

---

# Morfogenesi: una sfida interdisciplinare

*Everything is made of atoms. That is the key hypothesis. The most important hypothesis in all of biology, for example, is that everything that animals do, atoms do.*

*In other words, there is nothing that living things do that cannot be understood from the point of view that they are made of atoms acting according to the laws of physics.*

**R. P. Feynman**

---

**Silvia Grigolon**

Laboratoire Jean Perrin, Sorbonne Université & CNRS (UMR 8237)  
4, Place Jussieu, 75005 Paris, France

---

**Q**uali sono i processi e principi primi che permettono ad un organismo di assumere la sua forma funzionale per poter espletare tutte le funzioni vitali a partire da una cellula fecondata? Come questi processi interagiscono fra di loro attraverso le diverse scale biologiche, ovvero dalle cellule ai tessuti e viceversa? È possibile utilizzare le scoperte a questo riguardo per poter sviluppare terapie finalizzate a curare malformazioni e malattie dello sviluppo? In questo articolo, si illustreranno, dapprima, con una rassegna storica, la nascita dello studio dello sviluppo degli organismi da un punto di vista matematico quantitativo ed i diversi problemi da risolvere per rispondere alle domande sovraccitate; verrà poi fornita una panoramica degli approcci tecnici fisico-matematici utilizzati nella biofisica della Morfogenesi e di come questi possano integrare e siano integrati con la scienza speri-

mentale; si concluderà infine con una discussione delle sfide a lungo termine, non solo tecniche ma anche di sviluppo di una comunità interdisciplinare conscia delle potenzialità, limiti e soprattutto complementarità delle singole discipline e della sua fondamentale importanza per poter rispondere a domande fondamentali quali “Come si sviluppa la vita?”.

## Introduzione

La **Morfogenesi** (dal greco *morphê*=forma e *genesis*=creazione) è l'insieme di tutti i processi che regolano lo sviluppo degli organismi viventi a partire da una cellula uovo fecondata. Per poter espletare tutte le sue funzioni vitali, l'organismo deve poter raggiungere una determinata forma ottimale. Lo sviluppo pertanto richiede l'interazione coordinata di diversi processi su diverse scale, spaziali e temporali, quali la divisione cellulare, la formazione di tessuti (insiemi di cellule) in grado di crescere, deformarsi

ma ciononostante mantenere integrità durante l'intero sviluppo, e la formazione di diversi organi, che costituiscono le macrounità funzionali fondamentali.

Come questi processi avvengono a partire da una singola cellula o da un insieme di cellule apparentemente omogenee nella loro composizione biochimica è ormai oggetto di studio da più di un secolo. I primi pionieri nello studio teorico della Morfogenesi furono D'Arcy Wentworth Thompson e Alan Turing, che con le loro intuizioni possono essere considerati i padri non solo dello studio quantitativo della biologia dello sviluppo ma soprattutto dell'odierna biofisica teorica.

Il lavoro principale dello scozzese D'Arcy Wentworth Thompson del 1917, intitolato "*On Growth and Form*" (Della crescita e della forma) [1], fu tra i primi a contrastare l'idea che l'evoluzione potesse fungere da unica forza motrice nell'acquisizione della forma funzionale da parte degli organismi, postulando l'ipotesi che anche le leggi fisiche e meccaniche giocassero un ruolo fondamentale nel governare non solo la crescita degli organismi stessi ma anche la speciazione. Nel suo lavoro notò le correlazioni presenti tra forme geometriche e forze meccaniche ed allo stesso tempo fra un problema tipicamente biologico ed un problema tipicamente fisico. Riportò infatti l'analogia tra la forma delle meduse e quella assunta da una particella liquida in un fluido viscoso o tra la struttura esagonale di un alveare e gli aggregati di bolle di sapone. Identificò le relazioni matematiche presenti in determinati fenomeni di crescita, come la fillotassi, ovvero l'ordine in cui foglie o rami nelle piante sono disposti in modo da seguire una spirale di Fibonacci, e la relazione fra superficie e volume di un organismo e la sua lunghezza.

Nonostante l'opera racchiuda in sé una quantità di esempi largamente maggiore, la selezione sovraccitata dimostra ampiamente l'intuizione formidabile, se non geniale, dell'autore, ovvero la correlazione fondamentale tra forze meccaniche e forme geometriche riproducibili e funzionali e l'esistenza di relazioni matematiche precise che regolano lo sviluppo e la crescita degli organismi.

Le innumerevoli osservazioni di D'Arcy Thompson segnarono l'inizio di quella che poi

si sviluppò come un'intera branca interdisciplinare, la biofisica, dove biologia, fisica e matematica non sono più discipline separate ma diversi approcci di studio da integrare per poter svelare i meccanismi fondamentali ed i principi primi su cui si basa il mondo naturale ed il suo funzionamento collettivo.

Tuttavia, pur non avendo pretesa di essere un testo quantitativo, l'opera di D'Arcy Thompson resta un testo descrittivo. Per poter osservare un primo studio quantitativo della Morfogenesi, bisogna aspettare un'altra opera fondamentale, ovvero l'articolo del 1952 di Alan Turing [2]. Turing fu il primo ad ipotizzare che la crescita degli organismi sia anzitutto regolata dalla distribuzione non uniforme nei tessuti di sostanze chimiche, che lui stesso denominò *morfogeni*, finalizzate a guidare il differenziamento cellulare, ovvero specificare la funzione che le diverse cellule espletano nel tessuto stesso.

Turing propose un modello inizialmente minimale, in seguito divenuto più complesso, composto principalmente da due specie soggette a diffusione e reazioni chimiche di diverso tipo, identificando le relazioni necessarie tra questi diversi fenomeni chimici per la formazione di pattern. Grazie a questo modello minimale, Turing aprì un'altra strada per lo studio dello sviluppo degli organismi viventi, ovvero quella della modellizzazione di reti molecolari su grande scala.

In parallelo, la scoperta della struttura del DNA da parte di Rosalind Franklin, James Watson e Francis Crick, gli enormi sviluppi tecnologici sul sequenziamento genico, la manipolazione e l'interferenza genetica, la microscopia ad alta risoluzione e la misura di forze *in vitro* ed *in vivo*, hanno permesso non solo di validare molte delle intuizioni di D'Arcy Thomson e Turing, ma di andare oltre analizzando sempre più in dettaglio problemi quali le intricate ed innumerevoli interazioni fra le molecole e la generazione e trasmissione di forze all'interno delle cellule e dei tessuti.

Lo scopo di questo articolo è di fornire una panoramica degli approcci teorici generali utilizzati oggi nello studio dello sviluppo dei sistemi biologici a diverse scale spazio-temporali e della loro interazione con la branca sperimentale per validazione e predizione. Pertanto, si inizia-

rà dapprima discutendo la scala microscopica della cellula, l'espressione genica e la modellizzazione delle reti molecolari, riassumendo gli approcci ed i risultati fondamentali; si proseguirà con lo studio della formazione degli organuli cellulari, ovvero della formazione di pattern a livello della cellula stessa, a cui seguirà la scala dei tessuti, dove si illustrerà come le idee di Turing vengono applicate oggi giorno alla comprensione del differenziamento cellulare; si discuterà poi la questione dell'identificazione di modelli di reti geniche ottimali e di come metodi basati su o ispirati all'intelligenza artificiale possano contribuire a risolvere questo problema.

Nella seconda parte dell'articolo, ci si concentrerà invece sull'altra faccia della medaglia, la meccanica, e si illustrerà una delle teorie maggiormente utilizzate nello studio della generazione e trasmissione di forze meccaniche e di come queste possano essere validate sperimentalmente.

In finale, si concluderà discutendo le sfide a lungo termine, riguardanti non solo le domande fondamentali ancora aperte e gli sviluppi tecnici teorici e sperimentali necessari ad uno studio completo della Morfogenesi ma anche l'importanza dello sviluppo di una vera comunità interdisciplinare in continua sinergia.

## La cellula e la modellizzazione delle reti geniche e di segnale

Come discusso nell'introduzione, la morfogenesi ha inizio a partire dalla singola cellula fecondata, la quale, per poter espletare tutte le sue funzioni vitali, necessita della produzione di proteine, le macromolecole funzionali fondamentali. L'informazione necessaria alla produzione di proteine, ossia l'informazione sulla loro struttura, è contenuta nei geni, frammenti di DNA, i quali vengono espressi tramite una serie di processi i cui stadi intermedi fondamentali richiedono la produzione dell'RNA messaggero (mRNA), anche detta trascrizione, e la sua traduzione nella sequenza primaria della proteina, la quale, attraverso il processo di *folding*, assume la sua struttura finale e, se non vi sono errori, funzionale. L'insieme di tali processi è anche noto come espressione genica.

Una volta che la proteina finale è prodotta, interagisce all'interno della cellula con le altre macromolecole, espletando un'innumerabile quantità di funzioni, come ad esempio l'attivazione stessa dell'espressione genica (ovvero i fattori di trascrizione), degradazione di altre proteine (come le proteasi), costituzione del citoscheletro (ossia la struttura fondamentale che controlla forma e motilità cellulare) e costituzione della membrana cellulare per la comunicazione con l'ambiente circostante.

Nonostante ad oggi non si conosca con esattezza l'intero quadro di interazioni intracellulari e la loro stechiometria, grazie a diverse tecniche sperimentali, negli anni è stato tuttavia possibile ricostruirne una grande quantità. Ciò ha permesso di catalogare parte di queste interazioni in circuiti molecolari, ovvero reti minimali composte da un basso numero di macromolecole [3].

Per poter comprendere il funzionamento di questi circuiti da un punto di vista teorico, nel corso degli ultimi decenni, si sono sviluppati approcci basati sulla teoria delle reti complesse e sulla teoria dei sistemi dinamici. In questo contesto, le diverse specie molecolari d'interesse vengono rappresentate come i **nodi** della rete, le interazioni come i *link* della rete (Figura 1A) e la variabile primaria di interesse è la concentrazione molecolare di ogni specie presente nella rete.

Supponendo quindi di considerare  $N$  specie, si può definire il vettore delle concentrazioni delle specie molecolari al tempo  $t$  come  $\underline{y}(t) = (y_1(t), \dots, y_N(t))$ . Date le interazioni tra le diverse specie molecolari, è quindi possibile definire le equazioni che governano l'evoluzione temporale della rete, ovvero:

$$\frac{dy_i(t)}{dt} = f_i(\underline{y}(t), t),$$

dove  $y_i(t)$  è la concentrazione della specie  $i$  al tempo  $t$  e  $f_i(\underline{y}(t), t)$  è una funzione, normalmente non lineare, contenente le reazioni chimiche che coinvolgono la specie  $i$ , come ad esempio produzione ed interazione della specie  $i$  con se stessa e con le altre specie. Pertanto, tale funzione dipende anche dai parametri del modello, ovvero dalle velocità delle reazioni chimiche, dalle soglie di reazione o dagli esponenti caratterizzanti

funzioni tipicamente non lineari <sup>1</sup>.

Tramite questo approccio è perciò possibile studiare una serie di proprietà legate alla rete stessa come ad esempio i) l'esistenza ed unicità di uno stato stazionario e la sua stabilità rispetto a perturbazioni esterne; ii) le esatte proprietà introdotte dalla presenza di determinate reazioni chimiche all'interno della rete.

A quest'ultimo proposito, un'ampia branca della biofisica teorica si è occupata di catalogare dapprima circuiti molecolari sovrarappresentati in biologia per poi comprenderne le proprietà dinamiche e statiche. Si è studiato i) come la presenza di reazioni fortemente non lineari potesse indurre una bistabilità nel sistema e, di conseguenza, la possibilità per la cellula di assumere diversi stati in condizioni stazionarie, come nel caso del *toggle switch* (Figura 1B); ii) come reazioni di inibizione (*negative feedback*) potessero indurre un'alta resilienza (Figura 1C) o oscillazioni in risposta a perturbazioni esterne [5]; iii) come reazioni di mutua attivazione (*positive feedback*) rendano gli stati stazionari instabili o addirittura caotici [9]. Nonostante vi siano già alcuni studi [8, 9], restano tuttavia ancora da comprendere in maniera sistematica le proprietà dinamiche indotte dalla combinazione di diversi circuiti molecolari e come questi possano risultare ottimali da un punto di vista evolutivo.

Nonostante l'approccio sovracitato permetta di studiare un'elevata varietà di sistemi, resta comunque di tipo deterministico, ovvero ignora completamente la stocasticità dovuta all'interazione fra le molecole in quanto quantità discrete (rumore intrinseco) o a fluttuazioni dell'ambiente circostante la rete (rumore estrinseco). Pertanto, questo approccio limita la comprensione di una grande quantità di fenomeni cellulari, come si discuterà di seguito.

Per superare questo limite, è quindi necessario adottare una descrizione stocastica del sistema, in cui la variabile di interesse non è più il numero di molecole (o concentrazione) di una determinata specie al tempo  $t$  ma piuttosto la probabilità

<sup>1</sup>In generale, è possibile semplificare tale descrizione introducendo al posto della concentrazione di ciascuna specie nel tempo una variabile binaria (ovvero che possa assumere due valori soltanto, ad esempio 0 e 1), che determini semplicemente lo stato di espressione o non espressione di un gene, come nel caso delle reti booleane.

di osservare un determinato numero di molecole al tempo  $t$ , ovvero  $P(\underline{n}, t)$ , la cui legge di evoluzione temporale è data dall'equazione maestra, ovvero:

$$\frac{dP(\underline{n}, t)}{dt} = \sum_{\underline{n}' \neq \underline{n}} \{T(\underline{n}|\underline{n}')P(\underline{n}', t) - T(\underline{n}'|\underline{n})P(\underline{n}, t)\},$$

dove  $T(\underline{n}'|\underline{n})$  rappresentano le frequenze di transizione dallo stato  $\underline{n} = (n_1, \dots, n_N)$  allo stato  $\underline{n}' = (n'_1, \dots, n'_N)$  e viceversa e sono definite dalle reazioni chimiche presenti nella rete.

La risoluzione di tale equazione è perciò uno strumento fondamentale per studiare il comportamento stocastico del sistema e, di conseguenza, per capire il ruolo giocato dalle diverse interazioni molecolari nel determinare il tipo di distribuzioni di probabilità caratterizzanti le specie presenti nel sistema. Tuttavia, nella maggior parte dei casi che vengono considerati, risolvere l'equazione maestra resta un compito arduo. Pertanto negli anni si è ricorso a diverse approssimazioni, tra cui una delle più celebri denominata di van Kampen (o in inglese anche *system-size expansion*) [6].

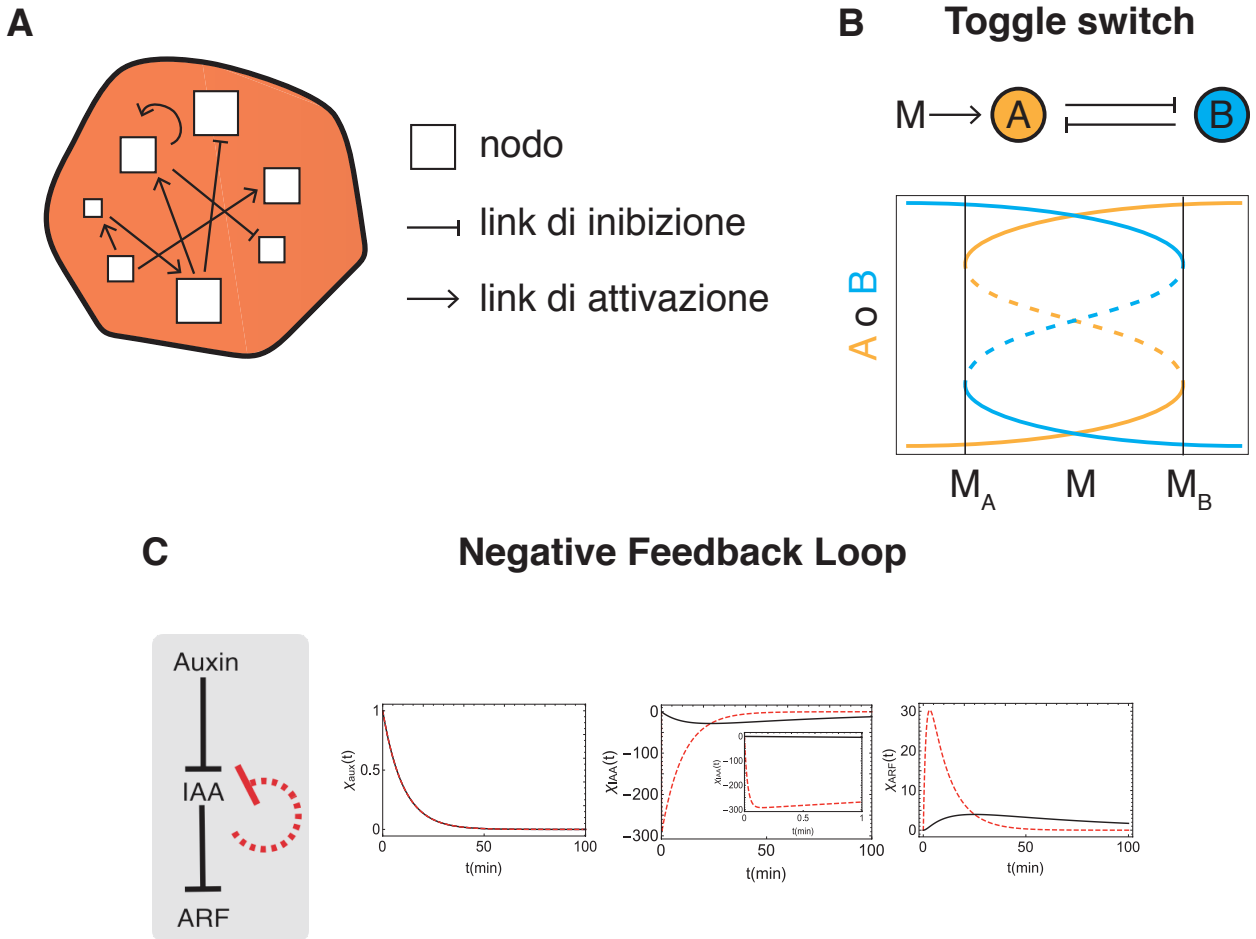
In tale espansione, nel caso di rumore lineare, si assume che il numero di molecole di una determinata specie  $i$  al tempo  $t$ ,  $n_i$ , possa essere approssimato come:

$$n_i = V\phi_i + V^{1/2}\xi_i,$$

dove  $V$  è il volume del sistema,  $\phi_i$  è la concentrazione di molecole della specie  $i$  nel caso deterministico e  $\xi_i$  è il termine di rumore.

Espandendo in termini di potenze del volume  $V$  e fermandosi al secondo ordine dell'espansione, è possibile riscrivere l'equazione maestra nella forma di un'equazione di diffusione per la distribuzione di probabilità, l'equazione di Fokker-Planck, più facilmente trattabile analiticamente.

Questo approccio di conseguenza permette il calcolo dei momenti della distribuzione delle diverse specie molecolari e la loro evoluzione temporale, ovvero media (corrispondente al caso senza rumore), varianza e correlazioni. Bisogna tenere a mente, però, che in tale regime e quadro specifico, si assume a priori che le distribuzio-



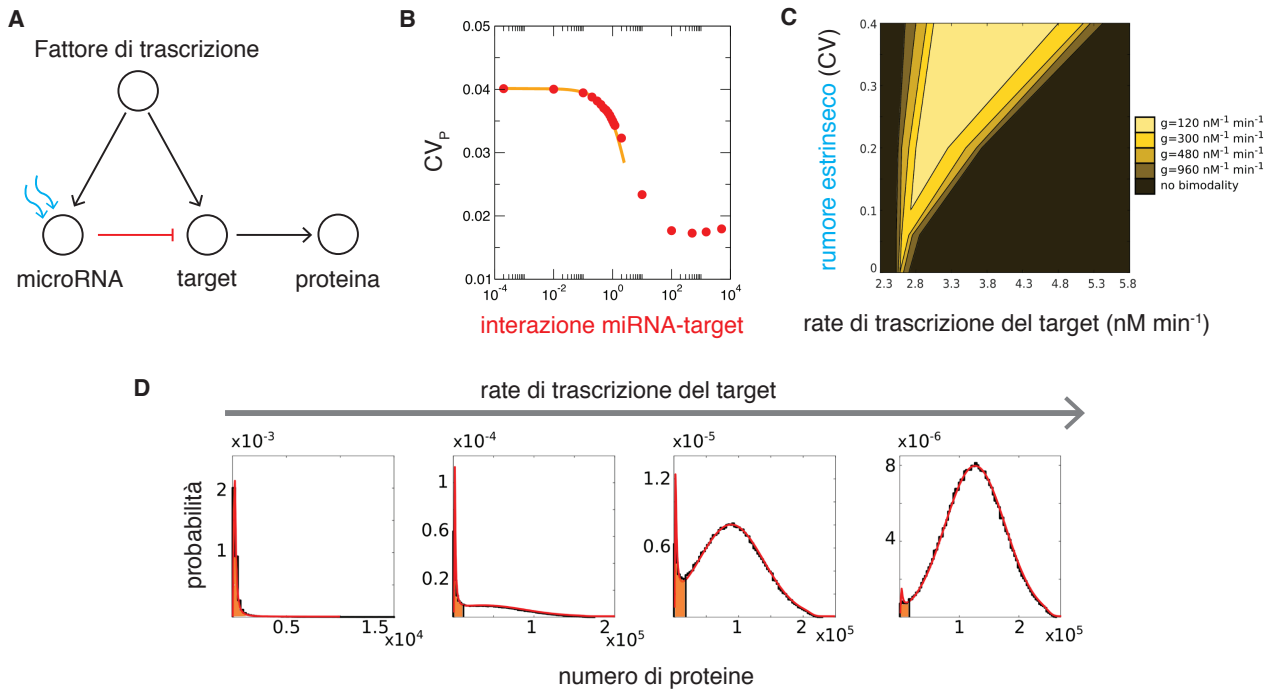
**Figura 1:** A) Schematizzazione della cellula con al suo interno un esempio di rete molecolare. I quadrati rappresentano i nodi della rete, ovvero le molecole coinvolte, mentre le frecce rappresentano le reazioni, descritte in maggiore dettaglio nel pannello laterale. B) Pannello in alto: esempio di toggle switch, comprendente due diversi nodi corrispondenti al gene A e gene B, due reazioni di mutua inibizione e un attivatore M. Pannello in basso: esempio di comportamento di stato stazionario per il gene A ed il gene B in funzione della concentrazione dell'attivatore M. Linee continue rappresentano gli stati stabili mentre le linee tratteggiate rappresentano gli stati instabili.  $M_A$  ed  $M_B$  rappresentano le soglie di concentrazione per l'attivatore M a cui si osserva l'apparizione degli stati stazionari instabili. Figura riadattata da [4]. C) Esempio di Negative Feedback Loop. Pannello di sinistra: circuito comprendente tre specie, Auxin, IAA e ARF, proteine regolatrici dello sviluppo della pianta A. *Thaliana* e 3 reazioni di inibizione. Pannello di destra: risposta temporale per le tre diverse specie indicate a sinistra a seguito di una perturbazione rispetto allo stato stazionario indotta da Auxin. Le linee continue nere rappresentano il caso senza il negative feedback indicato in rosso a sinistra; le linee rosse tratteggiate rappresentano il caso con il negative feedback. Si noti come la presenza del negative feedback effettivamente permette di osservare una risposta maggiore ma anche un'alta resilienza rispetto allo stato stazionario. Figura presa da [5].

ni di probabilità siano gaussiane e non vi siano termini di memoria all'interno del sistema.

A questo riguardo, è bene quindi, oltre che verificare i regimi di validità dell'approssimazione, comprendere se sia necessario tenere conto degli ordini superiori nell'approssimazione o dei termini di memoria o anche direttamente tentare di sviluppare tecniche analitiche più raffinate

(come ad esempio quelle basate sulla teoria dei campi [7]).

Tuttavia, vi sono vari esempi in cui questo studio si è rivelato potente per comprendere i) il ruolo delle interazioni chimiche o anche ii) il ruolo del rumore estrinseco nel determinare le distribuzioni di probabilità delle molecole. In questo contesto è importante ricordare il caso dei microRNA ed il loro ruolo nei circuiti molecola-



**Figura 2:** A) Schematizzazione di un circuito a *microRNA* composto da 4 nodi, ovvero un fattore di trascrizione che attiva l'espressione del *microRNA* e del suo mRNA target, che a sua volta codifica per la proteina. Le frecce nere indicano le reazioni di attivazione, mentre la freccia rossa indica l'inibizione. Le frecce azzurre ondulate sul *microRNA* indicano la trascrizione con rumore estrinseco del *microRNA*. B) Coefficiente di variazione della proteina come funzione del parametro di interazione fra *microRNA* e mRNA target. I punti rossi corrispondono ai risultati delle simulazioni mentre la linea arancione al risultato dell'espansione di van Kampen in questo circuito. Figura presa e riadattata da [12]. C) Diagramma di fase per l'emergenza di distribuzioni bimodali in presenza di rumore estrinseco sul *microRNA* in funzione del rate di trascrizione dell'mRNA target e del coefficiente di variazione del rumore estrinseco sul *microRNA*. Il parametro  $g$  rappresenta la costante di accoppiamento fra il *microRNA* e l'mRNA target. D) Esempio di emergenza di distribuzioni bimodali per le proteine a fissato coefficiente di variazione per il rumore estrinseco sul *microRNA*, variando il rate di trascrizione dell'mRNA target. Gli istogrammi vengono dal risultato delle simulazioni mentre le linee rosse continue dall'approssimazione di van Kampen modificata in presenza di rumore estrinseco. Figura presa e riadattata da [15].

ri. I **microRNA** sono piccole molecole di RNA scoperte negli anni '90 in *C. Elegans* [10], un verme nematode, codificate in regioni del genoma note per non contenere alcuna informazione sulla produzione di proteine (regioni non codificanti).

I *microRNA* agiscono legando gli RNA messaggeri *target* (mRNA) ed eventualmente degradandoli, portando in prima istanza ad una diminuzione del numero di proteine relative tradotte (Figura 2A). L'intera mappa di interazione tra i diversi *microRNA*, che risultano ad oggi essere più di 2000 solo nella specie umana, ed i loro *target* non è ancora completamente nota ma, dati i numeri, è chiaro che si tratti di una rete complessa coinvolta in svariati contesti, dallo sviluppo alla carcinogenesi.

L'uso delle tecniche teoriche sovraccitate in que-

sto caso ha quindi permesso di comprendere ulteriori proprietà caratterizzanti queste reti ed in particolare dell'azione dei *microRNA*, quali: i) la capacità dei *microRNA* di non diminuire unicamente il numero di molecole relative ad una certa proteina ma anche di diminuirne la varianza (Figura 2B), rendendo pertanto la rete più robusta a perturbazioni esterne [11, 12]; ii) in presenza di più *target*, l'azione dei *microRNA* può portare ad una varietà di effetti regolatori diversi e di conseguenza modulare le correlazioni fra mRNA *target* senza che vi sia una diretta interazione fra di loro [13, 14]; iii) la presenza di rumore estrinseco sui *microRNA*, ad esempio dovuto ad intermittenza nella trascrizione (*burst trascrizionali*), può rafforzare la bimodalità indotta dalla stocasticità delle reazioni delle distribuzioni delle proteine codificate dagli mRNA *target* [15] (Figura 2C-D).

Quest'ultimo punto è estremamente importante per due motivi: i) la presenza di rumore estrinseco, ovvero esterno alla rete, intuitivamente dannoso, può in realtà indurre una varietà di distribuzioni comunque ben definite all'interno della rete stessa; ii) la multimodalità indotta dal rumore può in principio risultare benefica nel caso della differenziazione cellulare. Infatti, la presenza di più picchi nella distribuzione di probabilità corrisponde alla presenza di più stati caratterizzati da un diverso numero di proteine, quindi possibilmente relativi a diversi stati di differenziamento della cellula.

Normalmente, l'ipotesi più accreditata per il differenziamento cellulare si basa sulla presenza di reazioni chimiche tali che permettano di ottenere più stati stazionari stabili. Ricordando il *toggle switch* sovraccitato, la presenza del rumore in questo caso permette al sistema di saltare da uno stato stazionario all'altro naturalmente, dove il tempo tipico di questo salto dipende dalla differenza in stabilità dello stato stazionario stesso.

In termini fisici, come anche proposto da Conrad Waddington durante il secolo scorso, definendo una funzione energia a partire dalle equazioni deterministiche, ove possibile, gli stati stazionari sono rappresentati come i minimi di questa funzione, la cui differenza regola quindi il tempo tipico di salto (legge di Arrhenius). Nel caso invece considerato per il microRNA, il sistema di per sé non contiene alcuna reazione chimica che definisca nel caso senza rumore una multistabilità del sistema e pertanto la multimodalità è unicamente indotta dalla combinazione fra struttura circuitale monostabile e presenza di rumore estrinseco.

Gli approcci teorici qui citati sono degli esempi di come i modelli matematici hanno permesso di svelare proprietà intrinseche di queste reti molecolari non accessibili sperimentalmente.

Tuttavia, grazie agli sviluppi tecnologici, ed in particolare grazie alla possibilità di costruire circuiti sintetici o tracciare finemente determinate molecole nel tempo, è possibile validare i modelli e le loro predizioni con esperimenti *ad hoc*, come effettuato nel caso sovraccitato dei microRNA [16, 17] o anche in altri contesti, come quello dei circuiti della pianta *A. Thaliana* [18]. Questo permette quindi di studiare quantitativamente le

proprietà dettagliate di un'enorme varietà di cellule, dalle cellule staminali, vegetali ed animali, fino alle cellule tumorali.

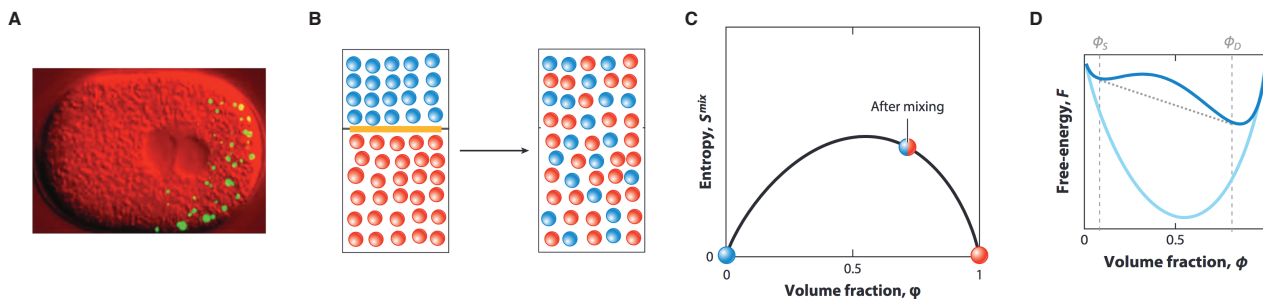
## Formazione di pattern cellulari: la separazione di fase

Negli approcci sovraccitati si assume a priori che il sistema in considerazione sia *well mixed*, ovvero che il numero di molecole presenti nel sistema sia talmente elevato da permettere alle molecole di interagire omogeneamente fra di loro indipendentemente da dove si trovino nella cellula: la nozione di spazio diventa pertanto irrilevante. Quest'ipotesi è ovviamente plausibile in molti casi, ma non in tutti.

A livello cellulare, infatti, la nozione di spazio può ad esempio essere di fondamentale importanza per la corretta formazione di strutture agglomerate, come gli organuli e, soprattutto, gli agglomerati privi di membrana, necessari per il mantenimento delle funzioni vitali della cellula.

Esempi di questo tipo possono essere trovati in diversi tipi di cellule come nella linea cellulare immortalizzata HeLa, nell'ovocita della rana o nell'embrione del verme nematode. Tra questi si possono ricordare i nucleoli, ovvero i siti di biogenesi dei ribosomi, i corpi di Cajal, importanti nel metabolismo degli RNA, ed i granuli germinali, tra cui i granuli P nel verme nematode, i granuli polari nel moscerino della frutta, o anche i granuli di stress. In tutti questi casi, è necessario che le molecole dapprima si aggregino e poi diffondano all'interno dell'aggregato affinché possano interagire fra di loro chimicamente. Ma quali sono i principi che permettono a tali molecole di aggregarsi e come? Come possono queste molecole passare da una fase disordinata in cui sono perfettamente mescolate con le molecole del mezzo circostante ad una in cui sono localizzate?

Per rispondere a questa domanda, è necessario prima di tutto ricordare gli ultimi studi riguardo i granuli P nel verme nematode (Figura 3A). Infatti, è stato recentemente osservato che dal punto di vista reologico i granuli P si comportano come liquidi, ovvero fondono nel



**Figura 3:** A) Visualizzazione dei granuli P (in verde) all'interno dell'embrione del verme nematode *C. Elegans*. Figura presa da [20]. B) Esempio di mixing in un sistema formato da due specie molecolari, rappresentate dai cerchi blu e rossi rispettivamente. Nel momento in cui si rimuove la barriera fra le due specie, l'aumento di entropia per diffusione porta al mescolamento spontaneo delle due specie. C) Entropia di mixing in funzione della frazione volumica  $\phi$ . D) Energia libera come funzione della frazione volumica  $\phi$ , in assenza di interazioni (linea continua azzurra) ed in presenza di interazioni (linea continua blu). I pannelli B, C e D sono stati presi da [19].

momento in cui interagiscono fra di loro assumendo poi una forma sferica, tipica di un liquido sottoposto a tensione superficiale, le loro deformazioni sotto *stress* sono simili a quelle di gocce liquide, e riassumono una forma sferica dopo deformazioni/perturbazioni [19].

Pertanto tali proprietà hanno naturalmente rievocato nella mente dei fisici il problema della separazione di fase, in particolare nei casi sovraccaricati fra due sostanze liquide, la cosiddetta separazione di fase liquido-liquido. Come spiegato in dettaglio in [19], nella separazione di fase liquido-liquido si definisce un sistema contenuto in un certo volume  $V$  formato da due specie differenti,  $A$  e  $B$ , definite dal loro numero di molecole nello spazio e nel tempo,  $n_A$  e  $n_B$ , e la loro rispettiva frazione di volume occupato,  $v_A$  e  $v_B$  (Figura 3B). Pertanto, la concentrazione delle due diverse specie sarà rispettivamente,  $c_A = \phi/v_A$  e  $c_B = (1 - \phi)/v_B$ , dove  $\phi = n_A/V$ .

Date queste definizioni, è possibile introdurre la cosiddetta **entropia di mixing per unità di volume**,  $S_{mix}/V$ , ovvero la quantità che misura il disordine nel sistema, data da:

$$\frac{S_{mix}}{V} = -k_B \frac{\phi}{v_A} \log\left(\frac{\phi}{v_A}\right) - k_B \frac{1-\phi}{v_B} \log\left(\frac{1-\phi}{v_B}\right),$$

dove  $k_B$  è la costante di Boltzmann.

Tale entropia ha la proprietà di essere uguale a 0 nello stato cosiddetto *unmixed*, ovvero quando solo una specie è presente nel sistema (Figura 3C). In tutti gli altri casi, invece, è diversa da

zero e positiva e tiene conto di tutte le possibili configurazioni in cui le particelle appartenenti alle diverse specie possono trovarsi.

In particolare, la positività dell'entropia, in accordo con il secondo principio della termodinamica, implica che il processo di *mixing*, ovvero di trasporto, avviene spontaneamente (Figura 3B). Tipicamente tali processi di trasporto spontanei avvengono via diffusione, il cui tasso di trasporto è dato da

$$J \propto -\frac{d\mu}{dx},$$

dove  $\mu$  è il potenziale chimico. Il potenziale chimico è a sua volta definito come

$$\mu = \frac{v_A}{V} \frac{dF}{d\phi},$$

dove  $F$  è l'energia libera del sistema,  $F = E - TS_{mix}$ ,  $E$  è l'energia interna che tiene conto delle interazioni fra le particelle e  $T$  è la temperatura del sistema.

Considerando il limite di interazioni deboli o trascurabili rispetto all'energia termica tipica del sistema, si ottiene che il potenziale chimico è uguale a

$$\mu = -T \frac{v_A}{V} \frac{dS_{mix}}{d\phi},$$

ovvero

$$-\frac{d\mu}{dx} = T \frac{v_A}{V} \frac{d^2 S_{mix}}{d\phi^2} \frac{d\phi}{dx}.$$

Dato che l'entropia di *mixing* è concava, si ottiene che la corrente di diffusione è direttamente proporzionale al gradiente di concentrazione secondo un coefficiente che è proprio la costante di diffusione  $D$  delle molecole. Pertanto ciò mostra



come, unendo due compartimenti occupati da diverse specie molecolari miscibili, via sola diffusione, sia possibile raggiungere spontaneamente uno stato caratterizzato da una concentrazione omogenea.

Tuttavia, a questo punto, sorge spontanea la domanda: come invece è possibile ottenere dei pattern non banali per le due specie molecolari e quindi mantenere uno stato *unmixed*?

In questo caso, bisogna tenere in considerazione invece le interazioni microscopiche fra le molecole e pertanto il contributo dell'energia interna  $E$  all'energia libera  $F = E - TS_{mix}$ .

Uno dei casi più semplici da considerare è quello in cui l'energia interna favorisca configurazioni in cui molecole della stessa specie siano vicine nello spazio. Una tale energia interna può essere scritta come:

$$E = \alpha V \phi(1 - \phi),$$

dove  $\alpha > 0$  è un parametro che quantifica l'interazione tra le diverse specie molecolari.

Tale energia ammette due minimi, uno a  $\phi = 0$  ed un altro a  $\phi = 1$ , che corrispondono al *packing* di una sola specie. Pertanto la presenza di interazioni favorisce il cosiddetto *demixing* e può rendere l'energia libera concava in determinate regioni dello spazio delle concentrazioni, ovvero portare all'esistenza di due minimi, come nella teoria delle transizioni di fase (Figura 3D).

Da ciò, in generale, ci si potrebbe domandare perché diverse fasi con composizioni diverse possano coesistere senza che flussi diffusivi tendano ad omogeneizzarne le concentrazioni. Ciò è dovuto principalmente al fatto che i flussi diffusivi, come suddetto, sono guidati dapprima da gradienti di potenziale chimico.

Nel caso della separazione di fase, due diverse frazioni volumiche possono essere possibili, ovvero una per il soluto,  $\phi_D$ , ed una per il solvente,  $\phi_S$ , che corrispondono ai due minimi dell'energia libera  $F$  del sistema, con valori uguali di potenziale chimico, e rendono il potenziale chimico stesso non monotono.

Essendo tuttavia il potenziale chimico uguale in corrispondenza di questi due valori, all'equilibrio all'interno del sistema non possono generarsi flussi diffusivi che possano portare ad un mixing delle due fasi. In particolare, si posso-

no ricordare due diversi casi in cui il potenziale chimico è costante spazialmente: i) il caso di una goccia in un solvente; ii) il caso di due fasi omogenee separate da un'interfaccia.

La differenza fra i due casi risiede nel comportamento della pressione, che mentre nel primo caso è uniforme, nel secondo presenta un salto, in particolare la differenza fra la pressione interna alla goccia e la pressione esterna,  $\Delta P$ , è data dalla legge di Laplace, ovvero:

$$\Delta P = \frac{2\gamma}{R},$$

dove  $\gamma$  è la tensione superficiale della goccia stessa ed  $R$  il suo raggio, che segue dal bilancio delle forze normali all'interfaccia. Pertanto, il potenziale chimico, oltre che dipendere dall'interazione molecolare all'interno di una certa fase, dipende anche dalla pressione interna della fase stessa. Questo implica che il potenziale chimico nella goccia a raggio minore sarà più elevato che in quella a raggio maggiore. Da ciò segue che il trasporto diffusivo avverrà in direzione della goccia a raggio maggiore a spese di quella a raggio minore, fenomeno anche chiamato *Ostwald ripening*.

Ovviamente, nel caso di sistemi biologici, tale fenomeno è limitato dal numero di molecole delle diverse specie presenti nel sistema, che risulta pertanto essere anche un fattore di controllo per la taglia degli aggregati che possono formarsi all'interno delle cellule.

L'intero quadro qui descritto, tuttavia, resta una descrizione all'equilibrio per sistemi fortemente fuori equilibrio, dove vi sono reazioni chimiche che permettono la formazione e rottura di legami molecolari, introducendo quindi flussi molecolari costanti che rendono il sistema attivo (il cui significato completo sarà chiarito nei prossimi paragrafi). Pertanto, questo tipo di sistemi viene più generalmente studiato definendo un'equazione di conservazione della materia per il soluto la cui dinamica è regolata dalle reazioni chimiche presenti nel sistema e da un'energia libera contenente i termini di energia interna ed entropia di mixing sovradescritti ed un termine di tensione superficiale atto a mimare l'effetto della differenza tra la pressione interna ed esterna (ossia generalizzando la teoria di Cahn-Hilliard combinata

con l'energia libera di Flory-Huggins).

Varianti e semplificazioni di tale tipo di sistema vengono studiate per comprendere le condizioni precise per cui si possa osservare la separazione di fase, la forma della soluzione allo stato stazionario, la legge dinamica con cui crescono i domini di separazione e le condizioni per cui si possa osservare il suddetto Ostwald ripening. Tra le applicazioni alla biologia, tra i numerosi esempi si possono ricordare appunto la comprensione della formazione dei granuli P nel verme nematode [21] e della separazione di fase come meccanismo per ridurre il rumore all'interno delle cellule [22].

La separazione di fase, in particolare in fluidi attivi, risulta essere quindi una sfida non solo dal punto di vista teorico, per poter costruire un quadro tecnico consistente in grado di descrivere la fisica fuori equilibrio di questo tipo di sistemi, ma anche dal punto di vista biologico per poter comprendere la formazione di pattern e strutture all'interno delle cellule.

## I tessuti e la formazione di pattern per la specificazione degli organi

Analogamente alla formazione di *pattern* cellulari finalizzati a strutturare l'interno della cellula, a livello dei tessuti possono essere osservate concentrazioni disomogenee nell'espressione delle proteine tra le diverse cellule. I principi di formazione di tali *pattern* sono alla base dell'iniziazione degli organi, le macrounità funzionali degli organismi.

Uno dei primi a postulare l'ipotesi che tali *pattern* potessero essere generati spontaneamente da fenomeni di trasporto all'interno dei tessuti fu Alan Turing nel suo articolo del 1952 intitolato "The Chemical Basis of Morphogenesis" [2]. In tale articolo, Turing postula l'esistenza di sostanze, definite **morfogeni**, i cui profili di concentrazione non omogenei inducono l'iniziazione degli organi in determinate regioni spaziali ed in determinate finestre temporali.

Il modello di Turing più minimale prende in considerazione due morfogeni, un attivatore ed un inibitore, diffusivi ed interagenti fra di loro. Analizzando tali equazioni, è possibile effetti-

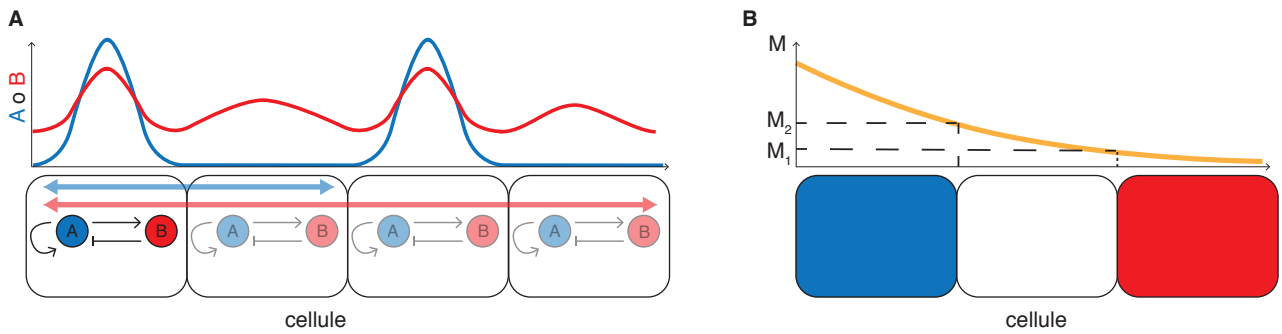
vamente mostrare che, in presenza di reazioni non lineari fra i due morfogeni e una diffusività maggiore dell'inibitore rispetto all'attivatore, lo stato omogeneo diventa instabile sotto perturbazioni, contrariamente a quanto ci si aspetterebbe intuitivamente dalla classica teoria della diffusione. In queste condizioni lo stato stabile è effettivamente dato da profili di concentrazione non omogenei, normalmente periodici, per i due morfogeni (Figura 4A), i cui dettagli strutturali dipendono dai parametri del sistema. Tale modello fornì le basi per lo studio di sistemi oggi chiamati reazione-diffusione oltre che introdurre una prima definizione di morfogeno come un segnale chimico diffusivo che permette la comunicazione intercellulare e la differenziazione stessa delle cellule.

Al contempo, intorno agli anni '60 dello scorso secolo, Lewis Wolpert introdusse invece un altro modello, conosciuto come French Flag Model (Figura 4B). Nel French Flag Model, la definizione di morfogeno è quella di una sostanza chimica che funge come segnale intercellulare e che permette di indurre determinate risposte cellulari a seconda della sua stessa concentrazione.

Osservando l'esempio della Figura 4B, un determinato morfogeno presente in alta concentrazione pertanto attiverà la risposta del 'gene blu, in media concentrazione del gene bianco ed in bassa concentrazione del gene rosso o di nessun gene. Nel French Flag Model la presenza del pattern del morfogeno fornisce alle cellule un'informazione posizionale che ne permette quindi il differenziamento.

I due meccanismi di Turing e Wolpert sono a priori profondamente differenti se non addirittura contrastanti. Il primo caso, infatti, richiede la presenza di due sostanze, chimicamente interagenti e con proprietà di trasporto profondamente diverse: in questo caso il *pattern* necessario per il differenziamento emerge spontaneamente e risulta essere periodico.

Il secondo caso prevede, invece, la combinazione di un solo gradiente di concentrazione, prestabilito all'interno dell'organismo (*prepattern*) o semplicemente indotto dalla presenza di una sorgente di morfogeno sottoposto a diffusione e degradazione, e di meccanismi a soglia nell'attivazione dell'espressione genica cellulare.



**Figura 4:** A) Schematizzazione del meccanismo à la Turing per la formazione di pattern spontanei. In ogni cellula è definito un circuito formato da un attivatore *A* ed un inibitore *B* rispettivamente segnali a corto e lungo raggio. Una soluzione tipica è quella ad onda, la cui lunghezza d'onda tipica dipende dai parametri del sistema. B) Schematizzazione del modello French flag. La presenza di un gradiente di morfogeno *M* nel tessuto (curva arancione) attiva diversi geni nelle cellule a seconda della sua stessa concentrazione. Le soglie di attivazione per i diversi geni sono indicate con  $M_1$  ed  $M_2$  rispettivamente.

Oggi giorno, grazie agli avanzamenti tecnologici nell'ambito della microscopia e nello studio del trasporto molecolare, si inizia a comprendere che i due modelli possono entrambi essere presenti, se non addirittura in combinazione. Esempi possono essere trovati nella coppia Nodal/Lefty presente nello zebrafish, nella rana e nel topo, che permette all'organismo di stabilire l'asimmetria sinistra/destra, nella coppia Dpp/Notch e EGF, fondamentale per la formazione delle vene nell'ala del moscerino della frutta o Wnt e Dkk per la rigenerazione in Hydra.

Moltissimi altri esempi sono presenti in letteratura [23, 24] e molti altri sono ancora da scoprire. Fra le domande fondamentali, tuttavia, bisogna ancora comprendere nella maggior parte dei casi quali siano le specie molecolari determinanti per la formazione degli organi nei diversi organismi, fino a che punto i meccanismi sovracitati possano effettivamente giocare un ruolo in maniera esclusiva nella formazione di pattern, come tali pattern si riorganizzano in maniera così precisa nel tempo e nello spazio nonostante la presenza di sorgenti di rumore e come altri processi possano contribuire alla formazione di *pattern*.

Infatti, i due meccanismi di Wolpert e Turing prevedono principalmente una sola forma di trasporto, la diffusione. In alcuni casi, come nelle piante, il trasporto di morfogeni può avvenire attivamente ovvero per azione di recettori di membrana dedicati a mediare il flusso uscente o entrante; nel caso animale, invece, la presenza di movimenti macroscopici morfogenetici può indurre nello spazio intercellulare fenomeni quali

l'avvezione, ovvero trasporto di molecole all'interno di un fluido a causa del movimento del fluido stesso. Comprendere quali siano effettivamente i diversi fenomeni di trasporto in gioco e capire come questi partecipino esattamente al differenziamento cellulare resta quindi un'altra delle grandi domande a cui rispondere.

## Selezione del modello e inferenza di reti geniche e di segnale

I quadri teorici ed i modelli illustrati finora sono semplificazioni dei fenomeni naturali il cui fine è quello di isolare porzioni dei sistemi di interesse per carpirne le proprietà intrinseche di funzionamento. Tuttavia, vista la complessità di tali sistemi, una domanda sorge spontanea: il modello che viene definito è davvero il modello migliore per poter studiare le proprietà di interesse del sistema sotto osservazione? Ovvero le variabili prese in considerazione *a priori* sono effettivamente le variabili più rilevanti per uno studio ottimale del sistema di interesse o per rispondere alla domanda che ci si è posti? Quali sono le interazioni fra tali variabili?

Chiaramente, vista l'enorme quantità di macromolecole, non si conosce a tutt'oggi l'intero quadro di interazioni intra- ed inter- cellulari né tantomeno il loro ruolo nell'espletare le diverse funzioni vitali. Tuttavia, l'identificazione delle variabili rilevanti e delle loro interazioni in un determinato contesto biologico costituisce un pun-

to fondamentale da chiarire per poter formulare modelli coerenti ed esaustivi.

A tal proposito, nel caso della ricostruzione delle reti molecolari, gli enormi progressi nel sequenziamento genico nel corso degli ultimi decenni hanno permesso di avviare un'intera branca di sviluppo di metodi finalizzati ad identificare la struttura di diverse reti biologiche, tra cui le reti geniche e di segnale. L'idea fondamentale dietro questi metodi è quella di identificare le osservabili maggiormente correlate in un certo insieme di dati ed estrarre, o meglio inferire, le interazioni fra tali osservabili in modo da riprodurre le correlazioni osservate. Vista la numerosa quantità di tecniche disponibili, i dati a disposizione in questo caso sono moltissimi e normalmente corrispondono a quantificazioni di trascritti e proteine in svariate cellule.

Per analizzare questo tipo di dati ed estrarre la struttura della rete genica sottostante, si è dapprima partiti con l'analisi delle componenti principali (PCA), già applicata in precedenza a svariati altri ambiti, dai mercati finanziari alle neuroscienze per citare alcuni esempi. Tale tecnica si basa sullo studio delle proprietà spettrali della matrice di correlazione a due punti ed è pertanto limitante nel contesto di interazioni a molti corpi.

A questo riguardo, sono state sviluppate ed applicate altre tecniche più raffinate, che tengono conto delle forti non linearità delle interazioni all'interno di questi sistemi e ricadono per l'appunto nella classe dei metodi non lineari di inferenza, come ad esempio:

- **Diffusion map:** è un metodo che modella i dati come il risultato di un cammino aleatorio markoviano (ogni *step* del cammino dipende unicamente dallo *step* precedente) e, tramite decomposizione spettrale, permette di ridurre la dimensionalità dei dati e risulta essere un equivalente non lineare di PCA. Tale metodo è stato ad esempio applicato per comprendere le traiettorie di differenziazione delle cellule staminali embrionali nel topo e in dati ottenuti da embrioni umani nella fase di pre-impianto [25]. Inoltre, recentemente, una variante di questo metodo è stata implementata nel contesto dell'algoritmo MAGIC (acronimo di Markov Affinity-based

Graph Imputation of Cells) ed applicata a diverse cellule tra cui quelle della corteccia cerebrale, dell'ippocampo e del midollo osseo nel topo [26], risultando più performante della precedente nel ricostruire le interazioni tra i geni.

- **t-Stochastic Neighbour Embedding (t-SNE):** introdotto da [27] e finalizzato anch'esso a ridurre la dimensionalità del sistema, t-SNE è un metodo che minimizza la distanza di Kullback-Leibler fra le due distribuzioni di probabilità  $p_{ij}$  e  $q_{ij}$  (dove  $i$  e  $j$  corrispondono agli indici delle variabili osservate), rispettivamente costruite nello spazio completo di partenza ed in quello ridotto di arrivo. Tale metodo è stato applicato ad esempio nello studio delle differenze fra cellule cancerogene e sane del midollo osseo [28].
- **Uniform Manifold Approximation and Projection (UMAP):** è un metodo che costruisce a partire dai dati un grafo di partenza definito su una varietà riemanniana, per definizione corrispondente ad uno spazio  $n$ -dimensionale curvo, e minimizza l'entropia incrociata fra lo spazio di partenza ad alta dimensionalità e quello ridotto a bassa dimensionalità. In [29] ad esempio, UMAP è stato utilizzato per ricostruire le traiettorie di differenziazione cellulare nel verme nematode a partire dalla sua espressione genica.

Nonostante questi metodi si siano già dimostrati essere un ottimo punto di partenza per ricostruire la mappa delle interazioni geniche, presentano delle limitazioni a causa di una serie di problematiche, come ad esempio la struttura intrinseca ed ancora non perfettamente sotto controllo dei dati utilizzati.

Un esempio è quello dei dati di RNA-seq, i quali rappresentano le letture di determinati RNA (precursori delle proteine) all'interno di determinate cellule. Tali dati normalmente appaiono in forma tabulare dove le righe della tabella rappresentano le diverse cellule campionate, le colonne rappresentano i diversi RNA letti dall'apparato sperimentale e ciascuno elemento della tabella è il numero di letture di un dato RNA in una data cellula.

Tuttavia, a causa dell'enorme quantità e varietà di RNA all'interno delle cellule e delle limitazioni sperimentali, questo tipo di dati possono essere inficiati da una serie di fattori quali: i) *undersampling*, dato che il numero di RNA campionati è normalmente molto maggiore del numero di cellule osservate; ii) *overdispersion*, a causa dell'alta variabilità nell'espressione dei diversi trascritti; iii) *zero inflation*, ovvero una sovrarappresentazione di entrate nulle a causa di mancate letture ed di eventi trascrizionali intermittenti (anche detti *transcriptional bursts*); iv) gli effetti di *hidden variables*, ossia di agenti che influenzano il comportamento del sistema ma che non vengono osservati o non è possibile osservare.

Questi ultimi effetti possono risultare in pratica da livelli della rete di cui il dataset in considerazione non tiene conto, come ad esempio il proteoma nel caso dei dati di RNA-seq. Tale punto richiede quindi lo sviluppo di metodi che possano integrare diversi livelli di complessità e strutturali all'interno delle cellule.

Di conseguenza, prima di poter applicare un qualsiasi metodo di inferenza, è opportuno comprendere la struttura dei dati di interesse ed il loro processo generativo sottostante, come ad esempio è stato recentemente proposto in [30].

Inoltre, lo sviluppo recente dei metodi di *deep learning* può permettere di controllare alcuni dei fattori sovraelencati ed in particolare di tenere conto delle forti non linearità presenti all'interno di questi sistemi.

Recentemente tali metodi sono infatti stati applicati con successo allo studio e predizione della struttura 3D delle proteine a partire dalla loro sequenza primaria, come in [31], ed in via di sviluppo, insieme ad altri metodi, nel caso dei repertori immunitari (cfr. "Il sistema immunitario attraverso la lente dell'inferenza statistica" di Barbara Bravi in questo numero di Ithaca).

Questi risultati, insieme al continuo e rapido sviluppo dell'intelligenza artificiale, aprono numerose nuove strade e possibilità nella comprensione della struttura delle macrounità che regolano il corretto funzionamento e perciò la vita delle nostre cellule.

## L'altro lato della medaglia: generazione e trasmissione di forze meccaniche

Finora si è parlato della Morfogenesi in termini principalmente biochimici, ovvero di reazioni chimiche a livello cellulare e di tessuto. Tuttavia, per poter espletare le funzioni vitali, gli organismi devono al contempo crescere.

La crescita di un organismo prevede chiaramente replicazione e, nella maggior parte dei casi, riorganizzazione cellulare che di conseguenza inducono deformazioni cellulari all'interno dei tessuti.

Questo tipo di fenomeni richiede pertanto la generazione e trasmissione di forze meccaniche a livello microscopico della cellula e macroscopico dei tessuti che permettano alle singole cellule di cambiare forma per riadattarsi all'ambiente circostante o modificarlo.

Dopo circa un secolo di esperimenti, sappiamo oggi che le proprietà meccaniche delle cellule sono regolate da una struttura biopolimerica detta **citochescheletro**, composta da diversi tipi di proteine filamentose, ovvero: i) i filamenti intermedi; ii) i microtubuli; iii) i filamenti di actina.

Mentre i filamenti intermedi si pensa entrino in gioco principalmente nelle deformazioni su grande scala spaziale e temporale e contribuiscono passivamente alle proprietà meccaniche della cellula, i microtubuli ed i filamenti di actina sono strutturalmente e, di conseguenza, meccanicamente, profondamente diversi. Essi infatti sono strutture semi-rigide lineari, polari e che pertanto definiscono una direzionalità per i processi attivi che avvengono su di essi, come ad esempio l'idrolisi dell'adenosina trifosfato, il trasporto di proteine che trasducono energia in lavoro meccanico come le miosine e chinesine ed il *treadmilling*, ovvero la polimerizzazione ad un estremo e depolimerizzazione all'altro estremo del filamento. Pertanto risultano essere strutture fortemente fuori equilibrio.

La maggior parte delle proprietà meccaniche delle cellule animali<sup>2</sup> è regolata da una sottile

<sup>2</sup>Nelle piante, una struttura molto più rigida e pertanto meno deformabile è presente, la parete cellulare. Le deformazioni vegetali quindi risultano essere profondamente differenti rispetto a quelle animali. Normalmente

rete formata dalle proteine sovraelencate e chiamata corteccia cellulare, che, essendo formata da un reticolo di polimerico immerso nel citoplasma cellulare, dal punto di vista meccanico si comporta come un gel, ovvero come un solido su scale di tempo brevi (ordine di secondi) e come un fluido su scale di tempo lunghe (ordine di minuti).

Tuttavia, tale gel è fortemente fuori equilibrio a causa delle proprietà elencate precedentemente che lo rendono in grado di trasdurre energia chimica in lavoro meccanico. Pertanto, bisogna introdurre una descrizione reologica fuori equilibrio, come la teoria dei gel attivi [32].

Una volta identificate le variabili rilevanti e le quantità conservate (come massa e quantità di moto), diversi approcci possono essere utilizzati per ottenere le equazioni costitutive del materiale, tra cui quello di perturbare il sistema dallo stato di equilibrio ed utilizzare il formalismo della termodinamica fuori equilibrio oppure costruire una teoria microscopica che tenga conto di tutte le scale di tempo e spaziali coinvolte in questi sistemi.

Ad ogni modo, lo scopo finale è di ottenere la relazione fra lo *stress* nel gel e le forze generalizzate. In questo tipo di sistemi, tale relazione è modificata da un termine che non deriva da alcun potenziale e denominato effettivamente *stress* attivo, che tiene conto della capacità di questi sistemi di trasdurre energia chimica in lavoro meccanico e pertanto autopropulsione. Identificando con  $v_i$  la componente della velocità del centro di massa nella direzione  $i$ , le relazioni costitutive per il tensore *stress*,  $\sigma_{ij}$ , in tale fluido attivo risultano pertanto essere date da:

$$(1 + \tau_M \frac{D}{Dt}) (\sigma_{ij} - \sigma_{ij}^a) = 2\eta (\partial_i v_j + \partial_j v_i - \frac{2}{3} \partial_k v_k \delta_{ij}) + \bar{\eta} \partial_k v_k \delta_{ij},$$

dove:

---

vengono infatti descritte come deformazioni elastiche principalmente governate da differenze di pressione indotte da flussi di acqua (pressione di turgore) e seguite da riorganizzazioni cellulari minime o completamente assenti. Di conseguenza, la descrizione meccanica per le piante prevede semplicemente delle modificazioni delle equazioni costitutive del materiale rispetto al caso animale.

- $\tau_M$  è il cosiddetto tempo di Maxwell, ovvero il tempo tipico di rilassamento elastico. Pertanto tale termine domina nel regime di tempi brevi;
- $\eta$  e  $\bar{\eta}$  sono rispettivamente la viscosità di *shear* e di *bulk* del fluido (ovvero associate allo sforzo di taglio ed allo sforzo normale). Tali termini dominano nel regime di lunghi tempi. La viscosità di *bulk*, in particolare, è associata alla conservazione del volume che a sua volta determina la pressione all'interno del fluido;
- $\sigma_{ij}^a$  è la parte attiva del tensore *stress*, che modella i fenomeni di trasduzione di energia chimica in lavoro meccanico. Tale parte attiva può essere di diversa forma a seconda del sistema in considerazione e del livello di dettaglio della descrizione e può in generale contenere un termine isotropo (tensione superficiale, autopropulsione cellulare, ecc.) ed uno o più termini anisotropi a seconda delle proprietà del sistema (nematico, polare o chirale).

Tali equazioni sono il punto di partenza per imporre il bilancio delle forze nel sistema, ovvero l'equilibrio meccanico fra forze interne e forze esterne (come ad esempio l'attrito con il mezzo circostante). Corredando tale sistema di equazioni con la conservazione della materia, si ottengono le leggi finali che regolano la meccanica di tali materiali attivi viscoelastici.

Grazie alla teoria dei gel attivi e la sua comparazione con esperimenti combinanti misure meccaniche *in vivo* ed *in vitro* (come ad esempio *laser cuts*, *micropipette aspiration*, *atomic force microscopy*), manipolazioni genetiche e microscopia ad alta risoluzione [35], nel corso dell'ultimo decennio, è stato possibile ricostruire il diagramma di forze che determina la dinamica e la statica della divisione cellulare e dei movimenti morfogenetici nello sviluppo iniziale dello zebrafish, per citare alcuni esempi.

Ovviamente tale descrizione continua e basata sull'assunzione di gel attivo è motivata dal dettaglio di interesse e dalle osservazioni reologiche sull'organismo (o più in generale sul materiale) in considerazione. Nei casi in cui, ad esempio, si è interessati a comprendere anche

fenomeni di riorganizzazione cellulare collettiva e direzionale, bisogna utilizzare altri approcci, come ad esempio la teoria dei cristalli liquidi [33], o passare ad una descrizione discreta, come nel caso del *Vertex Model* [34], tra le cui numerose applicazioni appaiono per l'appunto lo studio delle proprietà dinamiche di riorganizzazione cellulare dei tessuti e meccaniche dei movimenti morfogenetici.

Tra le sfide attuali vi sono la necessità di ottenere una teoria consistente che possa permettere di relazionare i parametri macroscopici delle equazioni costitutive ai meccanismi microscopici di interazione cellulare, comprendere la dipendenza della parte attiva dello stress dai processi chimici sottostanti e pertanto adottare descrizioni che includano una dipendenza dei parametri meccanici dallo spazio, dal tempo o dalle deformazioni stesse del materiale (come nel caso di materiali non-newtoniani) [36] e riuscire a misurare le differenti forze che agiscono in questi sistemi singolarmente.

## Conclusioni

Grazie alla cooperazione fra diverse discipline, è stato possibile nel corso dell'ultimo secolo aprire una nuova strada per una comprensione quantitativa di fenomeni naturali come la Morfogenesi.

Nel corso di questo articolo, si è visto come la complessità di tale processo coinvolga meccanismi a diverse scale temporali e spaziali nonché branche della fisica e della matematica normalmente studiate in maniera separata. La comprensione profonda di tali meccanismi, pertanto, non richiede soltanto lo sviluppo di nuovi strumenti o metodi, ma di un intero quadro teorico che possa essere in grado di descrivere consistentemente dei sistemi fortemente fuori equilibrio e multiscala. I sistemi biologici quindi diventano un naturale laboratorio di sperimentazione e validazione per tali teorie.

Grazie ai recenti sviluppi sperimentali, illustrati nell'articolo in ciascun paragrafo, al giorno d'oggi si ha accesso ad una serie di misurazioni e manipolazioni biofisiche, biochimiche e genetiche ad altissima risoluzione rispetto al secolo scorso che permettono di investigare più in dettaglio i meccanismi fondamentali alla ba-

se dello sviluppo degli organismi e soprattutto rappresentano un'ispirazione e motivazione per mettere in discussione le teorie già esistenti e formularne di nuove.

Da questa panoramica generale si può evincere quindi come lo studio di fenomeni considerati normalmente oggetto di interesse unicamente per la biologia e per la medicina possa essere fonte di grande interesse anche per la fisica e per la matematica e necessiti di queste discipline per poter assumere un carattere non solo quantitativo ma anche predittivo. In quest'ottica, le diverse discipline si devono fondere per adempiere al fine ultimo di comprendere in dettaglio i fenomeni naturali, ovvero la fisica, secondo il metodo scientifico.

## Ringraziamenti

L'autrice ringrazia Ada Altieri, Barbara Bravi, Ulisse Ferrari, Raimondo Grigolon e Valentina Ros per la loro lettura attenta e critica del manoscritto.



- [1] D'Arcy Wentworth Thompson, *On Growth and Form*, Cambridge University Press, 1917.
- [2] A. M. Turing, *The Chemical Basis of Morphogenesis*, Phil. Trans. R. Soc. London B, 237 (1952) 641.
- [3] R. Milo et al., *Network Motifs: Simple Building Blocks of Complex Networks*, Science, 298 (2002) 824-827.
- [4] R. Perez-Carrasco, P. Guerrero, J. Briscoe, K. M. Page, *Intrinsic noise profoundly alters the dynamics and steady state of morphogen-controlled bistable genetic switches*, PLoS Computational Biology, 12 (2016) 10, e1005154.
- [5] S. Grigolon, B. Bravi, O. C. Martin, *Responses to auxin signals: an operating principle for dynamical sensitivity yet high resilience*, Royal Society Open Science, 5 (2018) 1, 172098.
- [6] N. G. van Kampen, *Stochastic processes in physics and chemistry*, North-Holland Personal Library, 1981.
- [7] U. C. Täuber, M. Howard, and B. P. Vollmayr-Lee, *Applications of field-theoretic renormalization group methods to reaction-diffusion problems*, Journal of Physics A: Mathematical and General, 38 (2005), 17, R79.
- [8] P. François & V. Hakim, *Design of genetic networks with specified functions by evolution in silico*, PNAS, 101(2004) 2, 580.
- [9] E. Azpeitia et al., *The combination of the functionalities of feedback circuits is determinant for the attractors? number and size in pathway-like Boolean networks*, Scientific Reports, 7(2017) 42023.

- [10] R.C. Lee, R. L. Feinbaum, V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14*, Cell, 75 (1993), 843-854.
- [11] M. Osella, C. Bosia, D. Corà, M. Caselle, *The Role of Incoherent microRNA-Mediated Feedforward Loops in Noise Buffering*, PLoS Computational Biology, 7 (2011), 3, e1001101.
- [12] S. Grigolon, F. Di Patti, A. De Martino, E. Marinari, *Noise processing by microRNA-mediated circuits: the incoherent feed-forward loop, revisited*, Heliyon, 2 (2016), 4, e00095.
- [13] C. Bosia, A. Pagnani, R. Zecchina, *Modelling competing endogenous RNA networks*, PLoS One, 8 (2013), 6, e66609.
- [14] M. Figliuzzi, A. De Martino, E. Marinari, *microRNAs as a selective channel of communication between competing RNAs: a steady-state theory*, Biophysical Journal, 104, (2013), 5, 1203-1213.
- [15] M. Del Giudice, S. Bo, S. Grigolon, C. Bosia, *On the role of extrinsic noise in microRNA-mediated bimodal gene expression*, PLoS Computational Biology, 14 (2018), 4, 1-26.
- [16] V. Siciliano et al., *MiRNAs confer phenotypic robustness to gene networks by suppressing biological noise*, Nature Communications, 4 (2013), 1, 1-7.
- [17] C. Bosia et al., *RNAs competing for microRNAs mutually influence their fluctuations in a highly non-linear microRNA-dependent manner in single cells*, Genome Biology, 18 (2019), 1, 1-14.
- [18] L. R. Band et al., *Root gravitropism is regulated by a transient lateral auxin gradient controlled by a tipping-point mechanism*, PNAS, 109 (2012), 12, 4668-4673.
- [19] A. A. Hyman, C. A. Weber, F. Jülicher, *Liquid-liquid phase separation in biology*, Annual Review of Cell and Developmental Biology, 30 (2014), 39-58.
- [20] C. P. Brangwynne et al., *Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation*, Science, 324 (2009), 5935, 1729-1732.
- [21] C. F. Lee et al., *Spatial Organization of the Cell Cytoplasm by Position-Dependent Phase Separation*, Physical Review Letters, 111 (2013), 088101.
- [22] A. Klosin et al., *Phase separation provides a mechanism to reduce noise in cells*, Science, 367 (2020), 6476, 464-468.
- [23] A. D. Lander, *Pattern, growth and control*, Cell, 144 (2011), 955.
- [24] S. Kondo & T. Miura, *Reaction-Diffusion Model as a Framework for Understanding Biological Pattern Formation*, Science, 239 (2010), 1616.
- [25] L. Haghverdi, F. Buettner, F. J. Theis, *Diffusion maps for high-dimensional single-cell analysis of differentiation data*, Bioinformatics, 31 (2015), 18, 2989-2998.
- [26] D. van Dijk et al., *Recovering Gene Interactions from Single-Cell Data Using Data Diffusion*, Cell, 174 (2018), 716-729.
- [27] L. Van der Maaten & G. Hinton, *Visualizing Data using t-SNE*, Journal of Machine Learning Research, 9 (2008), 2579-2605.
- [28] E. D. Amir et al., *viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia*, Nature Biotechnology, 31 (2013), 545-552.
- [29] J. S. Packer et al., *A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution*, Science, 365 (2019), 6459.
- [30] G. Durif et al., *Probabilistic count matrix factorization for single cell expression data analysis*, Bioinformatics, 35 (2019) 20.
- [31] A. W. Senior et al., *Improved protein structure prediction using potentials from deep learning*, Nature, 577 (2020), 706-710.
- [32] J. Prost, F. Jülicher, J.-J. Joanny, *Active gel physics*, Nature Physics, 11 (2015), 111-117.
- [33] P.-G. de Gennes & J. Prost, *The physics of liquid crystals*, Oxford Science Publications (1974).
- [34] S. Alt, P. Ganguly, G. Salbreux, *Vertex models: from cell mechanics to tissue morphogenesis*, Philos. Trans. R. Soc. Lond. B., 19 (2017), 372(1720):20150520.
- [35] N. Mishra & C.-P. Heisenberg, *Dissecting Organismal Morphogenesis by Bridging Genetics and Biophysics*, Annual Review of Genetics, 55 (2021).
- [36] E. Hannezo & C.-P. Heisenberg, *Mechanochemical Feedback Loops in Development and Disease*, Cell, 178 (2019), 12-25.



**Silvia Grigolon:** si è laureata in Fisica Teorica a Sapienza Università di Roma. In seguito, ha ottenuto il dottorato in Fisica Teorica presso il Laboratoire de Physique Théorique et Modèles Statistiques dell'Université Paris-Saclay per poi proseguire con un postdottorato, dapprima al Francis Crick Institute di Londra ed in seguito al Laboratoire Jean Perrin presso la Sorbona a Parigi. Dal 2021, è ricercatrice al CNRS presso il Laboratoire Jean Perrin (Campus Jussieu) alla Sorbona. I suoi interessi di ricerca spaziano dallo sviluppo di metodi di inferenza alla modellizzazione mecano-chimica di sistemi biologici, in particolare delle prime fasi dello sviluppo embrionale animale e vegetale, in costante sinergia con gli sperimentali.