de variables i de paràmetres? La resposta és «no». Un dels grans canvis de paradigma del segle passat, alimentat per la teoria dels sistemes dinàmics no lineals i la física teòrica, és demolidor: podem estudiar els sistemes complexos amb models simples.

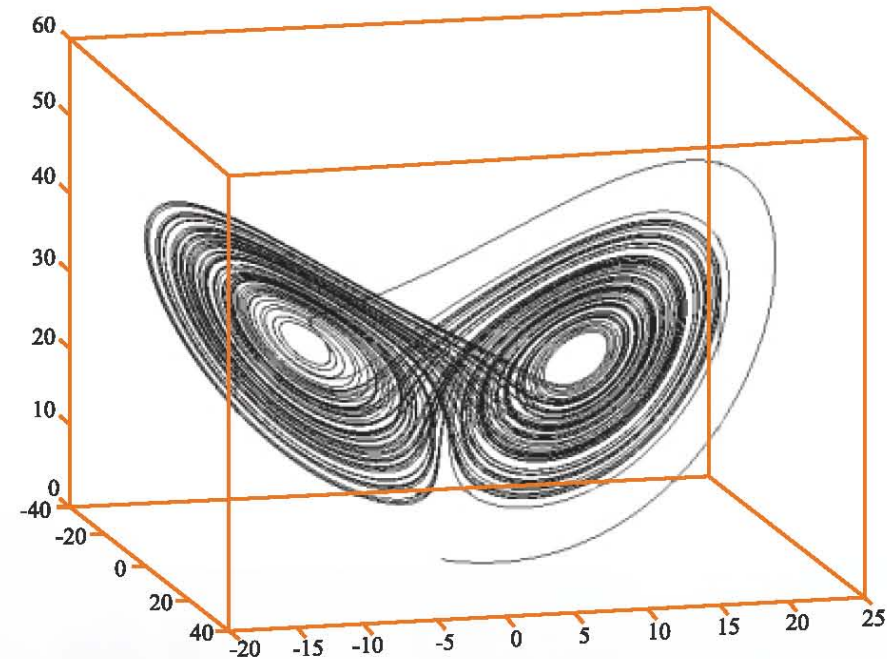
Escrit per

Josep Sardanyés i Cayuela

Centre de recerca: Institut de Biologia Molecular i Cel·lular de Plantes (IBMCP), Consell Superior d'Investigacions Científiques (CSIC) - Universitat Politècnica de València (UPV)

INICIADA al segle XVII gràcies als treballs d'Isaac Newton i de Gottfried Leibniz en càlcul integral i diferencial, la teoria dels sistemes dinàmics ha estat àmpliament utilitzada en diferents disciplines científiques, especialment a partir del segle XIX. Alhora, la visió geomètrica (o topològica) dels sistemes dinàmics desenvolupada per Henri Poincaré ha esdevingut una eina àmpliament utilitzada en diversos camps. Un dels grans paradigmes científics del segle passat, catalitzat per la teoria dels sistemes complexos, és que dinàmiques molt complexes (com les que observem en la natura, en els patrons climàtics o en els circuits electrònics) es poden estudiar mitjançant models «simples». El lector s'ha d'adonar que la

paraula «simple» està escrita entre cometes. Això vol dir que, encara que un sistema dinàmic tingui poques variables i pocs paràmetres, les no-linealitats poden fer que aquest model «simple» exhibeixi una complexitat enorme. Un dels millors exemples d'aquest fenomen s'esdevé en l'anomenat *sistema de Lorenz*. Nascut el 1917 a West Hartford (Connecticut, EUA), Edward N. Lorenz va construir un model matemàtic dinàmic molt abstracte que descrivia el fenomen de la convecció. Lorenz es va adonar que, en el seu model, els patrons temporals no sempre canviaven com ell predeïa, de manera que canvis infinitesimals en els valors inicials de les variables resultaven en patrons climàtics altament divergents al llarg del temps (la im-



◀ **Figura 1.** L'atractor de Lorenz, un dels atractors caòtics més famosos i estudiats, té forma de papallona. La dinàmica d'aquest objecte és molt complexa. Una de les propietats més importants dels sistemes dinàmics caòtics és la sensibilitat a les condicions inicials.

predictibilitat a llarg termini és una propietat del caos). La visualització de la dinàmica del model de Lorenz en l'espai de fases mostrava un objecte en forma de papallona (fig. 1). Aquesta troballa esdevindria el punt de partida definitiu de la teoria del caos, tot i que, prèviament, científics com Poincaré o Birkhoff ja havien tractat i estudiat alguns sistemes dinàmics caòtics.

L'anomenada *teoria de les quasiespècies moleculars*, originàriament desenvolupada en el camp de l'evolució prebiòtica, ha estat adoptada pels viròlegs com a eina matemàtica per a l'estudi de la dinàmica dels virus de RNA (grup que inclou el virus de l'hepatitis C o el VIH-1 i molts grups de virus de plantes, com ara el del torrat de la tomaca i els del mosaic del cogombre, de la carabassa i del tabac). L'àcid nucleic d'aquests virus és l'RNA d'una sola cadena, però també hi ha virus que tenen dues cadenes antiparal·leles i complementàries de RNA. Els virus de RNA monocatenaris es poden classificar segons el sentit, o polaritat, de la molècula. Així doncs, trobem virus de RNA de cadena positiva i negativa. El material genètic del primer tipus de virus és molt semblant als mRNA cel·lulars. Quan un d'aquests virus infecta una cèl·lula, aquesta pot confondre la cadena vírica amb un mRNA propi, traduir-lo i sintetitzar les proteïnes que el virus necessita per completar el seu cicle de reproducció i d'infecció. Les proteïnes virals es poden diferenciar en proteïnes no estructurals (encarregades de replicar l'RNA viral) o estructurals (encarregades d'empaque-

tar els genomes que infectaran altres cèl·lules). La replicació del virus, la realitza l'enzim RdRP (RNA-polimerasa dependent de RNA) que, en no tenir mecanismes de correcció d'errors, presenta una taxa de mutació molt elevada. Aquest fenomen fa que els virus siguin molt resistents als canvis ambientals i als fàrmacs, ja que evolucionen ràpidament i poden trobar solucions ràpides per escapar-se de les pressions selectives.

Les entitats replicatives amb una taxa de replicació baixa, quan competeixen amb entitats amb una taxa replicativa més alta, acaben essent desplaçades, ja que la seva eficàcia, o *fitness*, és més baixa. En el cas dels virus, l'eficàcia depèn de la funcionalitat de les proteïnes que codifica. Com que la seqüència mestra —el genoma més eficaç— coexisteix amb les seqüències mutants i ambdues classes genòmiques contribueixen a la producció de proteïnes, cal tenir en compte que les replicases sintetitzades per la seqüència mestra podrien ésser utilitzades per a la replicació de les seqüències mutants. Així doncs, el genotip de les seqüències mutants no en reflectiria l'eficàcia. Aquest fenomen, anomenat *complementació* i descrit experimentalment, es pot donar tant en les replicases com en les proteïnes encarregades d'encapsidar els genomes i formar nous virions. En aquest article presentarem uns quants resultats sobre els efectes de la complementació en la persistència de les seqüències mestres en un sistema simple format per genomes mestres, mutants i les replicases que sintetitzen.



Mitjançant un sistema dinàmic no lineal, analitzarem l'impacte de la taxa de mutació sobre les poblacions dels dos tipus de genomes que competeixen. Després, explicarem detalladament el significat del model, comentarem alguns dels resultats més rellevants i, finalment, presentarem algunes conclusions.

El model matemàtic

Per construir un model, primer definirem les variables (elements que canviaran amb el pas del temps) i certs paràmetres constants (no depenen del temps) que ens determinaran, en el nostre cas, les taxes de replicació i de degradació dels genomes virals. Per simplificar el sistema que volem modelitzar, només tindrem en compte la dinàmica dels genomes virals de RNA i de les proteïnes responsables de la replicació que codifiquen. Dividirem les poblacions de genomes en dues categories: la seqüència mestra (variable x_0) i una seqüència mitjana que inclourà tots els genomes mutants (x_1). A diferència de les seqüències mestres, les mutants tindran una eficàcia més baixa, ja que assumirem que les mutacions redueixen l'eficàcia. Per incloure la complementació, també hem de modelitzar la producció de proteïnes (replicases dependents de RNA), responsables de replicar els genomes virals. Assumim que les dues classes de seqüències són de polaritat positiva i, per tant, prescindim dels intermediaris de doble cadena que se sintetitzen durant la còpia de les cadenes. Això ens permet de considerar que cada tipus de genoma (x_0 i x_1) és traduït per la cèl·lula hoste i se sintetitzen les replicases mestres (p_0) i les replicases mutants (p_1) (fig. 2). El sistema dinàmic consta de les següents equacions diferencials no lineals (a l'equació 3 el subíndex $i = 0, 1$):

$$\dot{x}_0 = (1 - \mu)x_0\gamma[r_0p_0 + r_1p_1] - x_0\Phi(x), \quad [1]$$

$$\dot{x}_1 = (\mu\gamma x_0 + (1 - \gamma)x_1)[r_0p_0 + r_1p_1] - x_1\Phi(x), \quad [2]$$

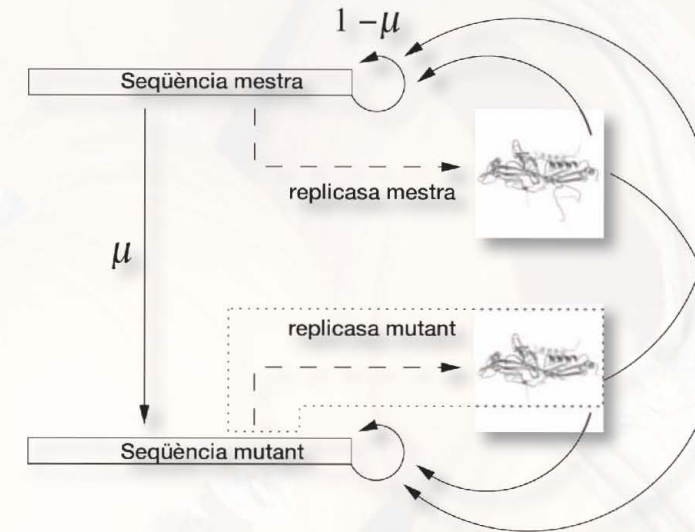
$$\dot{p}_i = k_i x_i - p_i \Xi(p), \quad [3]$$

L'equació 1 determina la taxa de canvi en la concentració de la seqüència mestra en funció del

temps (el punt sobre x_0 significa la derivada d' x_0 respecte del temps: dx_0/dt). La producció de cadenes mestres dependrà de la replicació sense error (indicada amb el paràmetre $1 - \mu$), que és proporcional a la presència de cadenes mestres (x_0) i a la quantitat de replicasa present en el sistema (p_0). La constant γ indica la fracció de replicases utilitzada en l'amplificació de les seqüències mestres. L'últim terme de l'equació 1 s'utilitza per evitar que la població de cadenes mestres pugui créixer indefinidament. Les constants r_0 i r_1 corresponen, respectivament, a les taxes de replicació de les replicases mestres i mutants. L'equació 2 ens indica com canvia la concentració de seqüències mutants amb el pas del temps. El primer terme dins el parèntesi modelitza la formació de seqüències mutants per canvis en la seqüència mestra. El segon terme correspon a la replicació de les seqüències mutants, que seran replicades amb la resta de les replicases que no estan copiant seqüències mestres, com denota el terme $(1 - \gamma)$. De la mateixa manera que en la primera equació, l'últim terme de la segona s'utilitza per restringir el creixement de les seqüències mutants. Finalment, l'equació 3 modelitza la producció de les replicases (p_i ; $i = 0, 1$) proporcionalment als paràmetres k_i ($i = 0, 1$). L'últim terme d'aquesta equació restringeix la producció de proteïnes a un valor finit.

Els resultats

Els sistemes d'equacions diferencials (ED) ens determinen com canvien certes variables en funció d'uns paràmetres a mesura que avança el temps (tot i que també es poden estudiar per temps negatius). Existeixen una multitud de tècniques per resoldre numèricament un conjunt d'equacions diferencials amb l'ordinador i poder visualitzar com evolucionen les variables al llarg del temps. Generalment, per a un sistema d'ED podem estudiar diferents característiques, com ara quina serà la dinàmica quan el sistema ha assolit l'equilibri dinàmic, quin és l'efecte quan canvien un o més paràmetres o

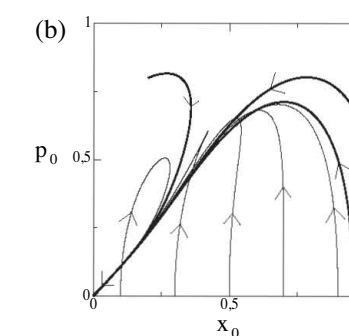
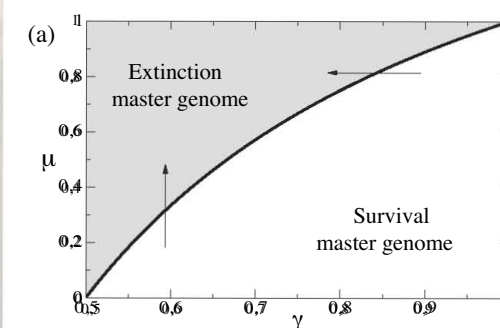


◀ **Figura 2.** Diagrama esquemàtic del sistema analitzat. El nostre model considera un sistema dinàmic on hi ha dos tipus de genomes virals de RNA, les seqüències mestres i els mutants. Cada tipus de genoma sintetitza un tipus de replicasa. A conseqüència dels fenòmens de complementació viral, cada genoma pot ser copiat per qualsevol tipus de replicasa. Quan les seqüències mestres són copiades, una fracció d'aquestes canvia de tipus de seqüència a causa de les mutacions. El nostre model assumeix que no hi ha mutacions que reverteixen els mutants a seqüències mestres.

com el sistema s'aproxima a l'equilibri. Una eina molt útil per esbrinar com es comporten les equacions és representar la dinàmica en l'anomenat *espai de fases* o *espai d'estats*. Aquest espai es construeix considerant les variables de les ED com a coordenades, en el nostre cas, d'un espai de concentracions. El nostre model té quatre variables, de manera que l'espai de fases tindria

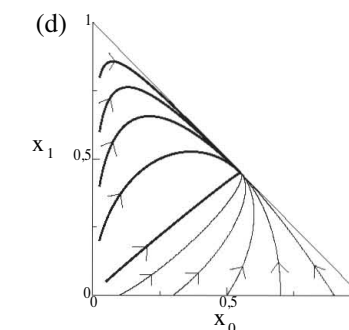
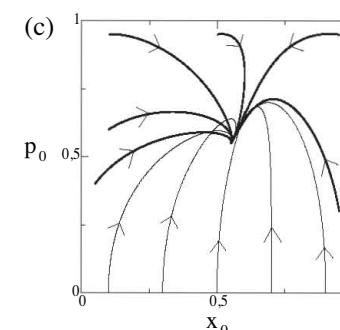
quatre dimensions. En podem estudiar la dinàmica fent una projecció en el pla o en un espai tridimensional. Per a cada punt de temps, cada variable del sistema d'ED té un valor concret, que anirà canviant o no en funció del temps. Si per cada punt de temps dibuixem un punt en l'espai

de fases corresponent al valor de cada variable, obtindrem una trajectòria que seguirà un cert camí en aquest espai fins a assolir un equilibri. El sistema dinàmic descrit per les equacions 1 - 3 és caracteritzat per dos possibles estats d'equilibri en funció dels valors dels paràmetres: I) manteniment de la seqüència mestra



(variable $x_0^* > 0$; on l'asterisc indica l'equilibri), que coexisteix amb les seqüències mutants; II) extinció de les seqüències mestres i domini dels mutants. En el diagrama de fases de la fig. 3.a, obtingut numèricament, mostrem la dependència d'aquests dos escenaris en funció

de fases corresponent al valor de cada variable, obtindrem una trajectòria que seguirà un cert camí en aquest espai fins a assolir un equilibri. El sistema dinàmic descrit per les equacions 1 - 3 és caracteritzat per dos possibles estats d'equilibri en funció dels valors dels paràmetres: I) manteniment de la seqüència mestra



◀ **Figura 3.** a) Diagrama de fases on es mostren els possibles estats d'equilibri del sistema dinàmic donat per les equacions 1 - 3 en funció de la taxa de mutació (μ) i l'afinitat de les seqüències mestres per les replicases (γ). La regió blanca indica per a quins valors d'ambdós paràmetres la seqüència mestra pot sobreviure. La regió fosca indica l'extinció dels genomes mestres, on la població és formada només per genomes mutants. La línia negra ens indica la frontera entre ambdós estats, on es produeix la bifurcació transcítica. b) Retrat de fase per a l'escenari d'extinció, on la taxa de mutació és $\mu > \mu_c$. En aquest escenari, els genomes mestres s'extingeixen per a qualsevol condició inicial dins l'espai de fases. c) i d) Retrats de fase per a l'escenari de supervivència i de coexistència entre seqüències mestres i mutants. Les fletxes indiquen la direcció de les trajectòries a mesura que el sistema evoluciona cap a l'equilibri.

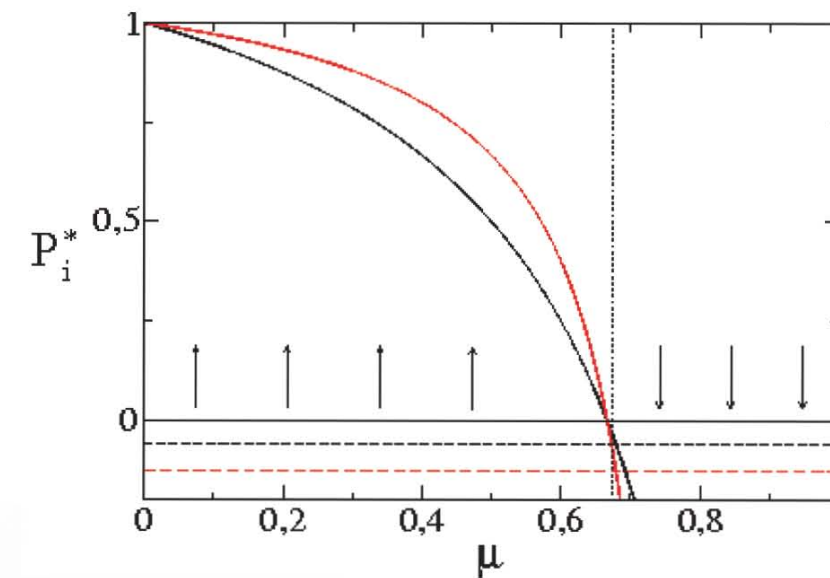
de la taxa de mutació (μ) i de l'afinitat de la seqüència mestra amb les replicases (γ). Podem veure que el canvi de la fase I a la II es pot assolir per l'increment de la taxa de mutació o mitjançant la disminució de l'afinitat de la seqüència mestra per les replicases. A les fig. 3.b i 3.c mostrem la dinàmica de les variables x_0 (concentració seqüència mestra) i p_0 (concentració proteïnes sintetitzades per la seqüència mestra), per a diferents condicions inicials en el pla de fase (x_0, p_0), mentre que en la fig. 3.d projectem la dinàmica en el pla de fase triangular (x_0, x_1), ja que el model també assumeix que $x_0 + x_1 = 1$. Aquest tipus de visualització, que es coneix per *retrats de fase*, ens mostra com evolucionen les trajectòries per a diferents condicions inicials fins a assolir l'equilibri. Per al cas mostrat en la fig. 3.b, les seqüències mestres s'extingeixen i les trajectòries es mouen cap al punt d'equilibri (0, 0) del pla. En els retrats de fase de les fig. 3.c i 3.d, les trajectòries assolixen punts d'equilibri on $x_0^* > 0$ i, per tant, les seqüències mestres persisteixen en el sistema.

Les transicions entre els diferents tipus de dinàmiques o escenaris d'un sistema dinàmic s'anomenen *bifurcacions*. Hi ha diferents tipus de bifurcacions. Per exemple, la bifurcació del nostre model és del tipus transcítica, de manera que, en moure un paràmetre (per exemple, en augmentar la taxa de mutació), es produeix un canvi qualitatiu suau en la dinàmica. El valor del paràmetre en què es produeix el canvi es coneix per *valor crític* o *valor de bifurcació*. En el nostre sistema (fig. 4), l'augment de la taxa de mutació implica una disminució de les concentracions d'equilibri de les seqüències mestres (corba negra) i de les replicases (corba vermella).

A un valor de mutació donat (línia puntejada vertical) corresponent al valor crític de mutació (μ_c) la concentració en l'equilibri de seqüències mestres i replicases s'extingeix (travessa la línia del zero). És a dir, per l'esquerra de la línia puntejada tenim l'escenari de supervivència de la seqüència mestra, mentre que per la dreta tenim l'escenari d'extinció dels genomes mestres. Aquests dos escenaris estan separats per μ_c : la bifurcació es produeix quan $\mu = \mu_c$.

Conclusions

La conclusió principal d'aquest article, il·lustrat amb un model per estudiar un fenomen descrit en sistemes reals, és que la complexitat inherent als sistemes biològics pot ésser estudiada formalment amb models matemàtics relativament simples. Per al sistema estudiat hem caracteritzat les bifurcacions en funció de dos paràmetres de rellevància biològica: la taxa de mutació i l'afinitat de la seqüència mestra amb les replicases. La complexitat d'un tipus de model no és determinada, en bona part, pel nombre d'equacions o paràmetres, sinó pel tipus d'interaccions modelitzades. En el model presentat en aquest escrit, la dinàmica consistia en dues possibles fases: supervivència o extinció de seqüències mestres. Altres models poden presentar dinàmiques diferents, com ara oscil·lacions regulars o caòtiques. És molt important tenir present que els models són simplificacions de la realitat i que tots arrossegueu una sèrie d'assumpcions. El nostre model, per exemple, només considera dos tipus de seqüències, i no té en compte l'espai físic ni el soroll present en els sistemes moleculars. Tot i això, ens ha permès entendre una mica més quins possibles estats dinàmics caldria esperar per a una població de genomes replicants sota escenaris de mutacions deletèries i sota processos de complementació. Per acabar, volem destacar que no hi ha models matemàtics certs, sinó útils. I



◀Figura 4. Bifurcació cap a l'extinció de la seqüència mestra (corba negra) i de la replicasa sintetitzada per la seqüència mestra (corba vermella). La línia puntejada vertical indica el valor crític de la taxa de mutació en el qual es produeix la bifurcació i les dues variables travessen el zero (línia horitzontal) i s'extingeixen.

Agraïments

Vull agrair, d'una manera molt especial, a Ricard Solé i a Santiago Elena el fet d'haver compartit amb mi l'entusiasme per la biologia i la recerca. Els treballs i els resultats presentats en aquest article han estat subvencionats pel projecte RGP12/2008: *Evolutionary implications of virus-encoded gene silencing evolution*, finançat per la Human Frontier Science Program Organization.

Referències bibliogràfiques

- BRENNER, S. (1957). «Genetic control and phenotypic mixing of the adsorption cofactor requirement in bacteriophages T2 and T4». *Virology*, núm. 3, p. 560-574.
- CICIN-SAIN, L. [et al.] (2005). «Frequent coinfection of cells explains functional in vivo complementation between cytomegalovirus variants in the multiply infected hosts». *J. Virol.*, núm. 79, p. 9492-9502.
- DOMINGO, E. (2002). «Quasispecies Theory in Virology». *J. Virol.* núm. 76, p. 463-465.
- EIGEN, M. (1971). «Self-organization of matter and the evolution of biological macromolecules». *Naturwissenschaften*, núm. 58, p. 465-523.
- FRAILE, A. [et al.] (2008). «A quantitative analysis of complementation of deleterious mutants in plant virus populations». *Span. J. Agricult. Res.*, núm. 6, p. 195-200.
- GARCÍA-ARRIAZA, J. [et al.] (2004). «Evolutionary transition towards defective RNAs that are infectious by complementation». *J. Virol.*, núm. 78, p. 11678-11685.
- HUANG, A. S. [et al.] (1974). «Pseudotype formation between enveloped RNA and DNA viruses». *Nature*, núm. 252, p. 743-745.
- HOLLAND, J. J. [et al.] (1989). «Virus mutation frequencies can be greatly underestimated by monoclonal antibody neutralization of virions». *J. Virol.*, núm. 63, p. 5030-5036.
- LORENZ, EDWARD N. (1963). «Deterministic Nonperiodic Flow». *Journal of the Atmospheric Sciences*, núm. 20, p. 130-141.
- NOVICK, A.; SZILARD, L. (1951). «Virus strains of identical phenotype but different genotype». *Science*, núm. 113, p. 34-35.
- SARDANYÉS, J.; ELENA, S. F. (2010). «Error threshold in RNA quasispecies models with complementation». *J. Theor. Biol.*, núm. 265, p. 278-286. [Treball referenciat en aquest article].
- SOLÉ, R. V.; MANRUBIA, S. C. (2001). «Orden y caos en sistemas complejos. Fundamentos - Aplicaciones». Edicions UPC.
- Wikipedia. <http://en.wikipedia.org/wiki/Newtonian_mechanics>.
- WILKE, C. O.; NOVILLA, I. (2003). «Phenotypic mixing and hiding may contribute to memory in viral quasispecies». *BMC Microbiol.*, núm. 3, p. 11.

Josep Sardanyés i Cayuela

(Granollers, 1976)

Josep Sardanyés és llicenciat en biologia per la Universitat de Barcelona (2001) i doctor en biologia per la Universitat Pompeu Fabra (2009). La tesi doctoral, la va desenvolupar en el marc del projecte PACE (Programmable Artificial Cell Evolution), sota la supervisió del doctor Ricard V. Solé. Els seus camps d'estudi són els sistemes dinàmics no lineals aplicats a processos biològics. Durant la tesi doctoral va investigar la teoria de l'hipercicle i de les quasispècies moleculars dins els camps de l'evolució prebiòtica, l'origen de la vida i els virus de RNA. Actualment està realitzant un curs postdoctoral en virologia teòrica i computacional al laboratori del professor Santiago F. Elena, a l'Institut de Biologia Molecular i Cel·lular de Plantes (CSIC-UPV, València).