ISTITUTO LOMBARDO
ACCADEMIA di SCIENZE e LETTERE



Rendiconti

CLASSE DI SCIENZE MATEMATICHE E NATURALI

Vol. 153
(2019)

Istituto Lombardo di Scienze e Lettere

MILANO
pISSN 1974-6989
eISSN 2384-986X

ISTITUTO LOMBARDO - ACCADEMIA di SCIENZE e LETTERE

RENDICONTI

CLASSE DI SCIENZE MATEMATICHE E NATURALI

Vol. 153
(2019)

Istituto Lombardo di Scienze e Lettere

MILANO

ISTITUTO LOMBARDO - ACCADEMIA DI SCIENZE E LETTERE

RENDICONTI

CLASSE DI SCIENZE MATEMATICHE E NATURALI

Responsabile scientifico
Carlo Domenico Pagani

Comitato di redazione
Silvio Beretta
Stefano Maiorana
Cinzia Bearzot
Carlo Domenico Pagani

Pubblicato con il contributo di
Politecnico di Milano
Università Cattolica del Sacro Cuore, Milano
Università Commerciale L. Bocconi, Milano
Università degli Studi dell'Insubria, Varese
Università degli Studi di Milano
Università degli Studi Milano-Bicocca
Università degli Studi di Pavia



Associato all'Unione Stampa Periodica Italiana

pISSN 1974-6989
eISSN 2384-986X
Copyright 2020



PAGEPress Publications
via A. Cavagna Sangiuliani, 5
27100 Pavia, Italy
T. +39.0382.1549020
F: +39.0382.1727454

NOTA SULLA NON LINEARITÀ DEL SISTEMA RESPIRATORIO: COME L'EMOGLOBINA LEGA GASTON TISSANDIER A REINHOLD MESSNER

Nota del s.c. GUIDO FERRETTI (*)

(Adunanza del 24 gennaio 2019)

SUNTO. – In questo articolo si prende spunto da una tragedia aeronautica per tracciare sommariamente la storia della curva di dissociazione dell'emoglobina per l'ossigeno e della creazione del modello di Hill. Dopo aver argomentato che le caratteristiche di tale curva causano il comportamento non lineare del sistema respiratorio, vengono discusse alcune conseguenze funzionali di tale comportamento, che spiegano come fu possibile che Reinhold Messner sia riuscito a compiere la straordinaria imprevista impresa di scalare la vetta dell'Everest senza ossigeno supplementare.

***

ABSTRACT. – This article starts from an aeronautical tragedy to track briefly the history of the oxygen equilibrium curve and the creation of Hill's model of it. After having argued that the characteristics of such a curve determine the non-linear behaviour of the respiratory system, I discussed some functional consequences of this behaviour, which explain how it happened that Reinhold Messner succeeded in the extraordinary enterprise of climbing Mount Everest without supplementary oxygen.

1. UN DRAMMATICO INCIDENTE

La tragedia aeronautica che avvenne nei cieli di Parigi il 15 aprile 1875 (*Fig. 1*) fece grande scalpore in Francia. Essa costò la vita

(*) Dipartimento di Medicina Molecolare e Traslazionale, Università di Brescia, Italy. E-mail: guido.ferretti@unibs.it

a due dei tre aviatori che tentavano un'ascesa in mongolfiera ad alta quota, mentre il terzo, il pilota, Gaston Tissandier (1843-1899), che pure perse conoscenza, rinvenne durante la discesa e riuscì a pilotare l'atterraggio. L'altimetro a bordo indicò il raggiungimento della quota di 8600 metri. Lo scalpore dipese anche dalla rilevanza scientifica del volo e dal coinvolgimento dell'Académie de France. Prima del volo, gli avieri erano stati studiati nella camera ipobarica costruita dal celebre fisiologo e accademico Paul Bert (1833-1886), che giudicò il tentativo fattibile, benché avesse suggerito la presenza a bordo di fiasche di ossigeno dotate di boccaglio, da usare in caso di necessità. Ovviamente, i tre svennero prima di riuscire ad afferrare la fiasca.

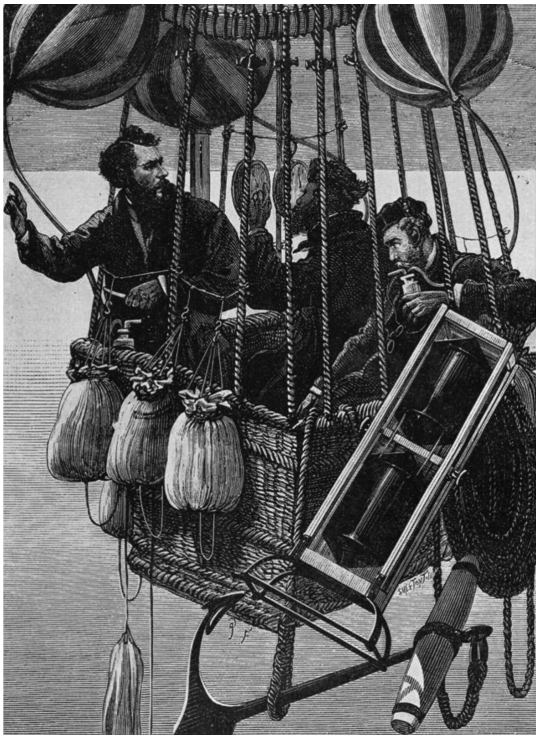


Fig. 1. Un'immagine del volo in pallone aerostatico del 15 aprile 1875, a Parigi, prima dell'incidente. Gaston Tissandier, capitano della navicella, unico sopravvissuto, è l'uomo a sinistra. Sul davanti della cesta si vedono le fiasche contenenti ossigeno, che Paul Bert volle venissero portate a bordo.

Paul Bert fu incaricato di studiare le cause della morte dei due avieri. I suoi primi risultati furono resi pubblici tre anni dopo, all'interno del suo celeberrimo volume intitolato *La Pression Barométrique* [1]. Egli infatti riuscì a misurare la concentrazione di ossigeno nel sangue di uomini esposti a gradi diversi di ipossia e costruì un diagramma (*Fig. 2*) in cui pose la concentrazione di ossigeno nel sangue in funzione della pressione parziale di ossigeno nell'aria inspirata. Egli dimostrò così che quella diminuiva in funzione di questa e concluse che la causa della morte degli avieri fu la scarsità di ossigeno. Il diagramma costruito da Bert appariva però sorprendente: benché la pressione parziale di ossigeno calasse più o meno in proporzione con la pressione barometrica, cioè con l'altitudine sul livello del mare, la diminuzione della concentrazione di ossigeno nel sangue non avveniva in proporzione ad esse, ma era modesta ad elevate pressioni parziali nell'aria inspirata, per divenire molto più intensa a basse pressioni parziali. Un elemento di non-linearità interferiva, ma quale?

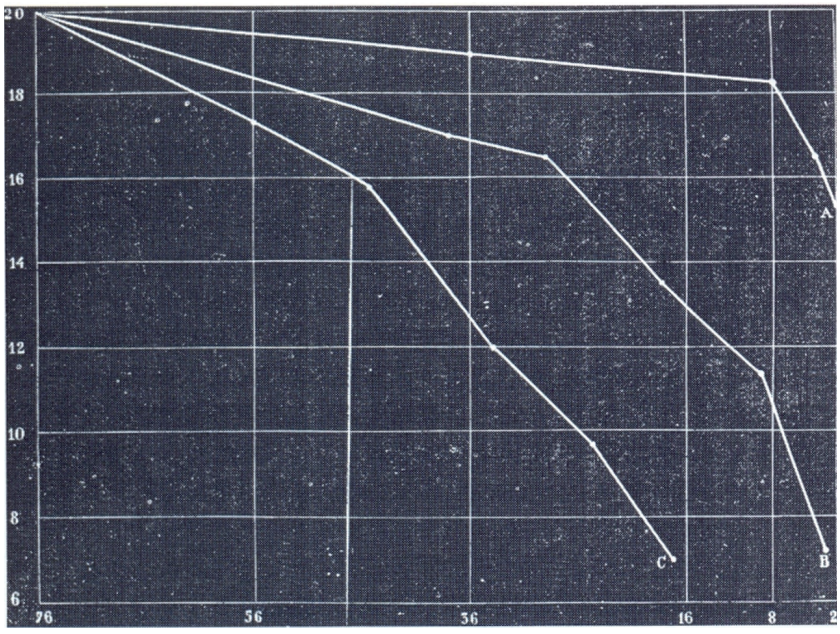


Fig. 2. La primitiva curva di dissociazione dell'emoglobina per l'ossigeno, costruita da Paul Bert a partire da dati di concentrazione e pressione parziale di ossigeno ottenuti in camera ipobarica, all'Académie de France.

2. Uno strano pigmento

Negli stessi anni, Felix Hoppe-Seyler (1825-1895), a Tubinga, isolò il pigmento del sangue, che egli chiamò **emoglobina**, e dimostrò che esso legava l'ossigeno. Alcuni anni dopo, a Copenhagen, Christian Bohr (1855-1911), al quale in séguito per questo fu assegnato il Premio Nobel per la Fisiologia o la Medicina, espose soluzioni di emoglobina a diverse pressioni parziali di ossigeno, costruendo così quella che oggi consideriamo la prima curva di dissociazione dell'emoglobina per l'ossigeno [2]: una curva quindi, non una retta, orizzontale ad elevata pressione parziale, e progressivamente più ripida man mano che la pressione parziale di ossigeno calava, alla quale Bohr diede una soluzione iperbolica. Ecco una possibile risposta al quesito di Bert sull'elemento di non-linearità. Fu quindi il successore di Hoppe-Seyler a Tubinga, Carl Gustaf von Hüfner (1840-1908), a sostenere la correttezza del risultato di Bohr anche sul piano teorico, in quanto compatibile con la legge di azione di massa per la reazione:

$$Hb + O_2 \leftrightarrow HbO_2 \qquad (1)$$

Il trattamento che Bohr e von Hüfner diedero della curva di dissociazione dell'emoglobina per l'ossigeno era simile a quello che successivamente Leonor Michaelis (1875-1949) e Maud Menten (1879-1960) avrebbero dato alle cinetiche enzimatiche.

Fu Bohr stesso a confutare [3] la sua ipotesi, e prima che Michaelis e Menten pubblicassero la loro teoria. Egli infatti pubblicò la curva di dissociazione dell'emoglobina determinata sul sangue intero, e non su soluzioni di emoglobina pura, dimostrandone la forma sigmoidale: una forma incompatibile sul piano teorico con l'ipotesi primigenia. L'anno successivo Bohr, coi suoi allievi Karl Albert Hasselbalch (1874-1962) e August Krogh (1874-1949), dimostrò che composti terzi (nel caso, il diossido di carbonio) avevano il potere di alterare la forma della curva di dissociazione dell'emoglobina per l'ossigeno [4], come mostrato in *Fig. 3*. In memoria del vecchio maestro, il fenomeno è stato tramandato sotto il nome di **effetto Bohr**. Fu John Barcroft (1872-1947) qualche anno dopo, a dimostrare che le alterazioni di forma della curva di dissociazione dovute all'effetto Bohr dipendevano dall'aggiunta o sottrazione di acidi [5]. Non fu però Barcroft, bensì il suo allievo Archibald Vivian Hill (1886-1977)

a dare una soluzione matematica elegante al concetto dell'effetto Bohr.

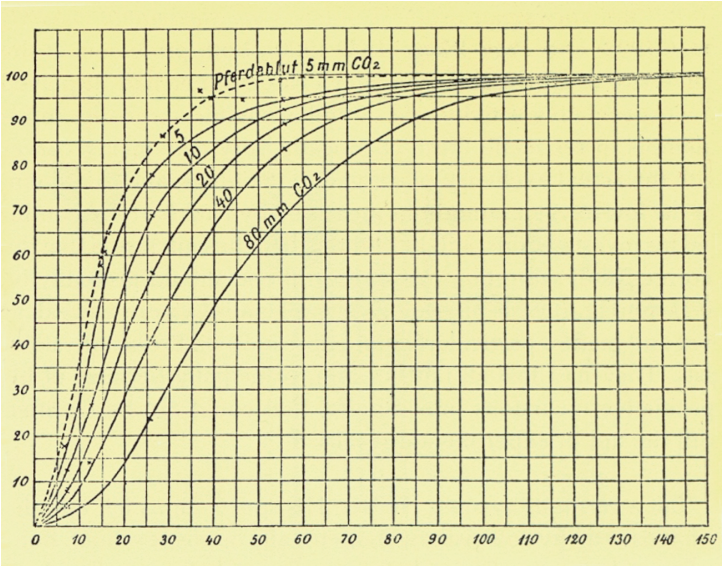


Fig. 3. Curve di dissociazione dell'emoglobina per l'ossigeno ottenute da Christian Bohr su sangue intero durante esposizione a diverse pressioni parziali di anidride carbonica. È questa la prima rappresentazione grafica dell'effetto Bohr.

3. Il modello e la congettura di Hill

Hill [6] propose che la curva di dissociazione dell'emoglobina riflettesse l'equilibrio della seguente reazione chimica tra emoglobina e ossigeno:

$$Hb + n\ O_2 \leftrightarrow Hb(O_2)_n \tag{2}$$

dove n indica il rapporto stechiometrico della reazione. Questa formulazione differisce da quella dell'equazione (1) per un piccolo dettaglio, però importante: questa ammette $n > 1$, quella no. Si noti che allora non si conosceva il valore di n, essendo la struttura chimica dell'emoglobina ancora ignota. Con la formulazione di Hill, la costante di equilibrio della reazione (K) è:

$$K = \frac{[HbO_2]^n}{[Hb^-][O_2]^n} \tag{3}$$

dove HbO_2 e Hb^- indicano rispettivamente l'emoglobina ossigenata e l'emoglobina non ossigenata (ridotta). Essendo l'emoglobina totale la somma di HbO_2 e Hb^-, possiamo definire la saturazione (S) dell'emoglobina per l'ossigeno come:

$$S = \frac{[HbO_2]}{[HbO_2]+[Hb^-]} \tag{4}$$

Considerato che pressoché tutto l'ossigeno presente nel sangue, in condizioni fisiologiche, sta legato all'emoglobina, e che HbO_2 sta in equilibrio con la pressione parziale di ossigeno (PO_2), Hill giunse, con una serie di passaggi algebrici successivi, alla seguente formulazione:

$$\frac{S}{(1-S)} = \Sigma = \left(\frac{PO_2}{P_k}\right)^n \tag{5}$$

dove P_k è una costante. Esprimendo l'equazione (5) in forma logaritmica, si ha:

$$\log \Sigma = n\ (\log PO_2 - \log P_k) \tag{6}$$

Pertanto, se mettiamo $\log \Sigma$ in funzione di $\log PO_2$ (*Fig. 4*), otteniamo relazioni lineari, in cui il coefficiente angolare è uguale a n e l'intercetta sulla ascisse è uguale a $\log P_k$. Per quanto riguarda n, Hill congetturò che esso corrispondesse all'equilibrio stechiometrico della reazione descritta dall'equazione (2), cioè al numero di molecole di ossigeno in grado di legarsi a una molecola di emoglobina. Egli ottenne $n = 1$ per la mioglobina (il pigmento, simile all'emoglobina, che troviamo nelle fibre muscolari) e $n = 2.8$ per l'emoglobina. Inoltre, nella rappresentazione grafica dell'equazione (6), le curve di dissociazione dell'emoglobina riprodotte in *Fig. 3*, che Bohr ottenne variando la concentrazione (o la pressione parziale) di CO_2, si traducono in una famiglia di rette parallele (n invariante ed indipendente da tutti gli altri termini dell'equazione 6), spostate progressivamente verso destra man mano che la concentrazione di CO_2, e quindi $\log P_k$, aumenta. Da qui deriva il concetto di spostamen-

to a destra (o a sinistra) della curva di dissociazione dell'emoglobina per l'ossigeno, di cui ancora oggi si parla agli studenti e di cui molto si abusa in àmbito clinico. Per parte sua, il valore di log P_k corrisponde a quello che log PO_2 assume quando log Σ è uguale a 0, e quindi Σ è uguale a 1. Ciò si verifica, secondo l'equazione (5), quando $S = 0.5$. Pertanto Hill definì la costante P_k come P_{50}, vale a dire la PO_2 corrispondente a $S = 0.5$. Un aumento di P_{50} corrisponde sempre a uno spostamento a destra della curva di dissociazione dell'emoglobina per l'ossigeno. Questi concetti sono validi solo all'interno del modello di Hill, e quindi parlare di «spostamento a destra» o di P_{50} comporta accettazione implicita del modello all'interno dei quali sono stati creati.

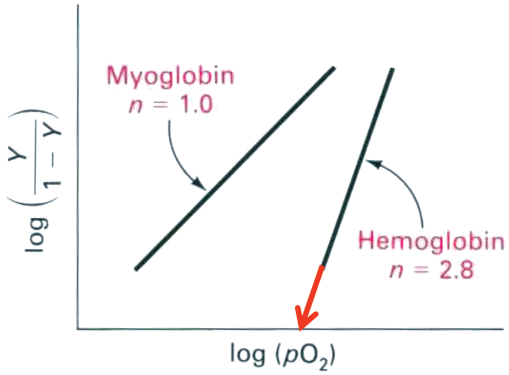


Fig. 4. Rappresentazione semilogaritmica del modello di Hill della curva di dissociazione dell'emoglobina per l'ossigeno, con indicazione della pendenza n*, che Hill congetturò corrispondesse al rapporto stechiometrico tra ossigeno ed emoglobina, per la mioglobina* (n = *1.0) e per l'emoglobina* (n = *2.8). Il secondo valore è indicativo di cooperatività. La freccia indica l'intercetta sull'asse delle ascisse, corrispondente a quella che Hill chiamò* P_{50}.

4. Confutazione e critica della congettura di Hill

La confutazione della congettura di Hill giunse quando Max Perutz (1914-2002) delucidò la struttura quaternaria dell'emoglobina, per cui ottenne il Premio Nobel per la Chimica nel 1962. Egli dimostrò che ogni molecola di emoglobina, composta di quattro apoproteine e quattro gruppi prostetici, era in grado di legare fino a quattro molecole di ossigeno [7], per cui $n = 4$ anziché 2.8. Per la mioglobina fu invece confermato $n = 1$. Dopo la confutazione della congettura di Hill, il mondo della fisiologia

respiratoria tese ad abbandonare il suo modello della curva di dissociazione, giudicandolo falso. Modelli teorici alternativi non furono però proposti. I pedanti, che non mancano mai, applicando anche tecniche analitiche sempre più precise, definirono modelli empirici polinomiali della curva di dissociazione dell'emoglobina sempre più complessi. Basta aumentarne i gradi, si troveranno sempre polinomi in grado di descrivere un andamento con precisione sempre maggiore. Tutti questi modelli, però, non avevano alcun legame teorico con modelli fisiologici soggiacenti: essi furono tutti esclusivamente descrittivi ed ebbero un impatto culturale infimo.

Benché confutato sul piano quantitativo riguardo al valore di n, il modello di Hill aveva un legame concettuale solidissimo con una realtà fisiologica soggiacente, forniva uno strumento di analisi quantitativa molto potente dell'effetto Bohr, e gli stessi pedanti lo considerano un descrittore sufficientemente preciso della curva di dissociazione, almeno nell'àmbito compreso tra $S = 0.2$ e $S = 0.98$. Credo che, invece di abbandonarlo come falso, sarebbe stato più prudente rivisitarlo, correggendone il significato, dopo la confutazione di Perutz: $n = 1$ (mioglobina, una sola apoproteina, con un solo gruppo prostetico) significa che le molecole di ossigeno si legano all'emoglobina in maniera indipendente l'una dall'altra; $n > 1$ significa presenza di cooperatività, che si ha quando le apoproteine (e i gruppi prostetici) sono più di uno, e quindi la proteina assume struttura quaternaria. Hill applicò l'equazione logaritmica all'intera gamma di valori di S: ciò comportò una distorsione della retta a causa dell'impatto dei valori di $S < 0.2$. La curva di dissociazione è in effetti sigmoide, causa allosteria. Pertanto il valore di 2.8 per n, riportato da Hill, potrebbe derivare dal fatto che l'emoglobina ha una difficoltà maggiore a tenere l'ossigeno legato quando predomina nettamente la forma ridotta.

E' comunque un fatto che il modello di Hill, o un altro modello teorico ad esso alternativo, è necessario per comprendere gli effetti che la curva di dissociazione dell'emoglobina per l'ossigeno ha sul comportamento non lineare del sistema respiratorio, e le conseguenze fisiologiche che da ciò derivano.

5. Comportamento non lineare del sistema respiratorio

Oggi si guarda al sistema respiratorio come a un sistema integrato in cui un gas fluisce lungo il percorso dalla bocca ai mitocondri (o viceversa) sospinto da un gradiente di pressione, necessario a vincere le

numerose resistenze in serie che si oppongono al flusso. La *Fig. 5* evidenzia le resistenze (o il loro reciproco, le conduttanze) identificate a varî livelli lungo il sistema e riporta le relative equazioni di conduttanza locali. La resistenza totale in un tale sistema è data dalla somma di tutte le resistenze parziali. Ogni resistenza costituisce quindi una frazione della resistenza totale al flusso di gas. In un sistema lineare, la frazione di resistenza totale rappresentata da ciascuna resistenza parziale è calcolabile semplicemente dal rapporto tra il gradiente di pressione che sostiene il flusso di gas a un dato livello e il gradiente di pressione totale lungo l'intero sistema. Questo principio vale per tutti i gas che nel sangue sono presenti solo in forma fisicamente disciolta, o la cui curva di dissociazione è, nell'àmbito di pressioni parziali considerate, lineare. Questo principio non vale però per il flusso di ossigeno, a causa della componente di non linearità introdotta dalla curva di dissociazione dell'emoglobina: l'equazione di conduttanza per l'ossigeno deve avere soluzione non lineare, qualunque modello si voglia applicare alla curva di dissociazione.

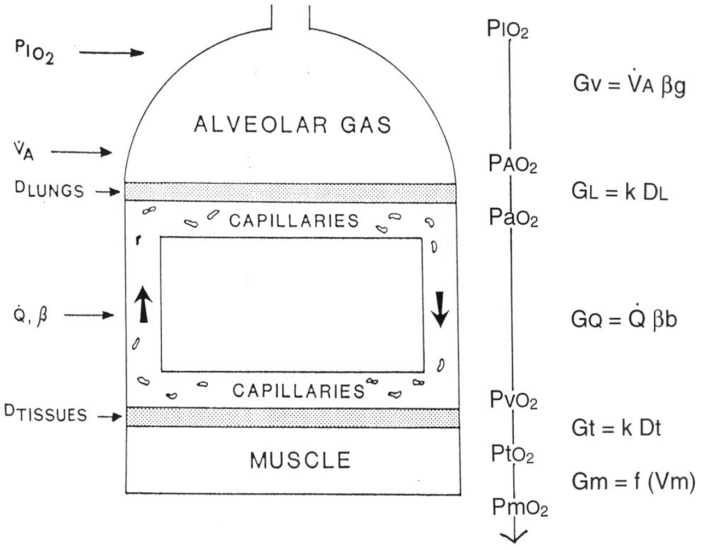


Fig. 5. Rappresentazione schematica del sistema respiratorio, con identificazione sia dei siti dove è possibile misurare la pressione parziale di ossigeno, sia delle diverse conduttanze di significato funzionale lungo il sistema, disposte in serie. Per ogni conduttanza, è indicata la relativa equazione di conduttanze. Di queste conduttanze, due (la conduttanza ventilatoria e la conduttanza cardiovascolare) risultano dal prodotto di due variabili fisiologiche, le altre sono proporzionali a variabili fisiologiche.

La prima soluzione non lineare dell'equazione di conduttanza dell'ossigeno fu proposta dal canadese Roy Shephard [8], ma passò rapidamente nell'oblio. Il principio però covò sotto la cenere del rogo delle idee e fu ripreso da Dick Taylor di Boston e Ewald Weibel di Berna come base teorica per il loro celebrato studio morfometrico sul sistema respiratorio dei mammiferi [9]. Da qui partì il socio corrispondente Pietro Enrico di Prampero per creare il primo modello multifattoriale dei fattori limitanti il massimo consumo di ossigeno nell'uomo all'esercizio [10], con cui egli riuscì a dimostrare l'inapplicabilità del principio di linearità al sistema respiratorio, e a stimare, in virtù di questo, che il 70% della limitazione del massimo consumo di ossigeno fosse da attribuirsi al trasporto cardiovascolare di ossigeno, anziché il 50% che si ottiene assumendo comportamento lineare del sistema. Tale stima ricevette poi solido sostegno sperimentale da parte di un gruppo coordinato dal sottoscritto [11].

6. Equazione di conduttanza per l'ossigeno a livello cardiovascolare e meccanismo delle risposte non lineari

Di Prampero e Ferretti [12] inferirono sul meccanismo con cui la non linearità della curva di dissociazione dell'emoglobina esplicava i suoi effetti sul massimo flusso di ossigeno (VO_2max) nel sistema respiratorio. Ammettiamo di ridurre improvvisamente e acutamente la resistenza al flusso di ossigeno nelle vie aeree, per esempio facendo respirare una miscela di gas contenente elio al posto dell'azoto. Ciò comporta un immediato rialzo della pressione parziale di ossigeno nell'aria alveolare, e così a valle, fino al sangue arterioso. Nel sangue arterioso però la concentrazione di ossigeno, che pure varia con la pressione parziale dello stesso, è dettata dalla forma della curva di dissociazione dell'emoglobina.

Nel modello delle conduttanze in serie, VO_2max a livello cardiovascolare è descritto dalla seguente equazione:

$$\dot{V}O_2max = \dot{Q}\,\beta b\,(PaO_2 - P\bar{v}O_2) \tag{7}$$

dove $\dot{Q}$ è il flusso ematico totale (gettata cardiaca) e βb è una costante definita *coefficiente di trasporto dell'ossigeno nel sangue*, mentre il loro prodotto rappresenta la conduttanza. La costante βb è uguale a:

$$\beta b = \frac{CaO_2 - C\bar{v}O_2}{PaO_2 - P\bar{v}O_2} \tag{8}$$

in cui il numeratore indica la differenza di concentrazione di ossigeno tra sangue arterioso (a) e sangue venoso misto ($\bar{v}$), mentre il denominatore corrisponde alla differenza di pressione parziale per il medesimo gas tra i due sangui. Come appare dalla *Fig. 6*, l'equazione (8) definisce βb come pendenza media della curva di dissociazione dell'emoglobina per l'ossigeno.

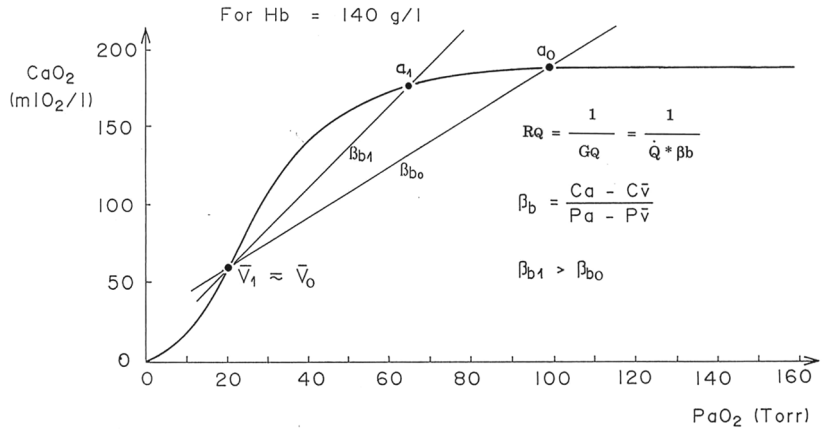


Fig. 6. Curva di dissociazione dell'emoglobina per l'ossigeno, con indicazione delle pendenze medie tra punto del sangue arterioso (a) e punto del sangue venoso misto (v)*, per casi in cui il punto* a *si situa sulla parte piatta della curva, oppure in prossimità della parte ripida. Sono indicate anche le definizioni di resistenza cardiovascolare (R_Q) e del coefficiente di trasporto dell'ossigeno nel sangue (β_b).*

Se PaO_2 aumenta in séguito ad una variazione di conduttanza ventilatoria, a tale aumento non consegue però una variazione di CaO_2, poiché a livello del mare operiamo sulla parte piatta della curva di dissociazione dell'emoglobina. Dal lato opposto, pressione parziale e concentrazione nel sangue venoso misto, se non variano, calano di poco. Pertanto, la pendenza media della curva, cioè βb, diminuisce. Ne conseguono la diminuzione della conduttanza e quindi l'aumento del suo reciproco, cioè la resistenza a livello cardiovascolare.

Ci troviamo pertanto in una condizione in cui, a séguito della variazione indotta di una conduttanza, quella ventilatoria, a causa della riduzione di densità della miscela respirata, si verifica necessariamente anche una variazione opposta di una seconda conduttanza, quella cardiovascolare, che tende a compensare ed annullare gli effetti della prima. L'esito è che una diminuzione di resistenza a livello ventilatorio, che naturalmente tenderebbe ad aumentare il VO_2max, non ha invece alcun effetto sul VO_2max a livello del mare [13]. Questo autorizza a concludere che, a causa della non linearità del comportamento del sistema respiratorio, le resistenze in serie situate a monte del trasporto cardiovascolare di ossigeno non contribuiscono a limitare il VO_2max. È per questo che di Prampero [10] ottenne una limitazione cardiovascolare del 70% anziché del 50%, come previsto da un modello lineare: egli si trovò ad analizzare un sistema dotato di solo due resistenze efficaci nel limitare il VO_2max, cioè in grado di ripartirsi la limitazione totale.

7. Massimo consumo di ossigeno in altitudine

Il VO_2max diminuisce in altitudine a causa della diminuzione della pressione barometrica. Numerosi studiosi, tra cui alcuni membri di questa Accademia [14, 15], riportarono, in circostanze diverse, che la caduta del VO_2max in altitudine è marcatamente non lineare: modesta a quote inferiori a 2000 metri, molto accentuata a quote superiori a 4000 metri.

La caduta del VO_2max in ipossia è rappresentata in *Fig. 7* per individui aventi a livello del mare un massimo consumo di ossigeno equivalente a quello medio della popolazione di sesso maschile non atletica [16]. Tale rappresentazione, teorica ma basata su una grande mole di dati sperimentali, equivale alla costruzione di una curva flusso-pressione: le rette convergenti sull'origine degli assi sono isoplete di conduttanza. Appare così chiaro che la non linearità della caduta del VO_2max in altitudine comporta un aumento della conduttanza totale del sistema respiratorio. Tale aumento appare come una conseguenza di un progressivo incremento di βb, cioè della conduttanza cardiovascolare [17].

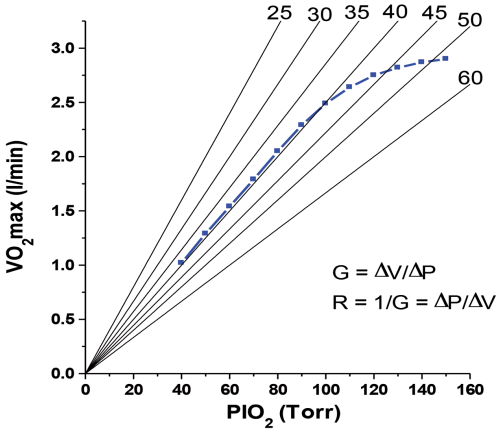


Fig. 7. La curva blu dimostra l'andamento della caduta del massimo consumo di ossigeno in ipossia. Le linee sottili nere, convergenti sull'origine degli assi, sono isoplete di conduttanza (o del reciproco di questa, la resistenza). Per ogni isopleta è indicato il corrispondente valore di resistenza. La forma della curva comporta calo di resistenza al flusso di ossigeno con il progressivo diminuire della pressione parziale di ossigeno nell'aria inspirata.

Ad altitudini inferiori ai 2000 metri, operiamo sulla parte piatta della curva di dissociazione dell'emoglobina per l'ossigeno. Pertanto ogni diminuzione della pressione barometrica, a cui si associa una diminuzione della PaO_2, non è seguita da una diminuzione della CaO_2. Così βb assume un valore progressivamente più elevato, che fa crescere la conduttanza cardiovascolare, riducendo pertanto la corrispettiva resistenza. Ciò mantiene il VO_2max più alto di quello che ci dovremmo attendere ad ogni altitudine in caso di comportamento lineare del sistema. Il fenomeno è tanto più accentuato quanto più intense sono le risposte ventilatorie all'ipossia [18]. È solo quando, salendo, si giunge ad operare sulla parte pendente della curva di dissociazione che la caduta del VO_2max diviene importante e tende a linearizzarsi (βb invariante). Wagner [19] dimostrò che il segmento di curva di dissociazione su cui si opera alle altitudini più estreme sulla Terra è a tutti i fini pratici indistinguibile da un segmento di retta. È pertanto verosimile che i due avieri del 1875 siano periti per aver raggiunto, in alta quota, una CaO_2 incompatibile con una sufficiente ossigenazione cerebrale, ma se ciò avvenne, fu ad una altitudine ben più elevata di quanto non ci si potesse immaginare allora.

8. Atleti in altitudine

Dempsey et al [20] furono i primi a riportare che atleti caratterizzati da elevati valori di $\dot{V}O_2max$ presentano desaturazione del sangue arterioso durante esercizio massimale a livello del mare. Questi autori attribuirono il fenomeno al fatto che tali atleti raggiungono anche valori di gettata cardiaca massima molto elevati, fino a 40 l min^{-1}, tali per cui il tempo di transito attraverso i capillari polmonari si riduce al punto di non consentire un'ossigenazione completa del sangue capillare.

Nel presente contesto tale osservazione assume valore in quanto ne consegue che durante esercizio massimale a livello del mare gli atleti operano già in prossimità della parte ripida della curva di dissociazione dell'emoglobina per l'ossigeno. Pertanto in essi il meccanismo, legato alle caratteristiche della curva di dissociazione, che mantiene il elevato a quote inferiori ai 2000 metri, non può funzionare.

Ne deriva che il $\dot{V}O_2max$ di atleti **deve** diminuire in altitudine in misura maggiore di quanto avviene in individui non atletici, a causa di un minor aumento della conduttanza cardiovascolare. Questa ipotesi, che ha ricevuto solida conferma sperimentale in ipossia acuta [21], ha come corollario che le naturali differenze di $\dot{V}O_2max$ tra la popolazione si riducono progressivamente e consistentemente col salire dell'altitudine, fino quasi ad annullarsi alla quota corrispondente alla cima dell'Everest.

9. Messner raggiunge la cima dell'Everest senza ossigeno supplementare

Il giorno 8 maggio 1978 è una data cardine della storia non solo dell'alpinismo, ma di tutta l'esplorazione della Terra da parte dell'uomo: per la prima volta due alpinisti, Reinhold Messner e Peter Habeler, raggiunsero la vetta più alta della Terra, il Monte Everest (8848 metri sopra il livello del mare) senza fare uso di bombole contenenti ossigeno supplementare. Il mondo della scienza riteneva l'impresa impossibile, anche se in realtà qualcuno aveva espresso per iscritto e in tempi non sospetti opinioni diverse ma neglette [22]. Messner rivoluzionò l'alpinismo himalayano e tracciò una via nuova all'esplorazione montana presto percorsa anche da altri. Al fine di capire meglio cosa avvenne o come l'impresa fu possibile, Paolo Cerretelli, oggi membro effettivo di

questa Accademia, e Hans Hoppeler di Berna organizzarono un ampio studio di valutazione funzionale di alpinisti che salirono vette di altezza superiore a 8500 metri sul livello del mare [23], tutti seguiti come medico da Oswald Oelz di Zurigo. L'assioma era che questi alpinisti dovessero avere caratteristiche fisiologiche uniche, da cui derivò l'ipotesi che il loro $\dot{V}O_2max$ fosse straordinariamente elevato a livello del mare. Questa ipotesi fu confutata sperimentalmente, tra lo sconcerto generale: il $\dot{V}O_2max$ degli alpinisti investigati era di gran lunga inferiore a quello degli atleti (*Fig. 8*). Non sapendo cosa dire, si concluse l'articolo con affermazioni metafisiche sul ruolo che potrebbe aver avuto la motivazione psicologica all'impresa: allora, non si conoscevano ancora gli effetti dell'effetto Dempsey sulla caduta del massimo consumo di ossigeno in altitudine, lo studio di Dempsey era stato pubblicato da troppo poco tempo. Oggi si sarebbe concluso che non serve avere massimi consumi di ossigeno straordinariamente elevati per raggiungere la cima dell'Everest senza ossigeno supplementare, a causa dell'effetto Dempsey e dell'effetto della non linearità del sistema respiratorio: il legame tra la tragedia del 1875 e il trionfo del 1978 è completo.

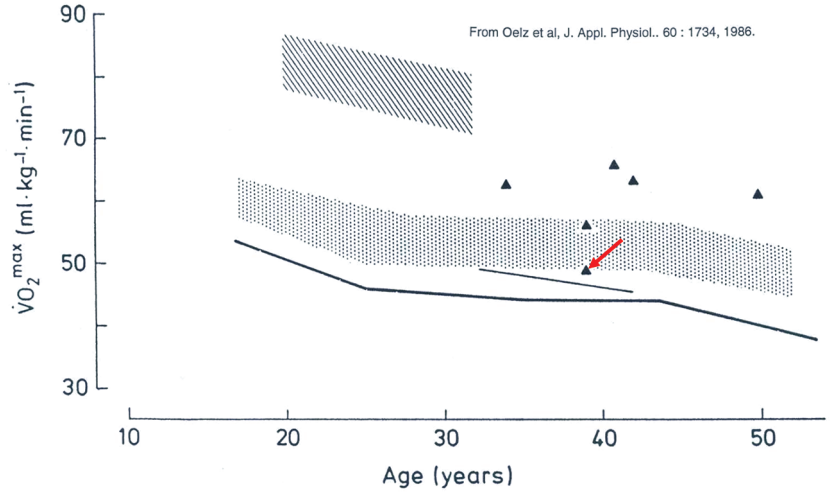


Fig. 8. Massimo consumo di ossigeno in funzione dell'età. La banda tratteggiata in alto a sinistra si riferisce ad atleti; la banda a puntini a non-atleti allenati. La traccia in basso definisce la normalità per il non atleta. I punti sperimentali riguardano sei alpinisti che hanno asceso le più alte vette della Terra senza ossigeno supplementare. La freccia indica il valore di Reinhold Messner.

10. Conclusioni

Esplorazione umana e conoscenza fisiologica seguono spesso percorsi paralleli, che talora s'incrociano. Fu un clamoroso incidente mortale avvenuto a Parigi nel 1875 a dare il via a studi che condussero a definire le caratteristiche del legame tra ossigeno ed emoglobina ed in séguito a delineare il comportamento non lineare del sistema respiratorio e a identificare alcune delle conseguenze funzionali che questo avrebbe avuto. È attraverso un tale percorso che si sono costruiti gli strumenti culturali per comprendere il significato fisiologico di uno straordinario evento di un secolo dopo. Molti uomini di scienza furono protagonisti di questo percorso, che unisce le vicende diverse di due avventurosi, Tissandier e Messner. È per me motivo di fierezza constatare il ruolo importante di membri di questa Accademia nella costruzione di questo percorso.

REFERENZE

1. Bert P, *La Pression Barométrique,* Masson, Paris, 1878.
2. Bohr C, *Experimentale Untersuchungen über die Sauerstoffaufnahme des Blutfarbstoffes*, OC Olsen & Co Buchdruckerei, Copenhagen, 1885.
3. Bohr C, Theoretische Behandlung der quantitativen Verhältnisse bei der Sauerstoffaufnahme der Hemoglobins, *Zentralbl Physiol*, 1903: 17: 682-691.
4. Bohr C, Hasselbalch KA, Krogh A, Über einen in biologischer Beziehung wichtigen Einfluss, den die Kohlensäurespannung des Blutes auf dessen Sauerstoffbindung übt, *Skand Arch Physiol*, 1904: 16: 402-414.
5. Barcroft J, *The Respiratory Function of the Blood*. Cambridge University Press, Cambridge, 1914.
6. Hill AV, The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves, *J Physiol Lond* 1910: 40: iv-vii.
7. Perutz MF, Stereochemistry of cooperative effects in haemoglobin, *Nature* 1970: 228: 726-733.
8. Shephard RJ, A non-linear solution of the oxygen conductance equation: applications to performance at sea level and at an altitude of 7350 ft, *Int Z angew Physiol*, 1969: 27: 217-225.
9. Taylor CR, Weibel ER, Design of the mammalian respiratory system. I. Problem and strategy, *Respir Physiol*, 1981: 44: 1-10.
10. di Prampero PE, Metabolic and circulatory limitations to at the whole animal level, *J Exp Biol*, 1985: 115: 319-331.
11. Bringard A, Pogliaghi S, Adami A, De Roia G, Lador F, Lucini D, Pizzinelli P, Capelli C, Ferretti G, Cardiovascular determinants of maximal oxygen con-

sumption in upright and supine posture at the end of prolonged bed rest in humans, *Respir Physiol Neurobiol*, 2010: 172: 53-62.
12. di Prampero PE, Ferretti G, Factors limiting maximal oxygen consumption in humans, *Respir Physiol*, 1990: 80: 113-128.
13. Esposito F, Ferretti G, The effects of breathing He-O_2 mixtures on maximal oxygen consumption in normoxic and hypoxic men. *J Physiol Lond*, 1997: 503: 215-221.
14. Cerretelli P, Margaria R, Maximum oxygen consumption at altitude, *Int Z angew Physiol*, 1961: 18: 460-464.
15. Cerretelli P, Gas exchange at high altitude, in: *Pulmonary Gas Exchange*, West JB editore, volume II, 97-147, Academic Press, New York, 1980.
16. Ferretti G, *Energetics of Muscular Exercise*, Springer, Heidelberg, 2015.
17. Ferretti G, Limiting factors to oxygen transport to Mount Everest 30 years after: a critique of Paolo Cerretelli contribution to altitude physiology, *Eur J Appl Physiol*, 2003: 90: 344-350.
18. Marconi C, Marzorati M, Grassi B, Basnyat B, Colombini A, Kayser B, Cerretelli P, Second generation Tibetan lowlanders acclimatize to high altitude more quickly than Caucasians, *J Physiol Lond*, 2004: 556: 661-671.
19. Wagner PD, A theoretical analysis of factors determining at sea level and altitude, *Respir Physiol*, 1996: 106: 329-343.
20. Dempsey JA, Hampson PG, Henderson KS, Exercise-induced arterial hypoxaemia in healthy human subjects at sea level, *J Physiol Lond*, 1984: 355: 161-175.
21. Ferretti G, Moia C, Thomet JM, Kayser B, The decrease of maximal oxygen consumption during hypoxia in man: a mirror image of the oxygen equilibrium curve, *J Physiol Lond*, 1997: 498: 231-237.
22. West JB, Alexander M, Kellas and the physiological challenge of Mt. Everest, *J Appl Physiol*, 1987: 63: 3-11.
23. Oelz O, Howald H, di Prampero PE, Hoppeler H, Claassen H, Jenni R, Bühlmann A, Ferretti G, Brückner JC, Veicsteinas A, Gussoni M, Cerretelli P, *J Appl Physiol*, 1986: 60: 1734-1742.

QUATTRO CHIACCHIERE A PROPOSITO DELLA COMUNICAZIONE DELLE SCIENZE. LA CHIMICA

Nota del m.e. ANGELO ALBINI (*)

(Adunanza del 14 marzo 2019)

SUNTO. – La completa separazione tra i linguaggi delle varie scienze rende impossibile comunicare tra specialisti di aree diverse. Il linguaggio della chimica si basa su nomi non ambigui, che mostrano come gli atomi sono disposti e fissati attraverso legami chimici (o covalenti) a formare molecole. Tali nomi suggeriscono in modo euristico le modalità di sintesi. Inoltre le equazioni chimiche indicano il meccanismo che porta da una situazione iniziale (reagenti) ad una finale (prodotti). In molecole molto grandi (etero ed omo polimeri) il legame chimico non è sufficiente a illustrarne la reale struttura, ma la conformazione nello spazio è in realtà già dettata dal modo in cui gli atomi sono legati. La risposta delle molecole a variazioni nell'ambiente sono adatte a un universo che si trasforma con passaggi dolci, non a scatti.

***

ABSTRACT. – The language used within different disciplines has become fully separated. The language of chemistry is based on chemical names, where it is shown how atoms are disposed through chemical or covalent bonds to yield molecules, On the other hand, assigning a name contains as an added bonus the indication of which heuristic way/ways are suitable to prepare them that to transform an initial situation (reagents) to a final one (Products). In big molecules (homo and hetero polymers) chemical bonds are not sufficient per se to describe molecules, but the spatial conformations such molecules take are dictated by the bonds structure. The way molecules change answering a change in the environment are appropriate for a universe changing though smooth steps, not via shots.

(*) Emerito presso il Dipartimento di Chimica dell'Università di Pavia, Italy. E-mail: angelo.albini@unipv.it

1. Nomi chimici

Oggi non ha senso parlare di scienza, intesa come un unico corpo organizzato di conoscenze, ma semmai di scienze, discipline diverse che scorrono su strade parallele e ricorrono ciascuna a uno specifico linguaggio. Da sempre gli umani hanno osservato l'ambiente nel quale vivono, notando alcune correlazioni fra un dato evento e un dato effetto, ad esempio il costante avvicendamento delle stagioni e il mutare delle temperature, o le scottature che seguono una lunga esposizione al sole. Erano fenomeni osservati, ma non spiegati. Con lo sviluppo della scienza occidentale, è stato possibile spiegare il perché di tali fenomeni e, quando da un dato evento risulta una malattia, si è cercato di proporre un possibile rimedio.

Purtroppo il linguaggio delle scienze, o meglio, il modo in cui ciascuna di esse racconta il proprio mondo, è complicato dal fatto che ognuna di esse ha un suo codice di riferimento, e non è affatto semplice 'tradurre' le cose dall'una all'altra (nelle numerose occasioni in cui le varie scienze si intrecciano), anche quando le parole sembrano le stesse. Scienziati di diverse discipline declinano segni e fonemi in maniera diversa.

Una particolare possibile descrizione dell'universo è quella chimica, che si basa sulla comunicazione di come siano legati gli atomi nelle molecole che costituiscono il mondo intero. Nel seguito si presentano tre aspetti generali del linguaggio della chimica e della loro relazione con la prima e la più importante (oppure la meno importante in quanto meno razionalizzata) tra le scienze, a seconda del punto di vista scelto), la medicina). È sempre stata un'aspirazione della chimica quella di avere un sistema in grado di assegnare un nome non ambiguo alle sostanze.

Nel clima ottimista della Belle Èpoque, in cui il progresso stava per raggiungere l'acme delle scoperte, si sviluppò l'idea che si potesse costituire una enciclopedia chimica di base con una struttura non più da variare, ma da aggiornare quando fosse il caso con nuovi dati da inserire nella griglia. Questa idea fu realizzata con il Beilstein Handbuch der organische Chemie,[1,2] la cui prima edizione fu pubblicata nel 1881 e universalmente considerata un contributo monumentale alla chimica del popolo tedesco (*Fig. 1*). Purtroppo, la grande guerra distrusse la fiducia nel progresso e bloccò lo sviluppo del piano di aggiornamento dell'opera, ma terminata anche la seconda guerra mondiale, il progetto ripartì e, al giorno d'oggi, è disponibile anche in una raffinata forma informatica.

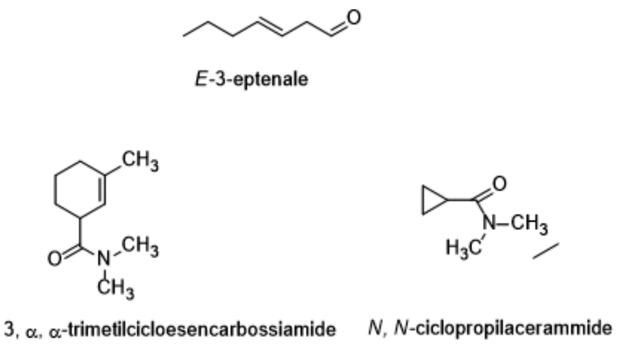


Fig. 1. Esempi di nomi chimici.

e regredendo man mano verso composi più semplici. Lo stesso vale, naturalmente, per la chimica inorganica, in cui la varietà dipende soprattutto dalle differenti possibilità che vi sono per i vari elementi di partecipare alla formazione di molecole o ioni con diversa valenza (Gmelin Handbook of Inorganic and Organometallic Chemistry).[3] Vedi *Fig. 2,3,4.*

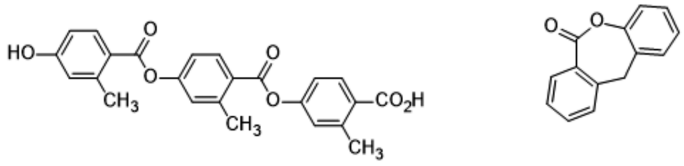


Fig. 2. Esempi di nomenclatura non sistematica, depsidi (a sinistra) e depsidone (a destra).

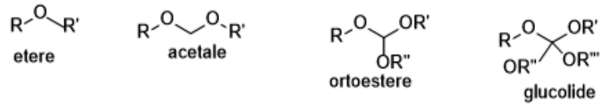


Fig. 3. Nomenclatura sistematica applicata a molecole con numeri crescenti di legami C-O
.

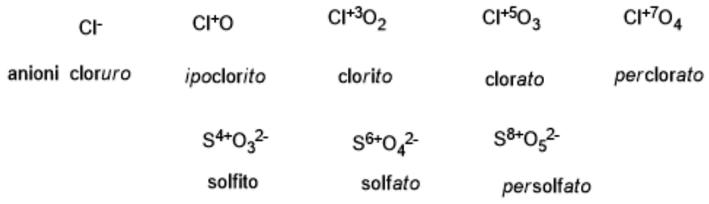


Fig. 4. Esempio di atomi che partecipano alla struttura di anioni con differente valenza.

Naturalmente, queste nozioni di stechiometria, per essere verificate, non hanno alcun bisogno di quantomeccanica, pur essendo perfettamente inquadrabile in essa; per accertare la composizione atomica è sufficiente una bilancia (omnia mensura sunt, sta scritto in lettere grandi sulle pareti dell'aula magna di chimica a Bologna).

Un paio di esempi potranno aiutare a capire quanto una nomenclatura strutturata possa essere di aiuto. Una delle rarissime volte in cui l'autore di questa nota sentì il professor Grünanger, Direttore dell'Istituto di chimica organica a Pavia, alzare la voce con uno studente durante un esame, fu perché l'esaminando non seppe individuare un gruppo funzionale acetalico nella molecola di cui si discuteva, e cercava di uscire dall'impaccio in cui si era cacciato chiamandola doppio etere, mentre il fatto che vi siano due legami C-O invece di uno solo, come avviene con gli eteri, non porta certo a un 'raddoppiamento' della somiglianza nelle proprietà fisiche e chimiche tra i acetali ed eteri. D'altra parte, mi piace ricordare un aneddoto che ho sentito più volte raccontare dal professor Silvio Pietra, che ricordava come durante una sessione di esame cui partecipava accanto al suo capo, questi aveva chiesto a uno studente cosa fossero i depsidi, ottenendo una scena muta. Nel corso dell'esame, era emerso che quello stesso studente nulla sapesse di cosa fossero i depsidoni, e la cosa non avrebbe dovuto meravigliare l'esminatore, perchè i due gruppi di composti presentavano solo una somiglianza nei nomi, mentre solo se si usa la nomenclatura sistematica, viene evidenziata ogni possibile differenza di struttura e di sintesi (*Fig. 2*).

Per avere un'idea più chiara della struttura delle molecole ci si può avvalere della spettroscopia che, grazie all'azione della luce UV e visibile assorbita (o non assorbita) dalle molecole, rivela esattamente come siano fatti i legami. Per questo l'interazione luce-molecole ha un ruolo speciale nella comprensione del perché le molecole, e in particolare le biomolecole, sono così stabili e come reagiscono, aiutando a chiarire meccanismi fondamentali di reazione e anche l'effetto sulle biomolecole e quindi sulla salute.[4][1]

[1] Ad esempio, nel caso della pelle, il maggiore degli organi umani per peso ed estensione, la natura ha predisposto un apposito meccanismo che limita i danni provocati dalla luce sul DNA (che comprendono alcune delle più pericolose forme di tumore, come il melanoma, particolarmente temuta, perché si manifesta a decenni dall'evento che lo ha originato), filtrando la luce a corta lunghezza d'onda. La ricerca scientifica sostituisce la formazione di melanine con filtri artificiali più efficienti. E' ormai noto a

D'altra parte è necessario che la pelle assorba sufficiente luce perché venga formata un'adeguata quantità di previtamina D (vedi *Fig. 5*), la cui conversione in vitamina verrà poi completata con ulteriore idrossilazione nei reni e nel fegato. Della vitamina D si sa molto,[5] ma non tutto, e vengono costantemente scoperti nuovi ruoli, oltre al ben noto governo dell'omeostasi del calcio, di ovvia importanza, dato che l'intero scheletro viene rinnovato ogni pochi giorni.[5]

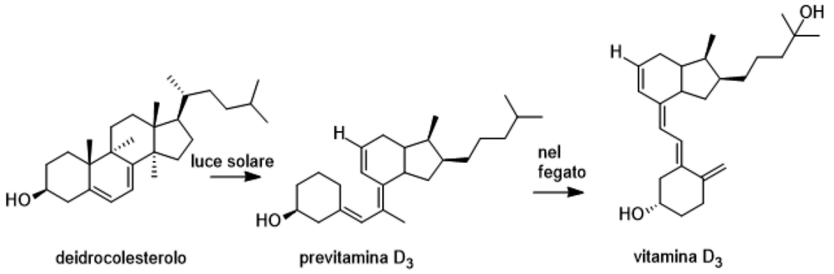


Fig. 5. Formazione della previtamina D_3 attraverso reazione ettrociclica a 6 elettroni.

I gravi effetti dell'insufficiente assorbimento di luce erano già noti a Erodoto, che scrive di essersi recato sui campi di battaglia e di aver notato che i crani dei guerrieri egiziani erano assai più resistenti di quelli dei nemici, un fatto attribuito al costume egiziano di stare a capo rasato a partire dall'infanzia. E' ormai assodato che l'insufficienza di vitamina D produce il rachitismo dei bambini e la osteomalacia, il morbo corrispondente degli adulti, detta anche Morbum Anglorum per la stretta correlazione tra le pessime condizioni igieniche caratteristiche della rivoluzione industriale in Inghilterra e questa carenza.[5,6]

D'altra parte il ricorso eccessivo ai lettini UVA per ottenere un abbronzatura costante, può causare un precoce invecchiamento della pelle o l'alterazione del controllo del livello di altre vitamine, in particolare quelle che intervengono su processi ossido-riduttivi, come le vitamine C, E, K.

---

tutti che, specie per persone di complessione molto chiara, è consigliato utilizzare una sufficiente quantità di crema solare, di cui bisogna controllare, più che il fattore di stabilità (che comunque non ha nessun effetto andare oltre il Sun Protecting Factor (SPF) di 30), mentre è fondamentale usarle in condizioni ottimali e riapplicarla spesso (il SPF si riferisce a una quantità tra 0.5 e 1 mg di principio attivo per cm^2, purché la si riapplichi abbastanza frequentemente e non si lascino parti del corpo scoperte).

Sono stati osservati numerosi altri effetti dell'esposizione alla luce solare, per fortuna meno correlati con la salute, ad esempio lo scoloramento e la perdita di robustezza di tessuti naturali e artificiali, tipicamente legate a processi radicalici innescati dalla luce. Tale meccanismo può manifestarsi anche dopo una fase di iniziazione in cui si formano prodotti che assorbono una parte maggiore di luce e sono buoni iniziatori della degradazione, tipicamente i chetoni).

2. Nomi chimici. Conformazioni e strutture tridimensionali

La considerazione dei legami chimici e dei nomi sistematici dei composti non è tuttavia sufficiente ad esprimere la struttura di molecole più grandi. Consideriamo, ad esempio, le proteine, gli eteropolimeri che si formano dalla policondensazione degli aminoacidi secondo le istruzioni depositate nel DNA. In effetti,

la struttura primaria di una proteina, cioè la sequenza lineare di amminoacidi che la compongono, è sufficiente per giudicarne la composizione, ma non il suo ruolo fisiologico. Questa informazione è tuttavia già presente nella struttura primaria, attraverso la conformazione tridimensionale della proteina *nativa*. Ogni sequenza polipeptidica e la posizione di ogni residuo aminoacidico nella catena *sa* come orientarsi nello spazio e come arrivare spontaneamente alla sistemazione preferita o *nativa* (con notevole crescita di entropia). Il raggiungimento di quella particolare conformazione dipende anche dall'ambiente in cui il processo di piegatura (folding) si completa, ad esempio con la formazione di legami a idrogeno intra e intermoleculari, o in presenza di forze di van der Waals o anche se, ad esempio, il folding inizi o no in maniera co-translationale, cioè mente ancora nuovi legami peptici vengono formati, dalla concentrazione di ioni, dal pH, dalla temperatura, dall'ambiente idrofilo o idrofobo e così via.[7]

Qualcosa di concettualmente simile avviene nella formazione di complessi organometallici a struttura specifica intorno a uno ione metallico. Sia il raggiungimento spontaneo della conformazione nativa che la formazione selettiva di complessi organometallici illustrano come la chimica sia la scienza dell'instabile, del provvisorio, di quanto di più lontano sia pensabile da Goethe e dai tedeschi in genere, nonostante la evidente affinità di quel popolo per tale scienza, e sia dunque l'unica possibile base per la vita, in un certo senso all'opposto della matemati-

ca, o della (quanto)meccanica, il sempre vero, l'assoluto, l'eterno, il das ewig-weibliche zieht uns hinan del Faust.

Le figure qui inserite danno un'idea della formazione di fogli antiparalleli per legami (β) ad idrogeno con ciò formando una struttura secondaria. (vedi *Fig. 6*)

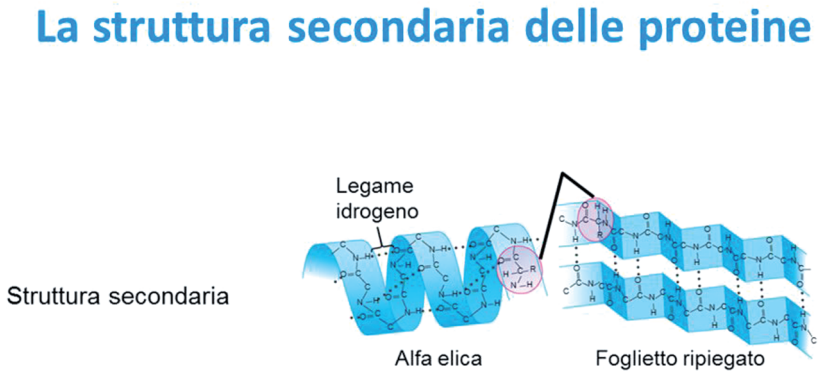


Fig. 6. Self assembing di una proteina, strutture secondarie (da SlidePlayer.it).

Le strutture secondarie offrono il primo passo in questa direzione. Tipiche delle strutture secondarie sono i fogli α e β, che si avvolgono rapidamente in quanto stabilizzati dalla formazione di legami a idrogeno intramolecolari. Per quanto riguarda le strutture terziarie, queste derivano dalla stabilizzazione tra fogli di natura anfipatica, cioè quelli che contengono una parte idrofila e una idrofoba, e si avvolgono in maniera da presentare la regione idrofila verso l'ambiente che circonda la proteina. Un contributo lo possono dare anche legami covalenti relativamente deboli, come i legami S-S della cistina. Infine, strutture quaternarie possono formarsi in certe proteine attraverso la sistemazione o risistemazione di subunità pre-esistenti.

Ad esempio la sequela di eventi che alla fine produce la contrazione di specifiche fibre muscolari comincia con un segnale, un neurotrasmettitore, ACh fornito dal neurone motore che innerva quella fibra attraverso un complesso meccanismo. Fondamentale per la definizione delle conformazioni pensabili è la rigidità del gruppo ammidico che pone limiti agli angoli formati dalle proteine, come si può valutare attraverso l'indice di Ramachandran.[8] Si può controllare sperimentalente che prodotti di denaturazione esposti ad alcuni fattori denaturanti si ripiegano di nuovo e prendono l'esatta conformazione nativa).[7-9]

Una proteina completamente denaturata non ha più struttura né secondaria, né terziaria ed esiste nella cosiddetta catena a caso (random coil). In alcuni casi, le cellule proteggono le loro proteine contro l'effetto denaturante del calore co enzimi detti proteine conto i colpi di calore (proteine 'heat shock', un tipo di molecole accompagnatrici),[9] che assistono altre proteine nel processo di folding, nonché nel rimanere poi nella conformazione preferita. Va tenuto presente che non necessariamente le proteine assumono una struttura denaturano, specie se la trasformazione avviene al di fuori delle condizioni di lavoro tipiche (ad esempio questo è motivo per cui le albumine che costituiscono il bianco d'uovo passino ad opaco). La stabilità termica delle proteine è per altro molto varia e son stati trovati batteri ipertermofili che funzionano perfettamente a 122°C.

Naturalmente, vi sono diverse malattie che dipendono da errori nella sequenza di aminoacidi o da errori nel folding di proteine, in particolare malattie croniche neurodegenerative dovuti al fatto che questi errori rendono poi le molecole incapaci di tornare alla normale struttura nativa. Le proteine indicate come con errori (misfolded), tipicamente contenenti fogli β organizzati in maniera supramoleculare, note come fogli-β reticolati, assai stabili, del tutto insolubili e molto resistenti all'idrolisi in condizioni acide (strutture amiloidi) ne sono un tipico esempio. La stabilità strutturale di queste componenti fibrillari ne favorisce l'accumulo, eventualmente con altre proteine in aggregati non più in grado di svolgere le previste funzioni fisiologiche. Queste proteine sono collegabili con i morbi di Huntington, di Parkinson e di Alzheimer che tutti esibiscono un accumulo di proteine con errori in aggregati insolubili extra o intra-cellulari (amiloidi); che poi sia questa la causa di tali malattia o invece una conseguenza della perdita di omeostasi di alcune proteine non sembra per ora definitivamente provato.

Gli agenti denaturanti comprendono sia alcune molecole o ioni (come l'urea, il cloridrato di guanidinio, il pH), come anche fenomeni fisici, come temperatura, pressione, e così via. L'equilibrio tra conformazione differenti e diverse d una proteina dipendono dallo stato nativo, gli stati intermedi non piegati a loro volta dipendono dal denaturante usato, e, con valutazioni attente, si può comprendere quale sia il profilo energetico del processo di folding e unfoldinge e identificare gli intermedi che si formano sul percorso seguito, nonchè ricavare i parametri termodinamici relativi. Va ricordato il numero elevatissimo di conformazioni da confrontare (dell'ordine di 3^{300}, cioè 10^{143}) che sem-

brerebbe difficile da raggiugere concretamente se si dovesse valutare singolarmente ciascuna conformazione.[10]

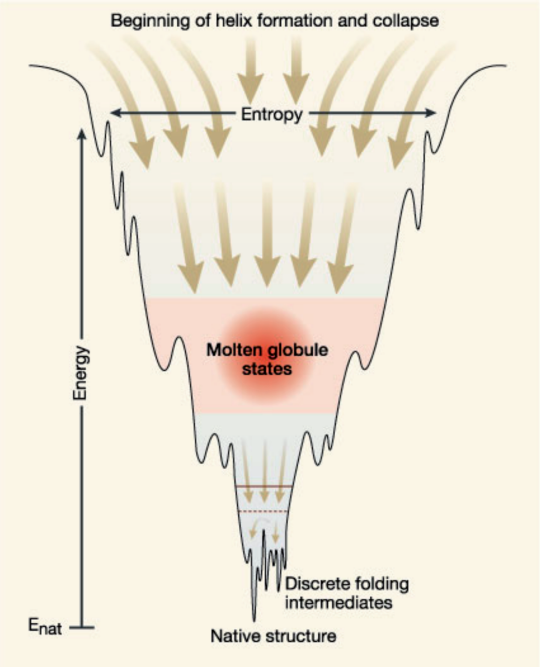


Fig. 7. Panorama degli spostameti nello spazio in funzione del livello energia quando il folding di una proteina e imbuto (funnel) energetico attraverso il quale un polipeptide arriva alla struttura nativa. (da Summary of the Horizon Symposium on 'Protein Misfolding and Disease'Christopher M. Dobson).

Secondo J Bryngelson and P. Wolynes,[11] le proteine seguono il principio di minima frustrazione, nel senso che le proteine come si sono evolute in natura devono aver minimizzato la loro energia, come indicato nella figura. Peraltro qualche grado di frustrazione rimane nonostante l'evoluzione, come indicato dall'esistenza minimi locali nel panorama energetico (vedi *Fig. 7*).[12]

Di fatto, il raggiungimento della conformazione preferita nelle proteine è stato stimato avvenga attraverso una molteplicità di imbuti, la maggioranza dei quali sono diretti verso lo stato nativo (folding funnel), e non attaverso un solo meccanismo. in tempi che vanno da100 ns a qualche ms, come tipico per i β turns.

Lo stato di transizione, per il quale ogni proteina deve passare per raggiugere la struttura nativa è sostanzialmente una forma non ancora completata di quella finale, non un altro intermedio, e determina la velocità, quandanche vi possano essere altri stati a maggior energia.[13,14]

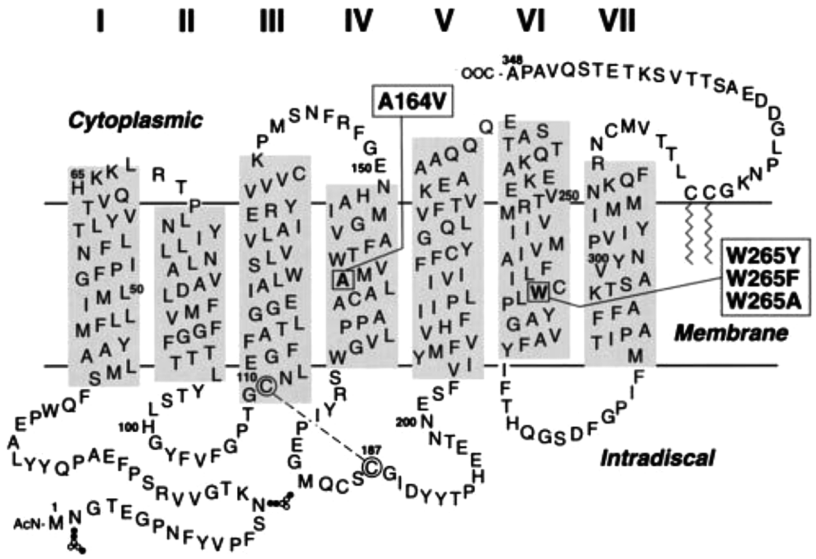


Fig. 8. Struttura bidimensionale della rodopsina. Come d'uso, i singoli amminoacidi sono indicati da una sola lettera.[15]

Nel caso della visione, basata sul ruolo della rodopsina (vedi *Fig. 8*), il pigmento porpora fotosensibile, che è uno solo anche se la struttura tridimensionale ne consente tre conformazioni differenti, che comportano differenti assorbimenti nello spetto visibile. Il processo fotochimico si deve alla transizione trans cis di un doppio legame C=C, da 11-trans a 11-cis nella molecola del retinale che è il gruppo prostetico, seguito da una serie di processi termici e infine del distacco del retinale, che viene poi rigenerato attraverso il corrispondente alcool. Le molecole di opsina, una proteina relativamente semplice (poche centinaia di residui amminoacidici) si sistemano in modo da formare sette canali attraverso la membrana che separa il citoplasma dalla soluzione esterna. La drastica variazione nella geometria della molecola di retinale (che è legata con un legame covalente come gruppo immonio, >C=N^+R sposta

notevolmente la sesta catena, e come risultato cambia il passaggio attraverso la barriera, e quindi la concentrazione finale degli ioni Na^+ e Ca^{2+} (si vedano le figure 8, 9). Non è questa la sede adatta per discutere la maniera assai complessa che porta alla visione tricromatica (rosso-blu-verde come spetro di sensibilità, tonalità, saturazione e luminosità, dal punto di vista quantitativo) degli umani. Per altro non è facile nemmeno pensare a come vedano animali che lo fanno un minor o maggior numero di dimensioni (ad esempio cani, cavalli e squali hanno una dimensione in meno, ma uccelli e pesci una in più, e alcune farfalle molte di più). È da notare però che tutti i vertebrati condividono lo stesso sistema di base con i sette canali della trans-membrana, che è perfezionata ed è rimasta sostanzialmente la stessa da gran tempo in tutte le specie dei vertebrati, molto di più di altre caratteristiche molto importanti, quali la mascella mobile, a indicazione che l'evoluzione si ferma quando una soluzione sodisfacente è stata trovata.

3. Ruolo del rasoio di Occam in chimica

Il concetto stesso di ipotesi scientifiche verificabili va considerato tuttavia con cautela. Ad esempio, il principio della minima azione e minimo spostamento domina la chimica e la scienza in genere, anche se non sempre in maniera consapevole. Si accetta semplicemente il valore euristico del rasoio di Occam, il cui fondamento comporta l'assunzione che il mondo sia semplice, cosa che non è però scontata. Alcuni aspetti di questo problema vengono accennati nel seguito, sulla base di una pubblicazione di R. Hofmann *et al.*[16]

Il rasoio di Occam può essere espresso con le seguenti affermazioni.

A Frustra fit per plura quod potest fieri per pauciora.
B Quando propositio verificatur pro rebus, si duae res sufficiunt ad eius veritatem, superfluum est ponere tertiam
C Pluralitas non est ponenda sine necessitate.
D Nulla pluralitas est ponenda nisi per rationem vel experientiam vel auctoritatem illius, qui non potest falli nec errare, potest convinci.

Si tratta, dunque, di stabilire la validità di singole ipotesi scientifiche, per quanto l'affermazione D possa servire a giustificare la reintroduzione della pluralità, pur entro limitazioni precise (Nec intra ulla pluralitas est ponenda nisi per rationem vel experientiam vel auctoritatem illius, qui non potest falli nec errare, potest convinci).

Se s'intende il termine 'meccanismo di reazione' come la descrizione di tutti i passaggi 'elementari' che portano dai reagenti ai prodotti al livello molecolare che quindi include principio almeno, la conoscenza della geometria e dell'energia relativa di tutte le strutture che hanno un ruolo, comprese quelle isolabili, almeno potenzialmente, con intermedi e stati di transizione, e i secondi rappresentanti i punti in cui l'evoluzione di un avanzamento della traiettoria della energia minima che congiungono le specie che si interconvertono. In una definizione alternativa, il meccanismo stabilisce l'evoluzione di un sistema chimico lungo la traiettoria di reazione, attraverso un cammino che congiunge i reagenti ai prodotti nello spazio di tutte le coordinate nucleari richiede l'introduzione del concetto di superfici dell'energia potenziale (PES). Di fatto, ben prima che l'idea di molecola prendesse piede, il principio di minima azione era già stato sviluppato dopo esser stato introdotto da Pierre Louis Morveau de Maupertuis e applicato universalmente da Leonardo Eulero in balistica, nel movimento di una forza centrale. Secondo questo principio, i moti spontanei avvengono sempre con la minima quantità di 'azione', e questa è una ben definita variabile fisica. In una presentazione svolta a Parigi nel 1744 su questo principio, de Maupertuis sottolineò che ad esempio la luce non segue linea più breve, né la più veloce, ma quella che è più economica (si confronti la legge della parsimonia), cioè quella in cui la quantità di 'azione' è minima.

Non sorprende quindi che, quando negli anni Trenta il meccanismo delle reazioni chimiche diventò più importante, anzi divenne un punto focale della chimica fisica organica, gli studi avvennero con lo spirito ed il vocabolario della meccanica. Peraltro, il primo passo in questa direzione era stato compiuto anche prima, in effetti da Muller nel 1886,[17] cioè in un tempo in cui la teoria molecolare era ancora recente e questo scienziato introdusse la regola della minor variazione possibile nel corso delle reazioni chimiche.[18]

L'idea era attraente e trovò posto in vari libri di testo,[19] ove venne indicato come il principio del minor variazione possibile e fu formulato nei termini più generali da F. Rice e E. Teller, che nel 1938 proposero il principio del minimo movimento (PLM),[20] secondo il quale verranno favoriti i passaggi 'elementari' che riguardano le variazioni minori nella configurazione degli atomi e nella configurazione elettronica'. In questo contesto di grande importanza furono le regole di simmetria degli orbitali molecolari (MOs) che furono introdotte nella chimica organica 27 anni dopo.[21]

Per applicare il PLM a una certa reazione, gli atomi che costituiscono reagenti e prodotti devono esser spostati l'uno rispetto all'altro in modo che i movimenti dei nuclei (di solito misurati dai loro quadrati) vengano minimizzati. Di fatto, un buon numero di riarrangiamenti, decomposizioni ed eliminazioni sono state dimostrate avvenire attraverso cammini che ubbidiscono alle richieste del PML. La grande semplicità dei calcoli richiesti e, cosa più importante, la chiarezza dell'idea provocarono l'ampia applicazione del PML, in particolare quando bisognava scegliere tra varie alternative. Tuttavia era sempre chiaro a chi usava questo metodo che il PLM rappresentava un modello teorico estremamente semplificato del movimento reale dei nuclei e degli elettroni nel corso di una reazione, un movimento che per esser descritto in maniera adeguata abbisogna della soluzione delle equazioni quantomeccaniche e questo lascia sempre il dubbio che la quantizzazione degli stati elettronici, vibrazionali e rotazionali cambi il risultato. Inoltre, bisogna tener in conto il fatto che è necessario un approccio dinamico per descrivere la reazione come avviene in realtà, cioè attraverso una miriade di percorsi accessibili. Tuttavia il PML va incontro al bisogno generale di semplicità, e quindi il fatto che l'allontanamento dal modello o anche l'incompatibilità con le previsioni del PML non resero mai questo principio un metodo per accettare definitivamente un meccanismo.

Consideriamo per primo una classica reazione pericliclica,[22] il riarrangiamento di Cope (spostamento sigmtropico 3,3, *Fig. 9*). In questo caso, anche una minima variazione nella struttura nell'idrocarburo di partenza, l'1,5-esadiene, sembra causare una variazione dal cammino più tipico (a) con il suo intermedio a struttura 'aromatica' (in due forme isomere), oppure attraverso i cammini (b) o (c), che son caratterizzati, rispettivamente, da uno stato di transizione biradicalico come stato di transizione o come vero e proprio stato intermedio.[23]

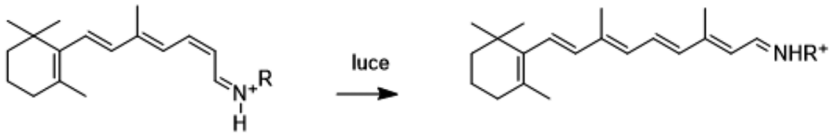


Fig. 9. Tre possibili meccanismi per il riarrangiamento di Cope.

Come esempio, consideriamo quello di fullereni, i prodotti poliedrici ottenuti dalla grafite per vaporizzazione a temperature del plasma (>3000 °C). Contrariamente a una previsione 'entropica' di forme poco

organizzate strutturalmente, forme della materia altamente organizzate sono la maggioranza, nello specifico molecole altamente simmetriche con poliedri C_{2n} con una struttura che ricorda le cupole geodesiche usate in architettura da R. Buckminster Fuller.[24] La configurazine C_{60}, che ha la forma di un icosaedro troncato, come un pallone da calcio ha attirato molta attenzione per la perfezione della sua struttura poliedrica, la sua relativa stabilità e gli orizzonti che apre essendo una nuova forma allotropica del carbonio (*Fig. 10*).

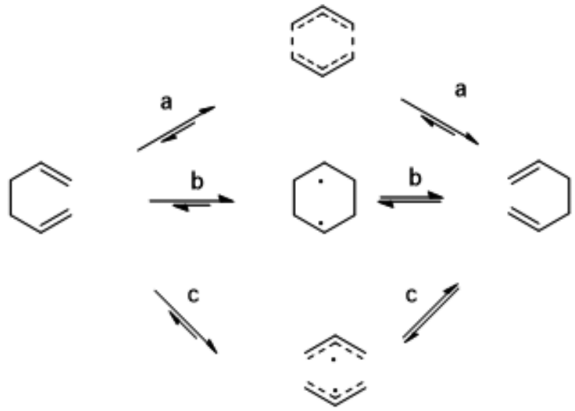


Fig. 10.

Come disse uno degli scopritori del C_{60}, R. Smalley, "Naturalmente, ci devono essere centinaia di meccanismi per i quali si possa andare al C_{60} che coinvolgono forze che portano alla spontanea auto sistemazione (self-assembly) degli atomi di carbonio. Questa affermazione evidenzia la grande complessità del problema e la terribile incompletezza del nostro sapere. Più capiamo delle origini di una reazione, più capiamo che i meccanismi sono intrinsecamente molteplici, e comportano un riarrangiamento geometrico delle posizioni relative dei nuclei che avvengono nelle molecole che interagiscono e si muovono lungo un percorso sulla superficie dell'energia potenziale (PES), traversando creste che costituiscono le barriere nella reazione. Vedere una PES ricorda un paesaggio collinare, e si può continuare la metafora paragonando la riuscita trasformazione della molecola con il successo di un montanaro che si muove dalla valle dei reagenti a quella dei prodotti attraverso uno dei passi più facilmente accessibili.

Ma i paesaggi collinari del mondo reale, come quelli calcolati, non sono così monotoni da presentare un unico passo tra le valli, cosicché traiettorie che portano a reazione non sono rare e il loro numero cresce rapidamente se si fa crescere di un po' l'energia cinetica, condizione nella quale la necessità di passare tutte per lo stesso punto di sella diviene una caratteristica meno stringente. Inoltre, se gli spostamenti dei nuclei nel corso di un riarrangiamento sono abbastanza piccoli, la reazione può avvenire anche attraverso (sotto) la barriera d'energia, per via di un tunnel quanto meccanico.

Consideriamo ora un fenomeno che si possa spiegare in due modi, l'uno per la via A, che ne è il fattore determinante

(1) $\Pi = A$

mentre l'altro si può spiegare col contributo di due modi, A e B

(2) $\Pi = ca\ A + cb\ B$

Può allora ben essere che nel caso di una osservabile sia valida una spiegazione *semplice* che spieghi bene i fatti, e con il rasoio di Occam sarà questa la via preferita. Ma c'è d'aspettarsi che nell'universo che si osserva, ci siano diverse osservabili, $\Pi 1$, $\Pi 2$, $\Pi 3$, che presentano una spiegazione più complessa, come nell'eq 2. Anche quando uno solo di tali modi è presente, questo fatto può portare ad individuare ulteriori esperimenti, e quindi può avere un ruolo propositivo nella ricerca.

Questo approccio potrebbe esser formalizzato con il metodo epistemologico proposto all'inizio del '900 dal geologo di Chicago, TC Chamberlain,[25] e poi usato da J. Platt come la base di un metodo per avere conclusioni rigorose, o che entrambe siano pensabili come gemmazioni.[26] Dal metodo induttivo di F. Bacon, sottolineando che il fatto di ottenere più test simultaneamente sperimentabili, ciascuno con le proprietà di arrivare alle conclusioni deve arrivare a conclusioni più ricche e più complete che l'insistenza di un solo modo. La storia della chimica offre moltissimi esempi di errori provocati dall'uso troppo stretto di ipotesi.

Chi si presenta nelle vesti di un critico del rasoio di Occam può forse esser accusato di assumersi capacità filosofiche non garantite, in quanto si implicherebbe senza dichiaralo la preferenza per un modello semplice e la credenza che viviamo in un universo semplice. Chi però

continua e dice che il mondo non può esser così semplice e su questa base invalida l'applicazione del rasoio di Occam negli studi scientifici ha ragione?. Ma è poi vero che il credere in un universo semplice è necessariamente implicato dall'essere filosoficamente onesti quado si invoca il rasoio d Occam? Dal punto di vista epistemologico, chi cerca dei mattoni da cui si possa ricostituire l'architettura miracolosamente complessa della natura? è questo l'uso corretto del rasoio di Occam? Non è meglio usare questo strumento per identificare questi moduli individuali, piuttosto che usarlo per l'intera Weltanschaung che si può costruire da essi? Il principio della parsimonia non è un'affermazione metafisica per cui l'universo è come è. Ciascuno sa bene che è estremamente complesso e il rasoio di Occam, che funziona, è una ricetta per capire come il sistema si organizzi. Secondo questo punto di vista prammatico, il rasoio di Occam serve come un principio operativo, non come una regola o legge della natura.

Nel cosiddetto metodo scientifico, si cercano esperimenti che possano falsificare delle affermazioni. La rinuncia a idee che avviene quando un modello non funziona si assume sia completo e irreversibile, purché questo esperimento sia accessibile a criteri di validità intellettuale e l' esecuzione sia competente e pertanto riproducibile

Nell'interpretazione prammatica del rasoio di Occam, non si applica un linguaggio tanto irrevocabile, ma si possono usare due modelli per ogni altro aspetto egualmente validi, e la scelta tra i due ipotesi deve bensì esser presa in favore della più semplice, ma il fatto di scartare quella più complessa deve esser inteso come condizionale. L'idea che è stata lasciata può essere riconsiderata in favore di un'altra ritenuta migliore al momento, e non scartata definitivamente perché la relativa ipotesi potrebbe non superare ulteriori test futuri. Non deve quindi essere indicata come scartata, per aver violato il rasoio di Occam, un'espressione che pertiene a un mondo in cui i test che li escludono sono più rigorosi.

Ma, a parte il senso intuitivo di quello che è giusto, perché si dovrebbe preferire questo modello a quello della retta? Una risposta può venire dai gradi di libertà di ogni modello che funziona. In statistica,[27] il numero di gradi di libertà è la differenza tra il numero di osservazioni sperimentali indipendenti e il numero di parametri aggiustabili nelle funzioni che cercano di descrivere la relazione tra y e x. è assiomatico che qualunque funzione con un numero di parametri aggiustabili eguale o maggiore del modello può esser fatto passare esat-

tamente per i punti (x, y) nel grafico, ma non è necessariamente vero che una funzione con un minor numero di aggiustabili del numero di osservazioni passi per i vari punti. Se capita che lo faccia, allora la funzione matematica che si è usata come modello ha avuto qualche successo nel descrivere gli avvenimenti che abbiamo studiato sperimentalmente.

I punti in cui passa la funzione, senza alcun motivo algebrico che così esiga, costituisce una misura, che sarà seguita da altre e le previsioni di ciascuna funzione saranno valutate. Ad esempio, si presuma che si cerchi una funzione matematica tra x e y che possa predire altri punti futuri sul grafico; si avrà più tendenza a scegliere quei modelli che abbiano avuto maggior successo nel predire le misure che abbiam fatto fino a quel momento. Questo sarà il modello con il maggior numero di gradi di libertà, o quello col numero minore di parametri aggiustabili – cioè il più semplice.[16]

La rappresentazione migliore di y in funzione di x serve anche a illustrare una qualche finale rappresentazione per applicare il raso di Occam come un metodo operazionale. Il numero di modelli egualmente ben funzionanti in una classe è in genere legato alla complessità di tale classe, ad esempio c'è una sola retta che passi attraverso tutti punti di (x, y), per la quale c'è solo una di quelle relazioni, ma per rappresentare la relazione x, y in generale con linee non rette che passino per tutti i punti c'è un'infinità di soluzioni.[28,29] D'altra parte, perchè il numero di parametri richiesto per descrivere line frastagliate necessarie per modelli complessi, occorre un'infinità di linee frastagliate che passino esattamente per tutti i punti. Con le osservazioni fatte fin qui, non c'è nessun metodo difendibile logicamente per scegliere tra uno e infinito.[16]

Il rasoio di Occam, costringe a limitare il numero di ipotesi da considerare e, di conseguenza, riduce il ruolo dell'immaginazione, ma è lo scotto da pagare per evitare di trovarsi di fronte a un numero illimitato di soluzioni. Qualcuno potrebbe chiedersi se gli scienziati pensino di capire tutto su come funziona l'universo. Naturalmente la risposta è assolutamente no. Ma, se vogliamo trarre un qualunque senso dal mondo, dobbiamo semplificare. Se dovessimo operare sotto le condizioni di eguale opportunità, avremo un mondo troppo complesso per essere descritto e una descrizione auto contradittoria della natura che ci impedirebbe di sentir mai che stavamo facendo qualche progresso nel capire o utilizzare l'ambiente.[30,31]

Abbiamo mai progredito? Con dolore ci rendiamo conto di tutte le ambiguità dell'idea di progresso nel ventesimo secolo, durante il quale la scienza è fiorita, ma altrettanto ha fatto la profonda sfiducia in tale idea da parte di persone ragionevoli.

Se ci distacchiamo dalle implicazioni filosofiche derivate dal trattare il rasoio di Occam come un semplice principio operativo, rischiamo di macchiarci di codardia intellettuale? Crediamo davvero che una correlazione lineare sia la più adatta? E in verità, se non otteniamo tali valori, non penseremo a un errore nell' esperimento? E questa aspettativa non è una fede in un universo semplice?

5. Conclusioni

A modesto parere dell'estensore di questa nota, il divorzio tra le varie discipline è ormai consumato, e, visto che sia le parole sia i simboli vengono usarti in maniera differente non c'è probabilità di tornare a una idilliaca unità, comunque impensabile per la veloce evoluzione delle discipline. Si è cercato sopra di indicare le caratteristiche del linguaggio della chimica, basato sulle formule che molto bene descrivono come sia fatto l'ambiente che ci circonda e come possiamo trarne cose utili. Per altro, già il passare dalle formule dei reagenti a quelle dei prodotti, attraverso le equazioni chimiche causa problemi, in quanto questa relazione è stata a lungo indicata col segno di eguaglianza (=), con un significato ben diverso, tuttavia, da quello che si addotta in matematica (Reagenti=Prodotti) . In chimica si poneva l'accento sul significato quantitativo niente si crea, niente si distrugge, ma tutto si trasforma. Ora del tutto abbandonato per i troppi carichi che si erano accumulati su di lui, il segno = è stato sostituito dalla freccia (→) che distingue un prius e un post, senza ulteriori complicazioni.

BIBLIOGRAFIA

1. Luckenbach, R, The Beilstein Handbook of Organic Chemistry, J. Chem. Inf. Comput. Sci. 1981, 21, 82-83.
2. Beilstein Handbuch der organische Chemie, Beilstein Institute for the develpment of chemistry, Frankfurt.
3. Gmelin Handbook of Inorganic and Organometallic Chemistry, Beilstein

Institute for the develpment of chemistry, Frankfurt.
4. Albini, A, Light, molecules, reactions and health. Elsevier, London, 2020.
5. Holick, MF, Vitamin D, a 3 steps strategy to cure our most common health problems, Hudson Free Press, New York, 2010.
6. Holick, MF, Am J Clin Natur, 2004, 80, 16885-1689.
7. Anfinsen CB ,1973, , Science, 1973, 181, 223–230.
8. Ramachandran GN, Ramakrishnan, C, Sasisekharan, V, Stereochemistry of polypeptide chain configurations". J. Mol. Biol. 7, 95–9.
9. R Krämer, JM Lehn, AM Marguis/Rigault, , Proc Natl Acad Sci USA, 1993, 90, 5394-5398.
10. Kim YE, Hipp MS, Bracher A, Hayer-Hartl M, Hartl, FU, Annu Rev Biochem, 2013, 82, 323–355.
11. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG, Proteins,1995, 21, 167–95.
12. Compiani M, Capriotti, E (2013). "Computational and theoretical methods for protein folding". Biochemistry. 52 (48): 8601–24.
13. Fersht, AR, Proceedings of the National Academy of Sciences of the United States of America, 2000, 97, 1525–1529.
14. Bryngelson, JD, Wlynes,P, J Phys Chem, 1989, 93, 6902-6915.
15. Hargrave, PA, Dowell, JH, Rhodopsin and phototransuction, Int. rev. Cytol, 37b, 1992, 9-97.
16. Hoffmann, R, Minkin, VI, Carpenter, BK, 1997, HYLE—International Journal for Philosophy of Chemistry, 3 (1997), 1-28.
17. Muller, A, Bulletin de Société Chimique, Paris, 1886, 45, 438.
18. W. Hückel, W, Theoretische Grundlagen der Organischen Chemie, Akademische Verlagsgesellschaft, Leipzig, 1934.
19. Wheland, GW, Advanced Organic Chemistry, Wiley, New York, 1960.
20. Rice, F,. Teller, Journal of Chemical Physics, 1936, 489.
21. R.B. Woodward and R. Hoffmann, "The Conservation of Orbital Symmetry, Verlag Chemie, Weinheim, 1970.
22. Bhunia, A, Rot, t, Pachfule, P, Rajamohanan, PR, Biju, AT, Angew. Chem, IEE, 2013, 52, 10224-10227.
23. Hammond, G, DeBoer, C, J. Am. Chem. Soc, 1963, 86, 899-902.
24. Kroto, HW, Heath, JR, O'Brien, SC, Curl, RF, Smalley, RE, Nature, 1985, 6042. 162-163.
25. Chaberlain, TC, The method of multiple working hypotheses, Studies for students, 1897, 837-848.
26. Platt, JJ, Science, 1964,146, 4.
27. Schuster, HG, Deterministic Chaos, Verlag Chemie, Weinheim, 1984.
28. Kalinowski, K, Tadeusz, Krygowski, TM, Croatica Chim Acta, 1988, 58, 107-126.
29. Rychlewski, J, Journal of Applied Mathematics and Mechanics, 1984. 48, 303-314.
30. Jeffreys, WH, Theory of Probability, Oxford University, Oxford, 1939.
31. Jefferys, WH, Berger, American Scientist, 1992, 80 64-72.

Istituto Lombardo (Rend. Scienze) 153, 41-56 (2019)

MATEMATICA E CALCOLO SCIENTIFICO

L. DONATELLA MARINI (*)

Nota presentata dal m.e. Piero Colli Franzone
(Adunanza del 28 marzo 2019)

SUNTO. – In questa breve comunicazione cercherò di dare un'idea di cosa si intenda per Calcolo Scientifico e di quale sia il ruolo della Matematica all'interno di esso. Illustrerò su un esempio estremamente semplice le fasi che portano dalla modellizzazione di un fenomeno fisico alla sua risoluzione. Elencherò vari problemi di interesse per le applicazioni, limitandomi a quelli più vicini alle mie competenze, e mostrerò le difficoltà che possono insorgere presentando casi famosi di "fallimenti" dovuti a piccoli, apparentemente trascurabili, errori in una qualunque delle fasi.

***

ABSTRACT. – This short note is meant to give an idea of what Scientific Computing is and what its goal is. In particular, I will focus on the role of Mathematics inside Scientific Computing. On an extremely simple example I will indicate the various steps that lead from the model of the physical phenomenon to the practical solution. I will provide a list of few problems of interest for the applications, among those closer to my field of expertise, and finally I will show some famous cases of failure due to small, almost negligible, errors in one of the steps.

1. INTRODUZIONE

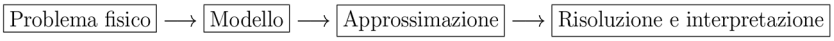
In questa breve nota cercherò di spiegare, seppure in modo inevitabilmente approssimativo, cosa si intende per Calcolo Scientifico e soprat-

(*) Professore Emerito Università degli Studi di Pavia, Dipartimento di Matematica, Pavia, Italy. E-mail: donatella.marini@unipv.it - marini@imati.cnr.it

tutto quale ruolo svolge all'interno di esso la Matematica cosiddetta Applicata. Sottolineo "cosiddetta", in quanto il termine Matematica Applicata, in opposizione a Matematica Pura, è discusso e criticato da molti, non senza valide ragioni: la Matematica è una sola, e tutta è applicabile. Sarebbe più appropriata la dicitura Matematica per le applicazioni.

Sotto il nome "Calcolo Scientifico" si colloca tutta una serie di attività di ricerca che vede coinvolti ricercatori provenienti da ambiti diversi, con competenze in diversi settori scientifici, allo scopo di predire l'andamento di fenomeni fisici di interesse applicativo tramite la simulazione numerica, come supporto alle scelte industriali e, più in generale, alle decisioni da prendere.

Molto grossolanamente, il Calcolo Scientifico si pùo schematizzare in quattro fasi di intervento. Nella prima fase viene individuato il problema fisico di interesse. Gli esempi di problemi da trattare sono numerosissimi, e vanno dalla Fisica alla Chimica, ai più svariati aspetti dell'Ingegneria, e più recentemente alla Biologia, all'Economia, alla Medicina, eccetera. Nella seconda fase si deve passare dal problema fisico alla sua modellizzazione: il problema va tradotto in opportune equazioni che descrivano in maniera esaustiva il fenomeno che si deve analizzare. Si passa poi alla fase di risoluzione delle equazioni descrittive del fenomeno e infine alla rappresentazione e interpretazione dei risultati.



2. Dal problema fisico alla sua risoluzione

Come dicevo, i problemi di interesse nelle applicazioni sono numerosissimi. Ne citerò solo alcuni, più legati alla tipologia delle mie ricerche: uno dei primi settori nei quali si è sviluppata la simulazione numerica (tralasciando le applicazioni belliche, alle quali non ho mai partecipato) è sicuramente quello della progettazione degli aerei da trasporto civile, per combinare costi (di costruzione e di consumo) e sicurezza. In particolare ci si deve muovere in ambiti ingegneristici molto diversi tra loro, come la dinamica dei fluidi, la dinamica delle strutture, l'elettronica, la combustione, l'acustica, eccetera. Questo, in generale, viene trattato da gruppi diversi (ciascuno con le proprie competenze specifiche), che vanno periodicamente messi in cantatto. Un altro settore (per alcuni versi simile) che

ha visto un enorme sviluppo è quello automobilistico, per combinare prestazioni, comodità, sicurezza ed estetica. Le applicazioni legate in qualche modo alla geologia hanno avuto anch'esse sviluppi enormi: dalla ottimizzazione dello sfruttamento dei pozzi petroliferi allo studio delle catastrofi naturali (terremoti, inondazioni, etc.) senza tralasciare la gestione delle scorie (da quelle radioattive alla più comune plastica).

In tempi più recenti si è avuto un grandissimo sviluppo delle applicazioni alla biologia e alla medicina, dove il problema di "descrivere matematicamente cosa succede nei vari casi" risulta *molto* più difficile che nei casi precedenti (che pure non sono facili).

Un altro settore che ha visto sviluppi enormi in tempi recentissimi è poi quello della intelligenza artificiale, che però, al momento, è molto più vicino all'informatica che al calcolo scientifico.

Per evitare di entrare in dettagli troppo tecnici, che tra l'altro richiederebbero competenze che, come praticamente avviene in tutti i settori, si avviano verso forme di specializzazione sempre più spinte, mi limiterò a pochi accenni superficiali a qualche caso molto particolare.

2.1. Alcuni esempi di modelli matematici

Dato il problema di interesse, per poterlo studiare lo si deve descrivere in maniera corretta tramite un modello opportuno. Questo avviene nella seconda fase: il problema va tradotto in opportune equazioni che descrivano in maniera sufficientemente accurata il fenomeno che si deve analizzare. Nella grande maggioranza dei casi si tratta di equazioni differenziali, ossia equazioni in cui compaiono delle derivate. Un esempio è rappresentato dalle equazioni di Navier-Stokes per fluidi incomprimibili (come l'acqua e l'aria), che sono alla base di tutti i problemi di Fluidodinamica

$$\begin{cases} \dfrac{\partial u}{\partial t} - \nu(\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2}) + u\dfrac{\partial u}{\partial x} + v\dfrac{\partial u}{\partial y} + w\dfrac{\partial u}{\partial z} + \dfrac{1}{\rho}\dfrac{\partial p}{\partial x} = F_x \\ \dfrac{\partial v}{\partial t} - \nu(\dfrac{\partial^2 v}{\partial x^2} + \dfrac{\partial^2 v}{\partial y^2} + \dfrac{\partial^2 v}{\partial z^2}) + u\dfrac{\partial v}{\partial x} + v\dfrac{\partial v}{\partial y} + w\dfrac{\partial v}{\partial z} + \dfrac{1}{\rho}\dfrac{\partial p}{\partial y} = F_y \\ \dfrac{\partial w}{\partial t} - \nu(\dfrac{\partial^2 w}{\partial x^2} + \dfrac{\partial^2 w}{\partial y^2} + \dfrac{\partial^2 w}{\partial z^2}) + u\dfrac{\partial w}{\partial x} + v\dfrac{\partial w}{\partial y} + w\dfrac{\partial w}{\partial z} + \dfrac{1}{\rho}\dfrac{\partial p}{\partial z} = F_z \\ \dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} + \dfrac{\partial w}{\partial z} = 0 \end{cases}$$

Si tratta di un sistema di 4 equazioni differenziali alle derivate par-

ziali nello spazio e nel tempo per le 4 incognite: u, v, w (le tre componenti della velocità) e p (la pressione). I dati del problema sono le componenti della forza applicata F_x, F_y, F_z, e i coefficienti ρ e ν, rispettivamente densità e viscosità del fluido. Sono equazioni molto complicate da trattare e, anche se a noi Matematici piace scriverle nella forma più compatta

$$\begin{cases} \dfrac{\partial \mathbf{U}}{\partial t} - \nu \Delta \mathbf{U} + \mathbf{U} \cdot \nabla \mathbf{U} + \dfrac{1}{\rho} \nabla p = \mathbf{F} \\ \nabla \cdot \mathbf{U} = 0 \end{cases}$$

non cambia la sostanza: si guadagna in eleganza ma le difficoltà restano ovviamente le stesse.

Un altro esempio è costitutito dalle equazioni di Maxwell, che modellizzano i fenomeni di elettro-magnetismo. Nel vuoto si scrivono così

$$\begin{cases} \varepsilon_0 \nabla \cdot \boldsymbol{E} = \rho \\ \dfrac{\partial \boldsymbol{B}}{\partial t} + \nabla \times \boldsymbol{E} = 0 \\ \nabla \cdot \boldsymbol{B} = 0 \\ \varepsilon_0 \mu_0 \dfrac{\partial \boldsymbol{E}}{\partial t} - \nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J} \end{cases}$$

Le quantità incognite sono il campo elettrico $\boldsymbol{E}$ e il campo di induzione magnetica $\boldsymbol{B}$. Gli altri parametri che compaiono nelle equazioni sono dati, ossia sono quantità note: ε_0 e μ_0 sono rispettivamente la permittività elettrica e magnetica, ρ e $\boldsymbol{J}$ sono la densità di carica e di corrente rispettivamente.

Un altro esempio è l'equazione di Schr¨odinger, fondamentale in meccanica quantistica, che determina l'evoluzione nel tempo dello stato di un sistema, ad esempio di un elettrone o di una particella subatomica nello spazio e nel tempo sotto l'azione di varie forze:

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = \widehat{H} \Psi(\mathbf{r}, t)$$

Una descrizione delle quantità in gioco che possa essere chiara per non addetti ai lavori è praticamente impossibile: l'incognita è la funzione d'onda Ψ, i è l'unità immaginaria, $\hbar$ la costante di Planck e $\hat{H}$ l'operatore Hamiltoniano.

È evidente che si tratta di equazioni molto complesse, impossibili da risolvere con carta e penna. E come già osservato da Fourier, un'equazione, per quanto bella, è di scarsa utilità se non la si sa risolvere, almeno in modo approssimato. Bisogna allora cercare di trovare soluzioni approssimate di queste equazioni, individuando dei metodi opportuni che possano essere implementati su computers e diano informazioni il più possibile accurate del fenomeno che si sta studiando. Naturalmente l'attendibilità dei metodi utilizzati viene filtrata dall'esperienza dell'utilizzatore (Ingegnere, Fisico, Chimico eccetera) e dallo studio teorico del Matematico.

Si passa quindi alla terza fase, ossia alla approssimazione numerica, e successivamente alla rappresentazione e interpretazione dei risultati ottenuti.

2.2. Risoluzione numerica e visualizzazione dei risultati

Non essendo in grado di ottenere un'espressione analitica della soluzione esatta, la soluzione viene calcolata, in modo approssimato, in alcuni punti del dominio del problema.

Per dare un'idea di come si procede nella risoluzione numerica considererò come esempio un problema particolarmente semplice, accessibile anche a chi non ha dimestichezza con la matematica. Immaginiamo di volere calcolare e visualizzare la distribuzione della temperatura in una stanza, che per semplicità assumiamo bidimensionale e quadrata. Su una parete si trova una porta-finestra a zero gradi, su un'altra parete un termosifone a 50 gradi, e le restanti pareti sono ben coimbentate, in modo che non ci sia dispersione di calore (*Fig. 1*)

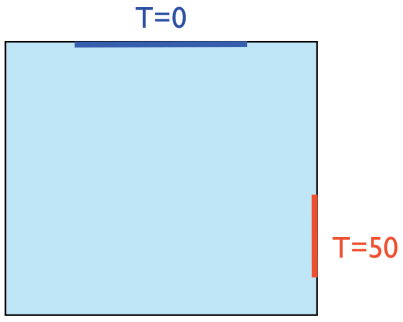


Fig. 1.

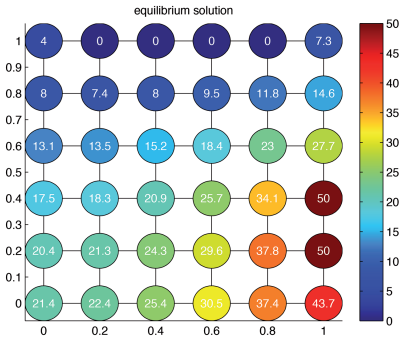


Fig. 2.

A regime, cioè dopo che la distribuzione del calore si è assestata e non varia più, la temperatura T è soluzione del seguente problema differenziale

$$\begin{cases} \Delta T = 0 & \text{all'interno della stanza} \\ T = 0 & \text{in corrispondenza della porta-finestra} \\ T = 50 & \text{in corrispondenza del termosifone} \\ \dfrac{\partial T}{\partial n} = 0 & \text{lungo le restanti pareti} \end{cases}$$

La lettera greca Δ indica l'operatore di Laplace, cioè la somma delle derivate seconde rispetto alle variabili spaziali, nel nostro caso x e y: $\Delta T = \dfrac{\partial^2 T}{\partial x^2} + \dfrac{\partial^2 T}{\partial y^2}$.

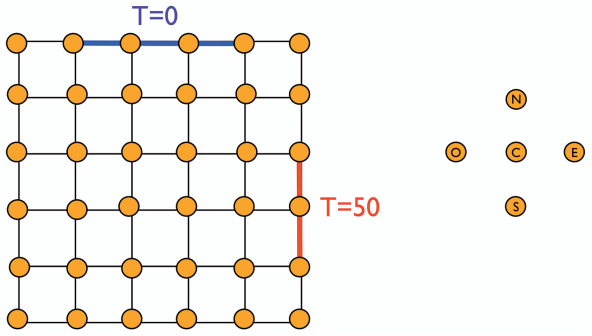


Fig. 3.

Nonostante si tratti di un problema molto semplice, non siamo in grado di risolverlo in modo esatto, e ci accontentiamo di calcolare la soluzione in modo approssimato. Il metodo numerico più antico, nonch e il più semplice da spiegare, è il cosiddetto metodo delle differenze finite: dopo avere suddiviso la stanza con una griglia di quadrati, in ogni punto interno della griglia si approssimano le derivate seconde con opportune differenze dei valori della temperatura nei punti vicini (*Fig. 3*).

Nel caso dell'esempio, assumendo per semplicità che la distanza tra i punti della griglia sia pari a 1, questo equivale ad approssimare la soluzione in ogni punto interno come media dei valori nei quattro punti vicini (Nord, Sud, Est e Ovest):

$$T^C = \frac{T^N + T^S + T^E + T^O}{4}$$

Nei punti lungo il termosifone e la porta-finestra la temperatura è nota e data da $T = 50$ e $T = 0$ rispettivamente. Nei restanti punti lungo il bordo della stanza si pùo imporre che il valore di T sia uguale al valore nel punto interno più vicino. In questo modo si scrive un sistema di 36 equazioni che viene risolto su computer. Il risultato fornito dal computer è una lista di numeri (in questo caso 36, uno per ogni punto della griglia) che rappresentano i valori della temperatura nei 36 punti della griglia e che vanno opportunamente visualizzati. Un esempio di rappresentazione è quello della *Fig. 2*, ma chiaramente non fornisce un'idea accurata di come si distribuisca la temperatura. Una rappresentazione come quella della *Fig. 4* sembra essere migliore per cui, procedendo in questo modo e suddividendo in griglie sempre più fitte si arriva ad una rappresentazione molto chiara, come si vede nella *Fig. 5*. Lo scopo di questo semplice esempio è duplice: da una parte sottolineare come la scelta del metodo numerico per approssimare la soluzione sia importante per ottenere risultati con un elevato grado di accuratezza, dall'altra mostrare come la rappresentazione dei risultati sia anch'essa pure di fondamentale importanza per una corretta lettura dei risultati stessi e conseguente comprensione del fenomeno fisico.

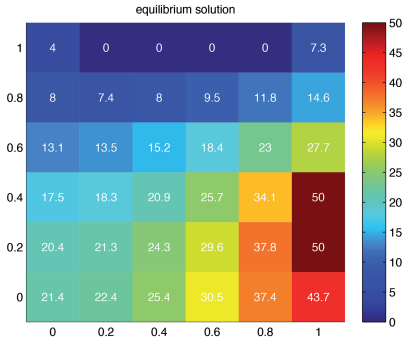


Fig. 4.

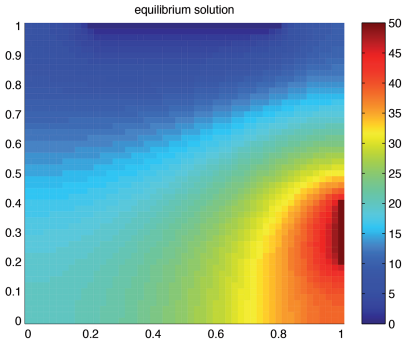


Fig. 5.

Nel caso dell'esempio appena descritto le varie fasi sono ben individuate ed in sequenza: modello, approssimazione numerica, risoluzione e visualizzazione dei risultati. In generale però, come si pùo facilmente intuire, le nostre fasi non vengono portate avanti in successione (come le

ho accennate), ma con un continuo *va e vieni*, dove gli esiti e le difficoltà nelle fasi *seguenti* inducono a ritoccare (a volte in qualche dettaglio, altre volte in modo sostanziale) la formulazione e il trattamento delle fasi *precedenti*. Se vi è mai accaduto di vedere all'opera un gruppo di formiche (o di api, o di molti altri animali) sappiate che la stessa apparente confusione e lo stesso disordine avviene, molto spesso, anche nella ricerca scientifica. Così come, spessisimo, in entrambi i casi il risultato finale è di ottima qualità (potenza del disordine..).

Ma attenzione: praticamente sempre le equazioni differenziali e/o i modelli su cui lavorano i matematici dipendono in generale da tutta una quantità di *coefficienti* (le proprietà dei materiali, le condizioni dell'ambiente, etc.) che spesso sono noti solo in modo approssimato, e conseguentemente le loro soluzioni dipendono anch'esse da questi coefficienti.

Se non si conoscono abbastanza bene alcuni coefficienti si risolve, sempre in modo approssimato, il problema per diversi valori dei coefficienti, scegliendo casi particolari in cui siamo in grado (as esempio, in un esperimento di laboratotio controllato) di *misurare* il comportamento della soluzione esatta. Se la soluzione numerica e la misurazione in laboratorio non sono abbastanza vicine vorrà dire che la risoluzione numerica è stata fatta male (come sostiene il *matematico modellista*) oppure che la equazione da risolvere non descrive adeguatamente il fenomeno (come sostiene il *matematico numerista*). A volte i due si mettono d'accordo, e se la prendono con chi ha preparato l'esperimento....

L'esempio delle formiche comincia a sembrare meno paradossale...

Ma c'è di peggio. In altri casi, ancora più difficili, la mancata conoscenza dei coefficienti pùo essere drammatica e (dal punto di vista scientifico), quasi insormontabile. Un chiaro esempio di una situazione di questo genere l'abbiamo, purtroppo, di fronte proprio in questi giorni: la propagazione, in una popolazione, di un virus. Non è molto difficile scrivere equazioni (in fondo abbastanza semplici) che predicano la probabile evoluzione del numero di contagiati da un giorno all'altro. Purtroppo, la semplicità di tali equazioni viene pagata con l'incertezza sui coefficienti. Conoscendo, oggi, il numero di contagiati (sintomatici e asintomatici), potremmo predire abbastanza facilmente il numero probabile di contagiati domani se *solo* sapessimo: la probabilità media di guarigione, la probabilità media di decesso, la frequenza media dei contatti, la percentuale di popolazione immune, e la probabilità che un contatto tra un contagiato e un "non immune" dia origine a un nuovo contagio... Insomma: basterebbe conoscere una manciata di coefficienti e sapremmo fare delle

previsioni abbastanza accurate. Purtroppo questi stessi coefficienti non sono facili da stimare, e inoltre variano nel tempo e nello spazio. Inoltre, il numero stesso di contagiati (sintomatici e asintomatici) per ogni giorno passato è conosciuto solo in modo molto approssimativo. Quello che si pùo tentare di fare (e che viene tentato) è di procedere per tentativi ed errori rifacendosi ai risultati precedenti: partendo (con notevole ottimismo) dalla conoscenza del numero di contagiati a una certa data "D", si possono fare ipotesi approssimative sui valori dei coefficienti, e simulare (con le nostre equazioni) l'evoluzione del fenomeno: se i risultati delle equazioni collimano con le misure fatte nei giorni successivi a "D", possiamo pensare che la nostra stima dei coefficienti sia corretta. Altrimenti si ripete il calcolo con una stima differente. La bravura consiste nel capire (in base al confronto tra la stima precedente e la misura) in che direzione correggere i coefficienti per avere un risultato migliore: per niente facile!

2.3. Cenni ad altri metodi numerici

In molti casi pùo risultare complicato usare reticoli regolari come quello della *Fig. 3*. Per esempio, quando il dominio computazionale presenta una geometria complessa (come il telaio di un'automobile o il motore di un aereo), o il problema considerato ha coefficienti con brusche variazioni, o quando l'accuratezza richiesta alla soluzione non è omogenea sul dominio (in un crash test su un autoveicolo, l'accuratezza richiesta è maggiore in prossimit della zona di impatto). Una griglia di triangoli o quadrilateri (tetraedri o esaedri in 3 dimensioni) si presta meglio a seguire le possibili oscillazioni e/o geometrie complicate. Nella *Fig. 6* si vede un esempio molto semplice di una griglia bidimensionale; i pallini rossi sono i punti del domnio in corrispondenza ai quali verrà calcolata la soluzione approssimata. La *Fig. 7* mostra invece un esempio di una griglia tridimensionale per la discretizzazione dell'ala di un aereo.

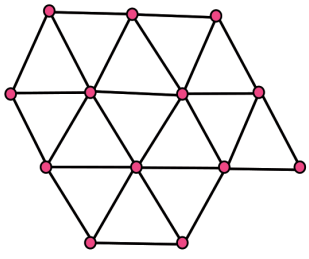


Fig. 6. Esempio di mesh.

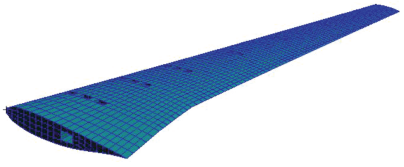


Fig. 7. Mesh per ala di aereo.

Si sono allora cercati altri metodi in alternativa alle differenze finite che permettessero di ampliare il campo di applicazioni e i problemi da trattare. Tra questi va certamente menzionato il metodo cosiddetto degli Elementi Finiti. Introdotto negli anni sessanta del secolo scorso, mantiene una posizione dominante nel panorama delle tecniche numeriche di approssimazione ed è tuttora il metodo largamente più diffuso e utilizzato nei codici sia di ricerca che industriali. Senza entrare nei dettagli, il metodo non si limita a calcolare solo i valori della soluzione approssimata nei nodi della griglia ma fornisce una soluzione approssimata come superficie polinomiale a tratti.

In tempi più recenti si sono esplorati nuovi metodi di approssimazione in cui, partendo da un approccio e una formulazione simili a quelli del metodo degli elementi finiti, si potessero usare decomposizioni in poligoni (o poliedri in tre dimensioni) di forma arbitraria. Tra i vari approcci mi limito a menzionare, per motivi di puro orgoglio, il metodo introdotto dal gruppo di ricerca di cui faccio parte, che abbiamo chiamato Metodo degli Elementi Virtuali, e che sta riscuotendo molto successo nella comunità scientifica internazionale. Le motivazioni che ci hanno spinto in questa direzione sono essenzialmente due. Da una parte griglie poligonali e poliedriche sono molto facili da generare con programmi esistenti. Dall'altra, in molte applicazioni il loro uso sembra essere la cosa giusta da fare, soprattutto quando si ha a che fare con materiali eterogenei e irregolari. Prendiamo ad esempio il problema dello sfruttamento dei giacimenti petroliferi (*Fig. 8*).

Dopo che il primo gigantesco getto di petrolio è fuoriuscito dal terreno, solo circa il 5% del petrolio presente in profondità è stato liberato e bisogna recuperare il restante 95%. Il modo di procedere attualmente in uso è il seguente: si inietta nel terreno un liquido (tipicamente acqua, in una prima fase) che deve spingere il petrolio verso il pozzo di raccolta. Se per esempio la pressione del liquido iniettato è troppo alta, si rischia di recuperare solo acqua. Se è troppo bassa, si recupera il petrolio a ritmi troppo lenti. Da cui l'importanza della simulazione numerica del fenomeno, che a sua volta ha come stadio iniziale la rappresentazione numerica del terreno, fatto di rocce di dimensioni e forme differenti che vanno ovviamente prese in considerazione con un livello sufficiente di dettaglio (e una mesh di poliedri ad-hoc opportunamente scelta pùo rivelarsi molto utile). La simulazione numerica di giacimenti petroliferi è uno dei molti problemi realistici per i quali il Matematico applicato (e in particolare l'Analista Numerico) pùo fornire, e di fatto fornisce, competenze utili per individuare il metodo più appropriato per la risoluzione numerica. La *Fig. 9* si riferisce al problema

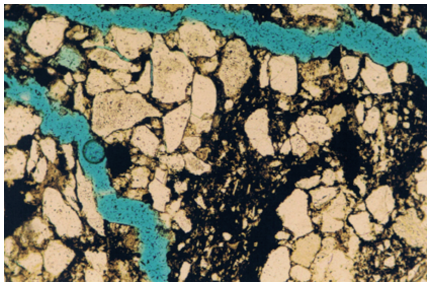


Fig. 8. Giacimento petrolifero.

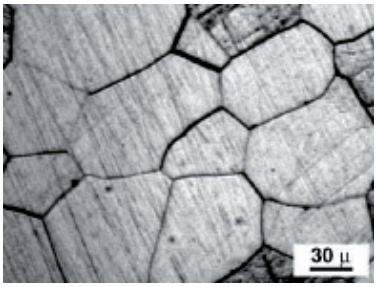


Fig. 9. Crescita di cristalli.

di studiare la crescita di cristalli. Ho volutamente accostato le *Figg. 8* e *9* per mostrare problemi in cui i domini computazionali hanno dimensioni molto diverse, di vari ordini di grandezza: dai chilometri cubici di un giacimento petrolifero a poche centinaia di micron cubici nel caso dei cristalli, ma i due problemi presentano strutture molto simili e in entrambi i casi per ottenere una accuratezza accettabile servono decomposizioni adeguate. I campi di applicazione di tutti questi metodi sono svariati e mi limito ad alcuni esempi a carattere indicativo: lo studio del flusso attorno ad un ostacolo (ad esempio l'aria attorno allo spoiler di un auto), oppure del flusso sanguigno attraverso la valvola aortica, e per problemi di questo e di altro tipo lo studio e lo sviluppo di metodi per raffinare la mesh in modo automatico in funzione della soluzione. Questo si rende necessario in problemi in cui la soluzione presenta dei bruschi cambiamenti, come avviene ad esempio per la pressione dell'aria sull'ala di un aereo: per passi successivi e tramite algoritmi opportuni si individuano le zone dove la soluzione numerica passa da valori bassi a valori alti e si infittisce la mesh solo in queste zone.

Passo ora a mostrare alcuni casi di fallimenti, dovuti a piccoli errori in una qualunque delle fasi, per far capire quanto possa essere delicato il calcolo scientifico.

3. Non sempre va tutto bene

Nel corso degli anni si sono purtroppo verificati anche fallimenti, che in alcuni casi hanno provocato disastri economici non indifferenti. Quasi sempre i fallimenti sono dovuti a piccoli errori, all'apparenza trascurabili, come negli esempi che vado a mostrare.

3.1. Un errore nel modello: il crollo del Tacoma bridge

Costruito tra il 1938 e il 1940 (senza computers), il ponte sospeso sul fiume Tacoma nello stato di Washington era un gioiello architettonico, il terzo ponte sospeso più lungo al mondo dopo il Golden Gate e il George Washington. Da subito il ponte mostrò tendenza ad oscillazioni verticali in presenza di forti venti, che tuttavia non destarono troppa preoccupazione, anzi furono da alcuni interpretate come indice di una buona elasticità. Dopo soli 4 mesi dall'inaugurazione il ponte fu investito da raffiche di vento a 42 miglia orarie, iniziò a torcersi e a produrre oscillazioni sempre più ampie fino a che, dopo un'ora e 10 minuti, cedette e crollò. La commissione di inchiesta non arrivò a formulare una causa precisa ma una serie di ipotesi di lavoro. Fu un ingegnere italiano, Giulio Krall, il primo a indicare una spiegazione tecnica del crollo. I calcoli da lui effettuati stabilirono che la velocità critica del vento per il raggiungimento del punto di rottura della struttura dovesse essere di 67 km/h, praticamente quasi coincidente con l'effettiva velocità del vento di quel giorno. Il progetto non aveva previsto che si potesse verificare una situazione così poco probabile di un vento che potesse entrare in risonanza con la frequenza di vibrazione della struttura. Fortunatamente non ci furono vittime, tranne un cane intrappolato in un auto e talmente terrorizzato da rendere impossibile l'intervento dei soccorritori.



Fig. 10. Crollo del Tacoma.



Fig. 11. Ariane 5.

3.2. Un errore di software: il disastro di Ariane 5

Il 4 giugno 1996 il vettore Ariane 5, progetto della Agenzia Spaziale Europea costato 10 anni di lavoro e 7 miliardi di dollari, fu lanciato nelle spazio e fatto esplodere dopo circa 40 secondi dal lancio per evitare che precipitasse a terra a causa di un malfunzionamento. Fu aperta un'inchiesta per individuare le cause del fallimento e dopo due settimane di lavoro la commissione preposta giunse alla conclusione che si era trattato di un errore di software, e precisamente di una erronea rappresentazione dei numeri nel computer. Il software di Ariane 5 era una modifica di quello relativo al modello precedente, Ariane 4. Nel processo di conversione un indicatore di velocità fu convertito da una rappresentazione decimale a 64 bit in una intera a 16 bit con segno: il numero era più grosso di 32.767, cioè del massimo numero rappresentabile come intero a 16 bit con segno e mandò in tilt il sistema. Ariane ricevette il comando di procedere con velocità negativa ed iniziò a puntare verso la terra. Per evitare danni maggiori la navetta fu fatta esplodere e questo pose fine a molte delle ambizioni spaziali europee.

3.3. Un errore nella approssimazione numerica

Si tratta del crollo dello Sleipner-A, una piattaforma petrolifera al largo delle coste della Norvegia: una costruzione gigantesca che poggiava su 24 pilastri piantati a circa 82 metri sul fondo marino. Il ponte superiore, con un'area di 16.000 metri quadri, pesava circa 57.000 tonnellate ed era in grado di ospitare 200 persone con un equipaggiamento di attrezzi per gli scavi di 40.000 tonnellate. Nella simulazione numerica ad elementi finiti con il codice NASTRAN, all'epoca il miglior codice esistente, la sezione circolare dei pilastri fu approssimata con una poligonale, e l'interstizio fra i pilastri (cosiddetto tri-cell) divenne un solo triangolo della griglia. All'epoca, nel 1991, i computers non erano potenti come quelli attuali, e l'uso di una griglia più fitta sarebbe stato praticamente impossibile. D'altra parte, considerando che la distanza fra i pilastri era di soli 6 metri (apparentemente trascurabile triangolo poteva sembrare più che sufficiente. Sfortunatamente la griglia usata era troppo rada per garantire una buona accuratezza dei risultati. Come conseguenza gli sforzi d itaglio che la struttura poteva sopportare furono sottostimati de l47%, e questa sottostima influì pesantemente sulla realizzazione del progetto. furono costruiti sufficientemente spessi per

sopp stimato in circa 700 milioni di dollari, il crollo d entità 3 sulla scala Richter.



Fig. 12. Sleipner-A.

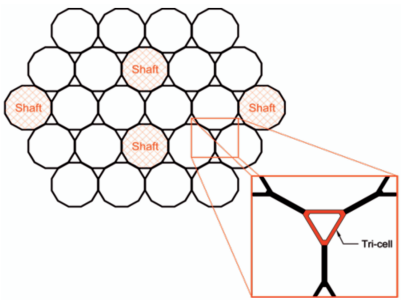


Fig. 13. Tri-cell.

4. Il ruolo del Matematico nel Calcolo Scientifico

Quanto ho esposto finora non chiarisce forse quale sia il ruolo del Matematico applicato nel Calcolo scientifico, che potrebbe sembrare del tutto simile a quello dell'Ingegnere. In effetti la linea di confine tra le due discipline è molto sottile e in alcuni casi addirittura indistinguibile. Tuttavia l'approccio è spesso sostanzialmente diverso. Per l'Ingegnere la priorità è risolvere un problema, o eventualmente una classe di problemi tutti di tipo molto simile. Ad esempio, si vuole decidere la forma e la struttura di un oggetto (diciamo, l'ala di un aereo) cercando di raggiungere un compromesso ideale tra peso, costo, prestazioni, robustezza, eccetera. Questo problema viene in generale affrontato simulando il comportamento di molte possibili scelte in molte differenti situazioni. E per ogni scelta della struttura e della forma dell'ala, e del tipo di situazione (volo a regime, atterraggio, decollo, vento trasversale, etc) si cerca di capire, con una simulazione al calcolatore, quale sarebbe il comportamento dell'aereo con la scelta in questione. In questo caso si faranno moltissime simulazioni di moltissime ali possibili, ma tutte solo leggermente diverse l'una dall'altra: come dicevo, si devono risolvere numerosissimi problemi tutti di tipo molto simile. L'Ingegnere si aspetta che, in generale, il metodo o i metodi di discretizzazione e risoluzione che hanno dato predizioni attendibili su uno di questi problemi daranno predizioni attendibili anche per tutti i problemi simili. E solo se palesemente questo, per

alcune tipologie, non accade cercherà di adattare (o a volte anche solo rabberciare) il metodo per farlo funzionare ancora. Nel fare questo userà soprattuto la sua grande conoscenza della natura fisica del problema, e cercherà di capire le caratteristiche fisiche dei problemi per i quali il metodo non funziona bene. Per contro l'obiettivo del matematico è spesso, in primo luogo, quello di tracciare con la massima precisione possibile la linea di demarcazione tra i problemi a cui un certo metodo pùo essere applicato con successo ed i problemi per i quali necessita una correzione. E nel fare questo la cosa più importante è l'incontro tra la struttura matematica del metodo e la struttura matematica del problema. Per valori molto diversi di certi parametri (diciamo: molto grandi, oppure molto piccoli) la struttura matematica del metodo potrebbe, in un certo qual senso, "sbilanciarsi", e mutare le proprie caratteristiche fondamentali. Sarà in questi casi che la correzione si rende necessaria, e la natura matematica della correzione da fare sarà resa evidente dalla comprensione della natura dello sbilanciamento. Per l'ingegnere il successo sarà un progetto di ala migliore di quanto realizzato finora. Per il matematico il successo sarà aver capito con precisione quando il metodo smette di essere affidabile, e cosa bisogna fare esattamente per rimetterlo in sesto. Per uno specifico problema, in generale, l'ingegnere avrà l'indubbio vantaggio di una molto maggiore comprensione della fisica coinvolta. Ma l'arma segreta del matematico applicato è spesso la capacità di riconoscere la stessa struttura matematica in problemi di natura fisica molto diversa. Questo gli permette di trasferire aspetti fondamentali della conoscenza da un dominio applicativo ad un altro, a volte totalmente differente, e di vedere quindi il problema sotto un punto di vista innovativo. Questo gli permette anche di esportare diagnosi e cure da un settore applicativo ad altri settori applicativi di natura molto diversa, ma con struttura matematica simile. Per quanto la cosa possa sembrare poco credibile ad un profano, equazioni quasi identiche descrivono il movimento di una navetta spaziale in un campo gravitazionale e quello di una particella carica elettricamente in un campo elettrico. O ancora, equazioni molto simili descrivono il moto di un fluido viscoso e la deformazione di un corpo elastico. Oppure la diffusione di un inquinante in un fiume e il movimento degli elettroni in un dispositivo semiconduttore.

Affido la conclusione di questo breve discorso a due grandi scienziati del passato e, tengo a precisare, nessuno dei due Matematico in

senso stretto, ma entrambi convinti sostenitori dell'importanza della Matematica.

> *La filosofia naturale è scritta in questo grandissimo libro che continuamente ci sta aperto innanzi agli occhi, io dico l'universo, ma non si pùo intendere se prima non s'impara a intender la lingua e conoscer i caratteri nei quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi e altre figure geometriche, senza i quali mezzi è impossibile a intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro labirinto*
>
> Il Saggiatore (1623), Galileo Galilei

> *Nessuna certezza delle scienze è dove non si pùo applicare una delle scienze matematiche.*
>
> Leonardo da Vinci

GASTRONOMY UNIVERSE BY EYES OF PHYSICISTS

Nota di ANDREY VARLAMOV (*) and ATTILIO RIGAMONTI (**)

(Adunanza del 4 aprile 2019)

SUNTO. – Dopo un breve inquadramento degli argomenti della presente nota nello scenario di una attività di divulgazione della Fisica attraverso la descrizione di fenomeni che intervengono nella vita quotidiana, si passa a considerare processi legati alla attività culinaria. In particolare, la cottura della carne o degli spaghetti, la preparazione della pizza o di un buon caffè. Sulla base di equazioni quali la equazione di diffusione del calore o del flusso e dei processi di trasferimento dello stesso e della temperatura (dedotte semplicemente attraverso considerazioni di analisi dimensionale) si perviene a formulare i principi che regolano questi processi e a fornire regole di esecuzione aventi validità scientifica.

***

ABSTRACT. – After a brief presentation of the general scenario to which the note is connected, the equations involved in some processes related to gastronomy and cuisine activity are derived, only by resorting to dimensional analysis. It is shown how the meat cooking time, or that one for spaghetti preparation can be quantitatively calculated on the basis of the heat conduction or thermal diffusion equations. The processes involved in baking of a good pizza (particularly in a wood oven!) or brewing a good coffee are addressed on the basis of scientific approaches.

(*) Istituto Lombardo Accademia di Scienze e Lettere, Milano. SPIN-CNR, Roma, Italy - MAGENTA. E-mail: varlamov@ing.uniroma2.it

(**) Istituto Lombardo Accademia di Scienze e Lettere, Milano. Università degli Studi di Pavia, Italy. E-mail: attilio.rigamonti@unipv.it

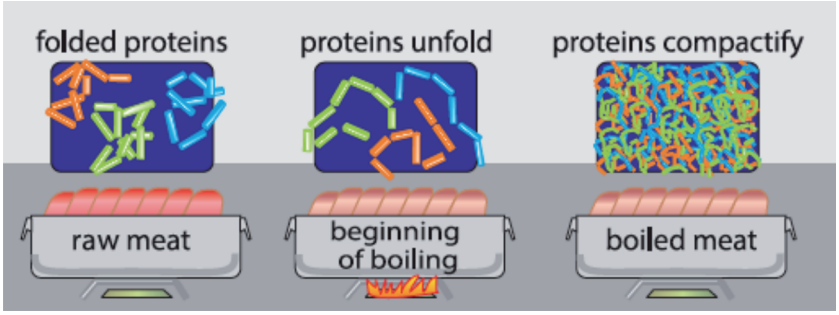
Foreword

The motivation and the roots of the present note go back to more than ten years ago. At that time the authors, strongly involved in researches on superconductivity of strictly scientific character, decided to devote some time to divulgation of Physics, in view of the increasing occurrence of this discipline in the everyday life. Thus several seminars in various institutions arose and finally the authors decided to publish a book. First in Italian language ("**Magico caleidoscopio della Fisica**", presented to our Istituto) and later on (2014) in an enriched form in French language ("**Le Kaleidoscope de la Physique**" with the additional contribution by Jack Villain, foreign Member of the Istituto Lombardo as well), by the editor Belin in Paris, even this presented and donated to the library of our Istituto).

After that, one of the authors, (A.V.) that is presenting this note to the audience, wanted to somewhat extend the description of the involvement of Physics into other aspects the every-day life, in the kitchen or directly in the preparation of foods typical of the Italian cuisine. Thus, a series of contributions arose and they are collected and deeply illustrated in the present note. As it will be seen, the general transformation of the proteins under cooking, the process of cooking of spaghetti or a big turkey or making a pizza or how to get a good coffee are illustrated, with the help of a few basic equation, still preserving scientific validity.

1. The basic formula of culinary

What does it mean to cook meat? There are biologists in the audience, and I apologize if using oversimplified picture. Meat is mainly composed of proteins. What are proteins? Geometrically, these are folded knots that unfold as the temperature rises. At a certain temperature - each type of meat has its own (for beef it is 74°C, for fish 47°C, for yolk - about 64°C) denaturation (compactification) of these chains occurs: they straighten form some kind of rugs. The difference between cooked and raw meat is whether these rugs have formed or not. So, at a temperature of about 75°C, such compactification occurs in beef, as a result of which boiled meat is obtained.



How does it differ from fried meat? The cooking process starts when we throw the piece of meat into the water. At normal atmospheric pressure, water boils at 100°C in an ordinary saucepan. You can calculate how long it takes for the temperature in the center of a piece of meat to reach the required 75°C and the compactification of proteins will take place in its entire volume. And frying?

Fried or grilled meat is very different from boiled meat. Different color, different taste. It turns out that if the temperature rises significantly above the denaturation one, then the so-called Maillard reaction can occur in the protein carpet. It was discovered in 1912 by the French chemist Louis Camille Maillard and consists of a chemical reaction between amino acids and sugar. It turns out that not all bonds are the same in our rug. In the plane, they are of one kind, across – of the other. At a sufficiently high temperature (for beef it is about 140°C), in the process of the Maillard reaction, some of the bonds are caramelized, which leads to appearance of the characteristic taste of a fried meat. Vegetable oil does not have a specific boiling point, there is a characteristic temperature when it starts to smoke. For olive oil this temperature is about 180°C.

I am not a chemist, and leave this topic, slippery for me. Let's go back to the equations, this is better. Let's formulate a model of the process of boiling a piece of meat. Martin Luther once advised that when bathing a child and pouring water out of the bathtub, be careful not to throw out the child together with the dirty water. This is what a physicist should always think about when formulating a model for some phenomenon. Unlike to mathematician, who solves exact problems, a physicist generally has to make some approximations. The world around us is very complex, and we must come up with a model that will

allow both to describe the essence of the phenomenon, and, at the same time, be solved by mathematical methods. Therefore, it is necessary to grasp the essence of the phenomenon and not overcomplicate the task.

Let's start the discussion with the case of a spherical piece of meat. Remember the anecdote about mathematician and physicist at the hippodrome? A physicist meets a mathematician there and asks:

- What are you playing on the sweepstakes?

The mathematician answers:

- No, I'm building a theory of the betting game ...

Physicist:

- Successfully?

Mathematician:

- So far I have managed to solve the problem of the movement of a spherical horse in an airless space ...

It happens that you take a boiled chicken and its meat is still raw near the bone ... I wish this did not happen and the denaturation process took place in the entire volume of the "meat ball". So, I have a spherical piece of meat with a diameter d, with an initial temperature of T_0, and it enters the environment with a temperature of $T_c = 100°C$.

It is clear that a heat flow begins to spread from the surface inward. To correctly describe this process, it would be necessary to write a differential equation known as the *heat conduction equation*. This equation states that at any point the time derivative of temperature is equal to the *thermal diffusivity* of the medium multiplied by the temperature Laplacian calculated at the same point

$$\frac{\partial T(r,t)}{\partial t} = \frac{\kappa}{\rho c}\frac{\partial}{\partial r}\left(r^2\frac{\partial T(r,t)}{\partial r}\right)$$

I think that several people sitting in the hall understood what I mean. And yet, we will not solve the differential equation, let's carry out its dimensional analysis.

In physics, many things can be understood without solving equations. It is enough to analyze the dimensionality of physical quantities: the dimensionality on the left side of the equation must coincide with that one on its right side. This allows one in many cases to understand what is happening. An intelligent student often manages to pass an exam without even being prepared in advance for the question asked, if he is able to write the correct formula from dimensional analysis. It is better if a person is able to write an estimate by dimensionalities, show-

ing that he understands the essence of the phenomenon, than simply is able to memorize.

Let us try to figure it out by ourselves using the dimensional analysis method. The temperature of denaturation of meat coincides by an order of magnitude with the boiling point of water (differs from it by 20 – 25%). Therefore, we assume that the time of "delivery" of the necessary temperature to the center of the solid sphere depends only on its material parameters and size: the thermal conductivity of the meat, its density, specific heat and radius. Therefore, we seek the dependence of the required time on the size of the sphere in the form:

$$\tau = \kappa^{\alpha} \rho^{\beta} c^{\gamma} R^{\delta}.$$

By comparing dimensions, we write:

$$[\tau] = [\kappa]^{\alpha} [\rho]^{\beta} [c]^{\gamma} [R]^{\delta}$$

The dimension of the thermal conductivity $[\kappa] = \frac{kg \cdot m}{s^3 \cdot K}$. Substituting it, side by side with the dimensions of density ($[\rho] = \frac{kg}{m^3}$), specific heat ($[c] = \frac{J}{kg \cdot K}$) and radius ([R] = m), and then comparing them in the right- and left-hand sides, one finds: α=−1, β=γ=1, δ=2. Thus, we conclude that

$$\tau = C_0 R^2 / \chi,$$

where $\chi = \kappa/(\rho c) = 1,4 \cdot 10^{-7} \frac{m^2}{s}$ and C_0 is an unknown constant of the order of unity.

From this formula is clear that if I take a ball with twice larger the diameter, then the time it takes for the required temperature to reach the center will quadruple.

You see, I, without straining you too much, got a nice formula. But there are more accurate people. One Englishman, Peter Burham, was not too lazy, and for boiling soft-boiled eggs, he derived the exact formula. How? He took the equation I wrote for you, rewrote it in elliptical coordinates (an egg is an ellipse with good accuracy), solved it carefully and got the exact formula:

$$t = 0.15 d^2 \log 2 \frac{(T_b - T_0)}{(T_b - T_f)}.$$

(in addition, it includes the initial temperature of the egg, the temperature of compactification of the yolk and the boiling point of water).

Now let's imagine to be in the USA. Today is the penultimate Thursday, of November, Thanksgiving day, about 40 million turkeys will be eaten. Usually, when you buy a turkey at the supermarket, the packaging says how long it needs to roast. However, each housewife has her own recipe. At times I have heard the advice for every pound of extra weight to add 15 minutes of turkey's stay in the oven. This is not true, we have already seen that the time of "temperature delivery" to the center depends on the mass nonlinearly, the obtained above equation for cooking time can be rewritten as:

$$\tau = C_1 \frac{c \cdot \rho^{1/3}}{\kappa} M^{2/3},$$

I went online and found this table of recommended times there:

Mass M (pounds)	Baking time τ (*hours*)
8	3,5
12	4,5
16	5,5
20	6,5
24	7,35

From the first glance I do not see the correlations here. The mass has grown by 4 times, and the cooking time has grown a little more than 2.

Let's use our formula. How the experimentalists elaborate their data? It is very important to choose the coordinates corresponding for the task. In our formula, we time appears in the first power while the mass appears in the power 2/3. It is why I propose to take the table and plot the corresponding data in the form of a graph in coordinates of time *versus* mass in the power 2/3. Look at what a nice straight line they fit!

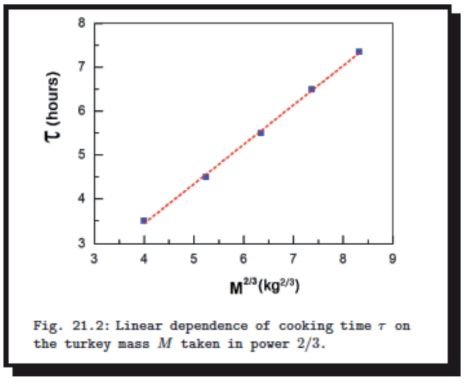


Fig. 21.2: Linear dependence of cooking time τ on the turkey mass M taken in power 2/3.

But this is not just the triumph of theory over practice. First of all, it is clear that from the linear dependence slope I can determine the thermal conductivity of meat yet is not the point now. It is much more important where this line goes. It can be seen that, with good accuracy, it goes to 0. This means that in reality only minutes are needed to undergo the denaturation process and the Maillard reaction. All hours of the turkey's stay in the oven are needed for the required temperature to reach the depth. What temperature? The temperature corresponds to the Maillard reaction, 140°C, if we want to eat baked turkey, not boiled one. And you need this turkey to be baked in full. You could set a higher temperature in the oven, but then there is a high probability of burning the turkey. The temperature should be set just above the required 140°C so that the heat wave goes inward but does not burn the periphery. Often, to save the surface, the turkey is pulled out of the oven, poured with fat, and then sent back. We will talk a lot about surface temperature when discussing the physics of pizza. In cooking, it is often necessary to try on two conflicting requirements: on the one hand, bake, on the other hand, do not burn.

2. Spaghetti and physics

Now the physics of pasta will be discussed. Today I will talk about the pasta of cylindrical shape. These are: capelli d'angelo - angel hair, the thinnest pasta with a diameter of only 0.85 mm, spaghettini, spaghetti, vermicelli, bucatini. All types of cylindrical pasta without a central hole along the axis of the cylinder vary in diameter from only one to two millimeters. Already when the diameter reaches 2.7 mm (bucatini), such a hole turns out to be necessary for some reasons. Below we will try to understand why.

What does it mean: to cook pasta? First, the pasta needs to be heated. If I put the pasta in cold water, it will not become eatable. Secondly, I have to soak its bulk with water, if I just put it into the oven, then the pasta also will not become eatable.

Above I discussed the problem of a heat transfer, here we must additionally solve the problem of the delivery of water to the center of the cylinder. The pasta was specially dried at the factory, and in the process of cooking one has to restore the water contents in it. You've

already seen a beautiful equation of heat conduction. The diffusion equation that describes how the concentration of water changes with time, turns out to be mathematically equivalent to the heat conduction equation. The same equations – the same solutions. When considering the issue of cooking pasta, it turns out to be more convenient to work not with the mass, but with the diameter, so we can rewrite the formula for cooking time as

$$\tau_{sp} = ad^2 + b.$$

The coefficient a is determined by the thermal conductivity and the diffusion coefficient, but this is not important to me, I will determine it empirically. As for the parameter b, we saw that for a turkey it is practically equal to 0. In the case of cooking pasta, the parameter b, as we will see, differs from zero, and I claim that its value depends on the nationality of the eater. Do you believe that Germans or Americans cook pasta for as many minutes as the Italian manufacturer writes on the wrapper? No, if they consider the pasta as the side dish. At the same time Italians undercook pasta, even compared to what is written on the wrapper! Why? Because the pasta, after being thrown into a colander, must still go to the pan with the sauce. And there it will still be cooked. Therefore, it is necessary to cook pasta for a minute or two less compared to what is written on the pack. Hence, for Italians, the coefficient b must be negative. I'll prove it to you now with numbers in my hands.

What does the term pasta "al dente" means? In my model, this corresponds to the fact that the water and temperature do not reach the cylinder axis. The spaghetti center remains uncooked. In practice, of course, this is not entirely true. We understand that complex physical and chemical processes take place in the dough during cooking - diffusion, heating, denaturation of starches... My model is simple - to deliver water and the required temperature almost to the center. Let's see if I am not throwing out the child with the dirty water, if I can build a reasonable theory.

I went to a supermarket and bought all the cylindrical pasta that was there: Capellini no. 1, Spaghettini no. 3, Spaghetti no. 5, Vermicelli no. 7, Vermicelli no. 8, Bucatini. The latter are already so thick that there is a hole in them. Then I went to the laboratory of my friend experimentalist, took a caliper and measured diameters of all of them. The results are collected in the following table:

Type of pasta	Diameter, external/internal (mm)	Experimental cooking time (min)
capellini no.1	1.15/-	3
spaghettini no.3	1.45/-	5
spaghetti no.5	1.75/-	8
vermicelli no.7	1.90/-	11
vermicelli no.8	2.10/-	13
bucatini	2.70/1	8

In our formula for cooking time there are two unknown parameters - a and b. I will not search for them on the Internet, they can be found in a simple way. Namely, I will take Spaghettini no. 3 and Vermicelli no. 8 as reference points (the extreme rows of the table should not be taken, this can lead to a larger error). For optimal preparation of spaghetti, the manufacturer recommends 5 minutes (as written on the pack). Let's take this as the value of the experimentally verified cooking time. For vermicelli no. 8 the corresponding time is 13 minutes. So, I have two equations connecting the known cooking times of two types of cylindrical pasta with their known diameters. The parameters **a** and **b** remain unknown. Solving two algebraic equations with two unknowns, we obtain

$$a = \frac{t_2 - t_1}{d_2^2 - d_1^2} = 3.4\,\text{min}/mm^2$$

$$b = \frac{d_2^2 t_1 - d_1^2 t_2}{d_2^2 - d_1^2} = -2.3\,\text{min}.$$

Using the found parameters, we can "predict" the cooking times for all the other types of pasta we have. We can see that my formula works great in the middle of the table and fails at the edges. Indeed, it doesn't go well with Capellini: we have 2.2 minutes instead of the recommended 3, but with Bucatini, it's a big problem - here our formula gives 25 minutes instead of 8. Let's try to figure it out.

Let's start with bucatini. Please tell me what do you think, why it is the hole in its center? There are two possible reasons. First, if there is a hole, then there is no need to deliver heat up to the very center through the outer surface, since water will fill in there anyway. As a second explanation, perhaps someone will tell me that, thanks to the inner

hole, there will be two flows of heat, one from the outside, the other from the inside? Do you think there will be a heat flow from within or not? No, it will not. For occurrence of the heat flow it is not enough for the water inside the bucatini to have a temperature of 100°C. Take a pot of boiling water and place a small saucepan in it, will the water boil in the latter? No, it will not boil! It is not enough to be surrounded by media with temperature 100°C, it is necessary to have constant flow of energy from the external source. The same thing happens with the long bucatini. Water really enters at the first moment with a temperature of 100°C, but then where to get energy for heat transfer into the dough? Therefore, I will take my formula for cooking time of cylindrical pasta which takes into account only the flow from the outside, and subtract in it the size of the internal hole from the diameter, which no longer needs to be heated, there is already water with its 100°C. In result I get the theoretical cooking time as 7.5 minutes instead of the recommended 8. This is really great!

We see that I came up with a theory, it works in a certain range, but as soon as I went beyond it, the theory stopped working. In 25 minutes my dough will just turn into porridge! Then I extended my theory, take into account the central hole, and the theory works great again. This is a good example that shows how a theoretical physicist works. You come up with a certain model, it works within some limits, which you must stipulate in advance, but at the borders of the region it may fail, then the theory must be revised.

Now let's deal with the second nuisance. Capellini are thin and have a diameter only 1.15mm. The recommended time is 3 minutes, but we got 2.2. Why this 30% discrepancy? The matter of fact, that from our formula follows that there is such a thickness of cylinder at which you can eat corresponding pasta just raw. "*Al dente*" in my model means that the heat flow does not reach the center. Let's see how thick the pasta should be in my model in order to be eaten raw. The cooking time is set equal to 0, therefore

$$ad^2 + b = 0.$$

Here **a** is a positive value, and **b** is negative, therefore this equation has a real solution. We find that the corresponding critical value of d is 0.85 mm. Hence, according to my logic, a hypothetical pasta with a diameter of 0.85 mm does not need to be cooked at all. The value 1.15 mm is

already very close. One can see that, when the diameter was 2 mm, everything is fine, but when we approach 1 mm, the theory breaks down.

Pasta serves the industry as well. One day, a Swiss firm set out to learn how to knit fishing nets. The company's engineers made a mathematical model from which it followed that the most dangerous places in terms of the break are these in the nodes. And they decided to test it experimentally, but they didn't succeed on the experiments with a real network. In those days, they did not have a fast enough camera, and they could not keep track of how the nylon breaks. Then an ingenious solution was proposed. The engineers took spaghetti, boiled it, doused it with olive oil, tied it in knots and began to study the tearing process on a slow camera.

Has anyone seen pasta knotted during cooking? Nobody has seen. Why don't they knot? Do you think it's that simple? I asked myself the question, how long should the spaghetti be to make it knotted? Why does the wire from the headphones always get entangled but no pasta? Moreover, out of 5,000 amino acids, only few are knotted. I have a colleague and friend, Alexander Grossberg, a great expert in knots. Many years ago he was a professor at Moscow State University, now he is a professor in New York. Once I asked him to answer the above questions, but he said that the problem of knotting a long rope has not yet been fully solved. A formula has not yet been found that would analytically determine the knotting probability. There are numerical models, but no good theory.

From the general principles of statistical mechanics, it is clear that the probability of self-knotting of a cord is determined by the exponential to the power of which its length is divided by some quantity of the dimension of length $\gamma\xi$, all with a minus sign:

$$w = 1 - \exp\left(-\frac{L}{\gamma\xi}\right)$$

What is ξ here, and what is γ? The quantity ξ is the characteristic length of the cord at which it can change its direction to 90°. With headphones, the wire bends very easily, we can say that ξ is less than one centimeter. So, ξ is the characteristic length of the material, which characterizes its elasticity in bending. And γ is the numerical coefficient that is still obtained only as a result of computing on a computer. Surprisingly, γ turns out to be huge - about 300, a rare case. But that's why spaghetti and amino acids don't entangle. Let's estimate the likeli-

hood of knotting for spaghetti. Let's take ξ = 2-3 cm. Then, in order for spaghetti to knot with a 10% probability, their length should be about 1 m. And the standard for pasta is 23 cm.

Finally, I would like to conduct an experiment. Everyone knows about the Nobel Prize. Thanks to journalists, everyone probably knows about Ig Noble Prize awarded by Harvard University. The Ig Nobel Prizes honor achievements that make people laugh, then think. The prizes are intended to celebrate the unusual, honor the imaginative — and spur people's interest in science, medicine, and technology. The Ig Noble Prize of 2007 was given to two Frenchmen, Basile Audoly and Sebastien Neukirch of the Université Pierre et Marie Curie, for their analysis that explains why uncooked spaghetti breaks into several pieces when it is bent.

Take the spaghetti by the ends and bend it into an arc. It seems obvious that with further bending, sooner or later it, somewhere near the middle, will break into two parts. It turns out that in this case our intuition lets us down, there will almost always be three or more fragments.

This unusual property of spaghetti attracted the attention of many scientists, among whom was Richard Feynman. And only quite recently, in August 2005, thanks to the research of two French physicists, Audoly and Neukirch, a quantitative explanation of this phenomenon appeared, which, possibly, will find its further development in the science of the strength of materials and applications in construction. Scientists have investigated the behavior of a thin elastic bar under bending deformation. To do this, they used the so-called Kirchhoff equation, writing it down for a curved elastic rod, first with two fixed ends. They then studied what would happen to the stress distribution in the bar when it breaks at an arbitrary point. The solution turned out to be possible to find only numerically, however, this also allowed understanding physics. Let us assume that, as a result of the applied mechanical bending stress, the first fracture occurs at the weakest point of the spaghetto. It would seem that after it both formed parts should return to their equilibrium positions. However, this process takes place in a very complex and ambiguous way. Immediately after the first fracture, elastic bending waves appear in both parts of the bar, which propagate along them. They lead to an increase in certain areas of the bar of the local bending stress in comparison with the already existing static bending preceding the first fracture. As a result, in these places of growth of the bending stress, further fractures of the bar can occur. Of

course, such elastic waves, generated by the first kink, decay with time, however, at a certain ratio between the length of the rod and its elastic properties, they can lead to subsequent fragmentation of the rod. It is remarkable that, having carried out their complex calculations, the scientists were convinced of the correctness of their conclusions with the help of slow-motion shooting of experiments with spaghetti with a high-speed movie camera.

3. Pizza physics

In ancient Egypt there was a tradition to celebrate the birthday of the pharaonic by eating cakes sprinkled with aromatic herbs. The ancient Greek historian Herodotus in his writings mentions Babylonian recipes for various kinds of tortillas baked from flour.



Virgil describes the process of making tortillas in the hearth by peasants from the time of ancient Rome, much like the baking of modern pizza. Apparently, the modern Italian word “pizza” has a Latin etymology: it is consonant with “pinsa” - the past participle of the verb “pinsere”, which means “to crush, to grind” in Latin. Traces of dishes similar to modern pizza can be found in historical documents from the Middle Ages and Renaissance of various Mediterranean countries. But she found her true home, from which pizza set out to conquer the world, in the labyrinths of Naples’ alleys in the first half of the 18th century.



In the triumphant procession of pizza, two of its ingredients were decisive: mozzarella cheese and tomatoes. The first appeared in Naples thanks to the Germanic tribe of the Lombards, who, after the fall of the Roman Empire, descended from the North to Campania and brought herds of buffaloes with them. It is from the fatty milk of buffaloes in the vicinity of Naples that the true mozzarella is still produced. It is very tender not only in taste - mozzarella is afraid of cold (after being in the refrigerator it becomes "rubbery") and lives only a few days, floating in its own whey.





The second ingredient - tomatoes - was brought to Europe from Peru by the Spaniards during the colonization of America. In the 16th

century, the Spaniards also ruled Naples; one of the positive consequences of their rule was the mountains of tomatoes on the shelves of greengrocers in the old city. At first, conservative Italians were distrustful of the overseas vegetable, but the Campanian climate turned out to be favorable for him, and gradually the tomato began its triumph in Italian cuisine, becoming the basis of many pasta sauces and, later, the most important ingredient in pizza.





Since the beginning of the 18th century, numerous shops, called "pizzerias", have opened in Naples, where, among other things, pizza is baked. Due to its cheapness and its wonderful taste, pizza is becoming an affordable dish both poor and rich. Soon the fame of pizza reaches the ears of the King of Naples and the Kingdom of the Two Sicily Ferdinand Bourbon, who, in order to taste this dish, neglects palace etiquette and goes to one of these pizzerias. From this moment on, pizzerias come into fashion, become establishments where only pizza is prepared. The most popular pizza in Naples at this time is the sailors' pizza - "Marinara". In its present form, it arose around 1730. To this day, this pizza is a flat disc of simple dough (yeast, water, salt and flour), covered with thinly chopped tomatoes and garlic, drizzled with a little olive oil, sprinkled with oregano (origan). Before getting on the table, Neapolitan pizza spends a couple of minutes in a hot oven, next to hot logs (the temperature is such that it's close to the melting of the aluminum shovel, with which the *pizzaiolo* puts the pizza into the oven: 450-480°C).

Another popular pizza recipe from this era dates from 1796 - 1810. Here, tomatoes, mozzarella and basil leaves are placed on a base of the same unleavened dough. These colors repeat the colors of the tricolor (vertical green, white and red stripes) popular in northern Italy at that time. In 1861, Garibaldi unites Italy, and the tricolor becomes a national symbol.





In 1889, the King of Italy Umberto I and his wife, Queen Margherita, came to Naples. According to legend, the best pizza maker in the city, Raffaele Esposito, bakes three different pizzas for them. Of all the three, the queen appreciated the pizza in the colors of the flag of Italy, so much so that since then this pizza has been named after her - "pizza Margherita". In 1989, the coauthor of this note A.V. was fortunate enough to be in Naples during the celebration of the centenary of Margherita pizza, when the whole city was celebrating the anniversary, the walls of houses were painted in red, white and green, and pizza was handed out on the streets for free.






In the process of the mass emigration of Italians at the beginning of the 20th century, pizza from southern Italy moved to northern Europe and overseas, conquering Grenoble and Hamburg, New York and Chicago, Tokyo and Melbourne. Pizza has taken root there so much that many are convinced of the American origin of pizza.

Having moved from Naples to Rome, I (A.V.) did not immediately recognize my favorite dish. In Roman pizzerias, instead of soft, fluffy, with towering wrapped edges of Neapolitan pizza, they usually serve thin, drier, the size and thickness of a thick pancake (4 - 5 mm), the so-called Roman pizza. On its surface, as in the case of Neapolitan pizza, all sorts of tasty things can lie: both mozzarella with ham and mushrooms ("Prosciutto e funghi"), and the contents of the same "Margherita", but still decorated with circles of peppered salami ("Diavola") , and a collection of four different cheeses ("Quattro for-

maggi") ... Served in Rome and "Napoli" - pizza with tomato, mozzarella, capers and anchovies. In Naples, on the contrary, pizza of such content is called "Romana" - Roman. It seems that you can combine anything with anything. However, this is not the case, there are classic combinations in Italian cuisine, and there are strict restrictions on imagination: for example, no one will sprinkle seafood with Parmesan cheese, mix fish and meat. This, however, does not apply to American pizzerias: pay for two "toppings" (This is the name of the agreed menu for the standard pizza toppings.) And in the kingdom of the "dibs" they will at least mix a herring with nutella.

"The neophyte is holier than the Pope," says the old proverb. Following her, I began to diligently delve into the secrets of making pizza. The first commandment that I heard from Italians is to try to always go to a pizzeria equipped with a wood-burning oven, not an electric oven. Good pizzerias take pride in their "forno", where the entire cooking process takes place in front of your eyes. Pizzaiolo masterly "sculpts pizza", puts it in the oven with a wooden or aluminum spade, in a moment - and it is already appetizingly bubbling with boiling cheese in front of you, calling to be immediately eaten and washed down with good beer.

Every time I (A.V.) visited the same "Two Lions" pizzeria near my home, I soon received valuable advice from a friendly pizza maker: come in for dinner either before 8 pm, or after 10 pm, when the place is half empty. This was also confirmed by behavior of one of the "lions" - a fat, smoky-gray cat, feeding at a pizzeria "client" [In ancient Rome, clients were called free citizens under the patronage of a patron - a noble citizen, often patrician]; during the influx of people, he went for a walk and was not at all interested in the contents of the plates sitting at the tables.

The reason for this limitation turned out to be very simple - in the throughput of the furnace. As the same pizzaiolo explained to me, the optimum temperature for the Roman pizza he makes in a wood-burning oven with a refractory brick bottom is 325-330°C. [This temperature, of course, depends on how the dough is prepared and stored. Friend Antonio prepares the dough in advance - a day before baking a pizza from it. After kneading well, he leaves the dough to "rest" for a couple of hours, then divides it into portions (for Neapolitan pizza - 180-250 g, for Roman pizza - less) - one per pizza, - and puts it in a wooden box in the form of balls. The dough ripens for another 4-6 hours, then it can already

be planted in the oven. But in our case, it is sent to the refrigerator.] At the same time, thin, Roman, pizza "matures" in the oven for 120-130 seconds. Thus, even putting them in the oven two at a time, the pizzaiolo will be able to serve 50-60 visitors per hour. And at "rush hours" there are about a hundred of them in the hall, and even a dozen people are piling up near the cash register, wanting to take the pizza, blazing with heat, home. Here Antonio, for the sake of the owner's income, raises the temperature in the oven to 390°C, pizzas fly out of the oven every seventy seconds, but their quality is not the same - the bottom and edges are slightly burnt, and the tomatoes are still damp ...

Since it is often difficult to find a pizzeria with a wood-burning stove (you need a chimney, but how can you make it in a multi-storey modern building?). Let's figure out what are its advantages over an electric oven, and is it possible to improve the latter so that he still made a passable pizza.

Let's start from afar. Imagine that your temperature rises, and your thermometer is broken during the last cold. Mom will put her hand to your forehead, hesitate and say: you have a cold, you have a fever, you are not going to school tomorrow. In physics, in order to understand a phenomenon, the problem should always be simplified first. Imagine that your mother is measuring the temperature of your forehead not with her fingers, but by touching her forehead to her forehead. Moreover, if the temperature of your forehead is 38°C, and her forehead is 36°C, then it is quite clear from the symmetry of the problem that on the border between the two foreheads the temperature will be set at 37°C, which mother will feel, because a heat flow will flow to her from your hot forehead.

Now let's imagine that the mother has a "copper" head, and the temperature is still the same 36°C. Then, it is intuitively clear that the temperature at the interface will drop, say, to 36.5°C: a copper head that conducts heat well (due to its high thermal conductivity) will take your heat away from the border between the foreheads into the depths. At the same time, it is clear that the removal will be the better, the less heat is required to be left in the border area; it increases (and the boundary temperature decreases) with a decrease in the specific heat capacity of the material.

In order to move from conversation to business, we remind the reader of the basic concepts and results of the physics of heat transfer. The term heat (heat energy) usually means the part of the body's energy

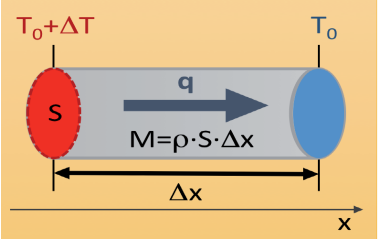
associated with the movement of its atoms, molecules or other particles of which it is composed. We inherited the concept of heat from the past of physics, heat, as scientists say, is not a function of the state of the body, its value depends on how the body came to this state. Heat, like work, is not a type of energy, but a quantity convenient for describing its transfer. The amount of heat required to raise the temperature of a unit mass of a substance by one degree is called the specific heat of the substance.

It is clear that in the SI system of units, the specific heat is measured in J kg^{-1} K^{-1}. When two bodies with different temperatures are brought into thermal contact, the heat from the more heated one rushes to the less heated one. In the simplest case of a homogeneous non-uniformly heated body, the heat flux q, *i.e.* the amount of heat flowing per unit of time in the direction of temperature change through a unit area perpendicular to this direction is

$$q = \frac{cM\Delta T}{S\Delta t} = c\rho\frac{(\Delta x)^2}{\Delta t}\left(\frac{\Delta T}{\Delta x}\right) = -\kappa\frac{dT}{dx},$$

where ρ is the mass density. Assuming that Δx is small, we identified the value in parentheses as the derivative of the temperature by the coordinate x and took into account the fact that the temperature decreases in the x direction. In the general case, q is a vector and the derivative should be replaced by the gradient of T, which describes the rate of temperature change in space. The coefficient κ is the thermal conductivity, which describes the ability of a material to transfer heat when a temperature gradient is applied. The presented equation expresses mathematically the so-called Fourier's law, which is valid when the temperature variation is small.

Next, let us analyse how a temperature front penetrates a medium from its surface, when a heat flow is supplied to it (see *Figure*).



Assume that during time t the temperature in the small cylinder

of the height L(t) and cross-section S has changed by ΔT. Let us get back to the equation and rewrite it by replacing Δx by L(t):

$$\frac{c\rho L(t)\Delta T}{T} = \kappa \frac{\Delta T}{L(t)}.$$

Solving it with respect to the length one finds:

$$L(t) \sim \sqrt{\frac{\kappa t}{c\rho}} = \sqrt{\chi t},$$

i.e., the temperature front enters the medium by the square-root law of time. The time after which the temperature at depth L will reach a value close to that one of the interface depends on the values κ, c, and ρ. The parameter $\chi = \kappa/(c\rho)$, as already was mentioned above, is called the thermal diffusivity or coefficient of temperature conductivity. The heating time of the whole volume can be expressed in its terms: $\tau \sim L^2/\chi$.

Of course, our consideration of the heat penetration problem into a medium is just a simple evaluation of the value L(t). A more precise approach requires solution of differential equations. Yet, the final result confirms our estimation just correcting the numerical factor:

$$L(t) = \sqrt{\pi \chi t}.$$

Now that we know how heat transfer works, let us get back to the problem of calculating the temperature at the interface between two semi-spaces: on the left with parameters κ_1, c_1, ρ_1 and temperature T_1 at $-\infty$,and on the right with parameters κ_2, c_2, ρ_2 and temperature T_2 at ∞. Let us denote the temperature at the boundary layer as T_0. The equation of the energy balance, *i.e.*, the requirement of the equality of the heat flowing from the warm, right semi-space through the interface to the cold, left semi-space, can be written in the form

$$q = \kappa_1 \frac{T_1 - T_0}{\sqrt{\pi \chi_1 t}} = \kappa_2 \frac{T_0 - T_2}{\sqrt{\pi \chi_2 t}}.$$

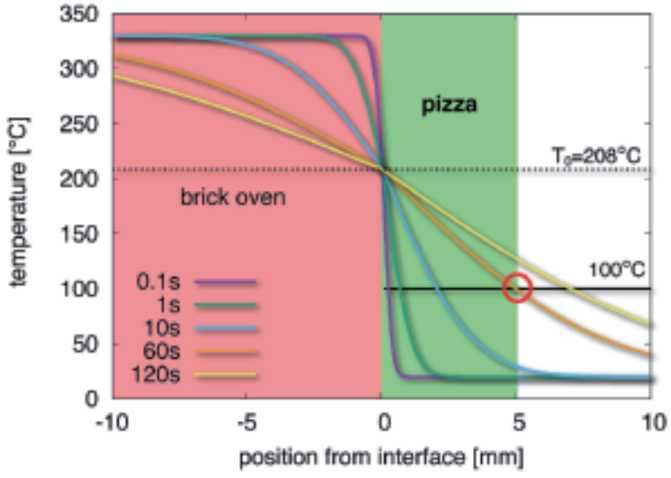
Here we simplified the problem assuming that all temperature changes happen at the corresponding temperature dependent length. Solving this equation with respect to T_0 one finds that

$$T_0 = \frac{T_1 + \nu_{21} T_2}{1 + \nu_{21}},$$

where

$$\nu_{21} = \frac{\kappa_2}{\kappa_1}\sqrt{\frac{\chi_1}{\chi_2}} = \sqrt{\frac{\kappa_2 c_2 \rho_2}{\kappa_1 c_1 \rho_1}}.$$

First of all, we note that time was not included in our formula, *i.e.* in the first approximation, the temperature at the boundary does not change. Thus, we quantitatively substantiated the intuitive answer given at the beginning at 37°C for the temperature at the border between two identical foreheads. If the right, cold, forehead is copper, then for him the temperature of the cold copper forehead almost does not increase from contact with the hot usual.



We are now ready to discuss the benefits of a wood-burning stove. Let's calculate the temperature at the bottom of the pizza we put in the wood-burning oven. All parameters necessary for us are given in the table.

Material	Heat capacity $c\,[\frac{J}{kg\cdot K}]$	Thermal conductivity $\kappa\,[\frac{W}{m\cdot K}]$	Mass density $\rho\,10^3[\frac{kg}{m^3}]$	Temperature conductivity $\chi\,10^{-7}[\frac{m^2}{s}]$	ν_{21}
dough (2)	$2-2{,}5\cdot 10^3$	0,5	0,6–0,8	2,5–4,2	1
steel (1)	$4{,}96\cdot 10^2$	18	7,9	45	0,1
fire brick (1)	$8{,}8\cdot 10^2$	0,86	2,5	4,0	0,65

Assuming the initial temperature of the pizza dough as 20°C, the temperature inside the oven, as our pizzaiolo claimed, being about 330°C, and the ratio $\nu_{d,fb} = 0,65$ we find for the temperature at the boundary layer between the oven surface and pizza bottom

$$T_0^{bp} = \frac{330°C + 0,65 \cdot 20°C}{1,65} \approx 208°C.$$

As we know from the words of the pizzaiolo, a pizza is perfectly baked in two minutes under these conditions. Let us now repeat our calculations for the electric oven with its baking surface made of steel. For an electric oven the ratio $\nu_{d,s} = 0,1$, and if heated to the same temperature of 330°C, $T^{eo}_1 = T^{wo}_1 = 330°C$, the temperature

$$T_{0*}^{eo} = \frac{330°C + 0,1 \cdot 20°C}{1,1} \approx 300°C.$$

That is too much! The pizza will just turn into coal! This interface temperature is even much higher than in Naples pizzerias, where oven temperatures between 400 – 500°C are common.

Well, let us formulate the problem differently. Let us assume that generations of pizza makers, who were using wooden spade to transfer pizzas into the oven, are right: the temperature at the (Roman) pizzas bottom should be about 210°C. What would be the necessary temperature for an electric oven with steel surface?

The answer follows from our equation with the coefficient $\nu_{d,s} = 0,1$ and solved with respect to T^{eo}_1 when the temperature at the bottom of the pizza is the same as in the wood oven: $T^{eo}_0 = T^{wo}_0$. The result of this exercise shows that the electric oven should be much colder than the brick one: $T^{eo}_1 = 230°C$.

Of course, as always in physics, in order to understand the essence of the phenomenon, we considered the simplified model. It is not always possible to install the correct wood-burning stove, and not all visitors will understand the difference between very good and good pizza. Therefore, engineers use tricks: in a modern professional electric oven, they make a bottom made of special ceramics, imitating the bottom of a wood-burning oven, in order to fry pizza evenly, they make the bottom rotate, thermal ventilation is added to the radiation from the oven walls, and other tricks are also invented. But still, the dry heat and the smell of the wood of the old stove remain the ideal to which any imitation and improvement can only strive.

4. The physics of good coffee

One of the most common home-brewed coffee makers in Italy is the moka. It consists of three parts of the lower truncated cone (heater, combi steamers), where water is poured, a metal filter, where coffee of medium fine grinding is placed, and, finally, the upper truncated cone, where the finished drink is collected. This coffee maker is already designed to prepare a drink of a certain consistency: water should be poured up to the level of the valve in the heater, the filter is filled with a full one - about 6 grams per serving in 50 ml of water.

The process of making moka coffee is very entertaining. After filling of moka by water and coffee powder the along the upper and lower cones are tightly twisted, the upper strainer covers the filter cylinder. Additional insulation from the outside is provided by a rubber gasket between the top and bottom cones. The coffee maker is put on low heat. The preparation process consists in bringing water to a boil in a heater, then running it through the coffee powder contained in the filter, further lifting the drink prepared in this way through the tube and draining it into the volume of the upper cone. The coffee is then ready to be dispensed into cups through the spout.

Everything seems simple and clear, but what is the driving force behind the described process? Fire, of course. First, it heats the water to a boil, then the boiling process occurs in a confined space, where the water occupies much more space than the steam above its surface. The temperature surpasses through 100°C, the steam above the water surface remains saturated all the time, its pressure exceeds 1 atm and continues to grow. The external pressure, at the upper level of the filter, is equal to atmospheric. The saturated steam with a temperature above 100°C in volume V begins to play the role of a compressed spring, forcing slightly overheated boiling water through the coffee powder contained in the filter. In this process, those aromas, oils and other components are extracted from it that turn water into a wonderful drink. It is clear that the properties of this drink depend both on the coffee powder itself in the filter and on the temperature of the water and the time it flows through the filter. The secrets of mixing, roasting and grinding coffee beans are the secrets of every manufacturer, based on talent, work and centuries of experience. What determines the time of fluid flow through the filter, we can understand without industrial espionage, proceeding only from the laws of physics.

In the middle of the nineteenth century, the French engineers A. Darcy and J. Dupuis made the first experimental observations of the movement of water in pipes filled with sand. These studies laid the foundation for the creation of the theory of filtration, which today is successfully used to describe the movement of liquids, gases and their mixtures through solids containing interconnected pores or cracks. In addition to creating the first perfect water supply system in Europe in the city of Dijon, A. Darcy formulated the so-called linear filtration law, which today bears his name. It relates the volumetric flow rate of the liquid through the filter with its length and area, with the characteristics of the porous medium filling it, as well as with the pressure difference on both sides of the filter. Today it is one of the basic laws applied in the oil extraction. It is not difficult to apply Darcy's Law to the study of our moka. For example, you can calculate to what temperature boiling water overheats at the bottom of the coffee maker. This temperature can be estimated from the pressure difference between the bottom and top sides of the filter, which is included in the Darcy formula and looking at the dependence of the boiling point pressure on the external pressure. To do this, it is enough to note the time during which the drink fills the upper vessel and measure the characteristic dimensions of the mocha filter. A little more searches in the tables of water viscosity, bewilderment where to find the porosity coefficient of coffee powder in the filter, a brilliant idea to replace it with the corresponding value for medium-grained sand (which is in the table for oil engineers) and now, according to the above graph, it can be seen that the corresponding boiling point of water is 115-120°C, depending on whether the moka is on low or strong fire.

Moka coffee turns out to be strong and aromatic, without sediment, but it is still inferior in its taste to "espresso", which is served in a good bar. The main reason for this, apparently, is the relatively high temperature of boiling water forced through the filter by superheated steam. Therefore, the recipe for improving the quality of coffee when preparing it in mocha is as follows: put the coffee maker on very low heat. In this case, the filtration process will go slower, however, the steam in the lower vessel will not overheat too much.

Antique Neapolitan coffee maker ("Napoletana") resembles a moka, but uses gravity filtration instead of excess steam pressure. It also consists of two vessels stacked one on top of the other and a filter filled with coffee between them. The water in the lower cylinder is brought

to a boil, then the coffee maker is removed from the heat and turned over. Filtration occurs under the action of a column of water pressure of the order of several centimeters. The process of making coffee here is noticeably slower than in moka. It is possible to experiment with brewing the same amount of coffee in both coffee makers and, based on the inverse proportionality of the coffee brewing time to the applied pressure following from Darcy's law, check our previous estimate of the pressure in the mocha heater. However, in practice, for "Napoletana" coffee is chosen with a larger grind than for mocha, otherwise the drink will be ready in half an hour and already cold.

Experts say that coffee from "Napoletana" is tastier than from mocha: here, although the filtration process is slower, there is no harmful effect of overheated boiling water on coffee. However, the high pace of modern life leaves no time for philosophical conversation on the terrace overlooking Mount Vesuvius and the beautiful Gulf of Naples, in a pleasant anticipation of finally having a cup of wholesome drink. This luxury remained in old paintings from Neapolitan life and in the works of Eduardo di Filippo.

Not all Neapolitans were patient in times gone by. They say that in the last century one of those residents of the capital of the Kingdom of the Two Sicilies, could not calmly wait coffee brewed in "Napoletana", persuaded his friend, an engineer from Milan, to design a fundamentally new coffee machine that prepares an individual portion of a wonderful fragrant, thick drink for half a minute. Each good cup of coffee is a repository of the secrets of growing and harvesting coffee beans, making a mixture and roasting it, grinding ... Behind the pinnacle of coffee art - a small cup of Italian "espresso" - there is also high technology. The espresso machine is much larger in sizes

and more impressive than its counterparts described above. Usually such machines are in bars and restaurants, but for connoisseurs and lovers of coffee there are also home versions of this machine.

In an espresso coffee maker, water with a temperature of 90-94°C is forced under a pressure of 9-16 atm through a filter with coffee powder of a special grinding, even finer than for mocha.

The whole process takes 15-25 seconds, resulting in 1-2 servings of coffee, 20-35 ml each, for you personally and, perhaps, your interlocutor. The process of fluid flow through a filter with coffee powder is described by the same Darcy's law as in moka, however, the pressure difference applied to the filter is 50 times greater here, and the temper-

ature, on the contrary, is below 100°C. These parameters are specially selected so that do not destroy unstable fractions of the coffee drink with high temperature. The relatively short time of interaction of water with the powder, together with high pressure, leave excess in it and extract all the best from it: emulsions of coffee oils form the drink that cannot be achieved in any other way; its aroma is preserved by the presence of a foam that does not allow volatile components to disappear.

Espresso, oddly enough, contains less caffeine - due to the short contact of water with the powder (20-30 seconds *versus* 4-5 minutes in the filter) and the smallness of its volume, all the caffeine does not have time to be extracted. The first example of an espresso machine was exhibited in Paris in 1855. In modern stationary devices that make up the equipment of bars and restaurants, water is supplied under the required pressure using a special pump in the design. In a classic espresso machine, hot water from the heating cylinder, by lifting the handle, fills the chamber above the filter and then is forced through the filter manually by lowering the handle; high pressure is generated by the dynamic resistance of the coffee filter and the leverage effect that multiplies the hand force.

Acknowledgements

I want to express my gratitude to my friends and coauthors – Giuseppe Balestrino, Andreas Glatz, and Sergio Grasso — with whom I discussed and published some pieces of the materials included in this presentation.

REFERENCES

Attilio Rigamonti, Andrey Varlamov, "**Il Magico Caleidoscopio della Fisica**", La Goliardica Pavese, Pavia, 2007.

L.G.Aslamazov, A.A. Varlamov, "**The Wonders of Physics**", IV edition, World Scientific Publishing Company, Singapore, 2018.

Attilio Rigamonti, Andrey Varlamov, Jacques Villain, "**Le Kaleidoscope de la Physique**", Belin, Paris, 2014.

Andrey Varlamov, Andreas Glatz and Sergio Grasso, **The physics of baking good pizza**, Phys. Educ. **53**, 065011, 2018.

SEPARAZIONE DI CARICA IN MISCELE DI MATERIALI ORGANICI: VALE LA LEGGE DI COULOMB?

GUIDO RAOS (*)

Nota presentata dal s.c. Stefano Valdo Meille
(Adunanza del 4 aprile 2019)

SUNTO. – Dopo una breve introduzione ai materiali fotovoltaici organici, viene affrontato il problema della separazione di carica in questi sistemi. Nei dispositivi migliori, le eccitazioni prodotte dall'assorbimento di un fotone possono essere convertire in portatori di carica positivi e negativi (rispettivamente buche ed elettroni) con una efficienza vicina al 100%. Queste efficienze sono apparentemente in contrasto con la legge di Coulomb, che porterebbe le cariche a ricombinarsi a causa della loro reciproca attrazione elettrostatica. Un modello teorico sviluppato dal mio gruppo consente di superare questa apparente contraddizione, tendendo conto di effetti quantomeccanici quali la delocalizzazione dei portatori di carica.

***

ABSTRACT. – After a brief introduction to organic photovoltaic materials, I discuss the problem of charge separation in these systems. In the best devices, the excitations produced by absorption of a photon can be converted into positive and negative charge carriers (holes and electrons, respectively) with near-100% efficiency. Such efficiencies are apparently at odds with Coulomb's law, which would lead to change recombination due to their mutual electrostatic attraction. A theoretical model developed within my group overcomes this apparent contradiction, by taking into account quantum mechanical effects such as the delocalization of the charge carriers.

La transizione dai combustibili fossili alle energie rinnovabili è una necessità non più prorogabile a causa dei cambiamenti climatici in atto, e

(*) Dipartimento di Chimica, Materiali e Ingegneria Chimica "G. Natta", Politecnico di Milano, Italy. E-mail: guido.raos@polimi.it

l'energia solare sarà sicuramente uno dei principali ingredienti nel mix energetico futuro.[1] Attualmente il silicio rappresenta la tecnologia più consolidata per la conversione di energia solare in energia elettrica. Tuttavia, esso presenta alcuni svantaggi, tra cui spicca in particolare l'alto costo energetico (e quindi anche economico) della produzione del silicio elementare, il peso e la rigidità meccanica dei pannelli. Diverse classi di materiali innovativi sono oggetto di ricerca, per arrivare a sostituire in tutto o in parte il silicio. Tra queste, ci sono i semiconduttori organici. Essi hanno alcune caratteristiche interessanti, quali il basso costo, la possibilità di stamparli in film sottili su substrati flessibili, la facilità di integrazione in edifici (in finestre colorate, ad esempio).

Una cella solare organica richieda la co-presenza di almeno due materiali distinti (detti accettore e donatore), all'interno del suo strato attivo (il "cuore" della cella, tipicamente con spessore di 50-100 nm).[2] Per avere le necessarie proprietà semiconduttive, entrambi i materiali devono essere π-coniugati, con una alternanza di legami carbonio-carbonio singoli e doppi. Possono essere polimerici, derivati dei fullereni, o altro. Inoltre, i loro livelli energetici più esterni (HOMO per *highest occupied molecular orbital* e LUMO per *lowest unoccupied molecular orbital*) devono essere opportunamente "sfasati" in energia, in modo da facilitare la conversione di un eccitone (lo stato elettronico generato dall'assorbimento di un fotone) in una coppia elettrone-buca.[3] L'elettrone (con carica negativa, $-e$) va sul materiale accettore, mentre la buca (con carica positiva $+e$) va sul materiale donatore. L'attrazione Coulombiana tra queste cariche tenderebbe a mantenerle localizzate all'interfaccia tra i due materiali, che tipicamente vengono preparati con una morfologie nanosegregata detta a *bulk heterojunction*.[4] Una volta localizzate all'interfaccia, entro pochi nanosecondi dalla loro formazione, le due cariche tenderebbero a ricombinarsi e l'energia del fotone andrebbe dunque perduta. Ciò nonostante, è noto che le cariche possono diffondere lontano l'una dall'altra e generare una corrente elettrica (quando raggiungono gli elettrodi sui lati opposti dello strato attivo), con un'efficienza che può arrivare al 100%.[5]

La facilità con cui una coppia elettrone-buca può essere convertita in due cariche libere è apparentemente in contrasto con la legge di Coulomb:

$$U(R) = -\frac{e^2}{4\pi\epsilon_0\epsilon_r}\frac{1}{R} \tag{1}$$

che descrive l'energia di interazione tra due cariche $\pm e$ a distanza R, in un mezzo con costante dielettrica relativa ε_r (il segno negativo nell'equazione esprime il fatto che l'interazione è attrattiva). Infatti, ponendo le cariche a $R \simeq 1$ nm (valore rappresentativo della distanza tra una molecola accettrice e una donatrice, a contatto tra di loro), con $\varepsilon_r = 3.5$ (valore tipico per i materiali organici in gioco), otteniamo un'energia di legame $U_b \simeq 0.4$ eV. Questa energia deve essere confrontata con l'energia termica, che a temperature ambiente vale k_B T = 0.025 eV (k_B è la costante di Boltzmann e T"T = 298K la temperatura assoluta). Questa rappresenta l'energia "disponibile" per la dissociazione, in base alle leggi della termodinamica statistica. Com'è dunque possibile che la dissociazione avvenga, nonostante la grande disparità tra l'energia necessaria per la dissociazione (U_b) e quella effettivamente disponibile a temperatura ambiente?

Sono state avanzate varie ipotesi e meccanismi, per risolvere il paradosso della separazione di carica nei materiali fotovoltaici organici:[6,7]

- nanostrutturazione delle miscele donatore/accettore
- disordine energetico
- campi elettrici interfacciali
- formazione di eccitoni "caldi", ad alta energia
- delocalizzazione degli elettroni e buche
- alte mobilità delle cariche
- effetti di coerenza quantistica.

Senza entrare nel dettaglio di questi meccanismi, possiamo discutere brevemente l'effetto della delocalizzazione. Questa è possibile grazie alla natura π-coniugata dei materiali, ma è anche ostacolata dalla presenza di disordine (localizzazione di Anderson).[8] L'interazione Coulombiana tra due cariche diffuse (rappresentate per semplicità da due distribuzione Gaussiane $\rho(\mathbf{r})$, di ampiezza σ e a distanza $R = \|\mathbf{R}\|$) può essere modellata come segue:

$$U'(R) = -\frac{e^2}{4\pi\epsilon_0\epsilon_r}\int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2-\mathbf{R})}{R} = -\frac{e^2}{4\pi\epsilon_0\epsilon_r}\frac{(R/\sigma)}{R}\,. \qquad (2)$$

La *Fig. 1* illustra come, quando l'ampiezza o delocalizzazione delle cariche diventa confrontabile con la loro separazione ($R \leq \sigma$), l'interazione Coulombiana risulta molto attenuata rispetto a quella tra cariche puntuali.

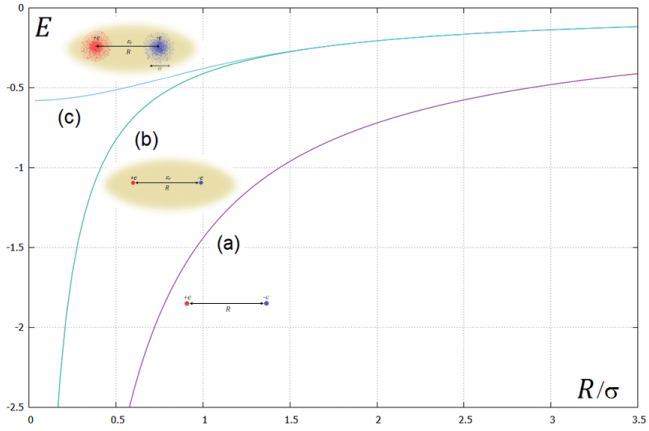


Fig. 1. Confronto dell'interazione Coulombiana tra due cariche puntuali nel vuoto (a), due cariche puntuali in un mezzo dielettrico (b), e due cariche delocalizzate in un mezzo dielettrico (c).

I moderni strumenti di calcolo consentono di simulare in modo dettagliato le interazioni all'interfaccia tra materiali donatori e accettori e, a partire da questa, calcolare le relative proprietà elettroniche.[3,9] Questi approcci (dinamica molecolare, chimica quantistica) richiedono ingenti risorse di calcolo e, proprio perché partono da modelli dettagliati, non sempre forniscono risposte e indicazioni generali, utili ad aumentare l'efficienza dei processi di separazione di carica. Per questo è utile studiare il problema anche con modelli semplificati, contenenti un numero limitato di parametri ed applicabili alla mesoscala (5-30 nm, approssimativamente).[10]

Negli anni passati, il mio gruppo di ricerca ha sviluppato e applicato un modello mesoscopico delle interfacce donatore-accettore.[11,12] Il modello riduce il problema all'essenziale, trattando ciascuna molecola accettrice o donatrice come un "sito" con due orbitali (HOMO e LUMO) e due elettroni. Le energie dei livelli energetici delle molecole sono parametrizzabili, in modo da riprodurre dati sperimentali o calcoli *ab initio*. L'interazione Coulombiana tra cariche, gli effetti di disordine e di delocalizzazione possono esere facilmente modulati, per verificare il loro effetto sugli stati eccitati del sistema. Il calcolo di questi stati viene fatto a livello quantomeccanico, senza approssimazioni di tipo classico. Il nostro risultato principale è che, con una scelta molto ragionevole dei parametri del modello, anche i più bassi stati eccitati di que-

sti sistemi (eccitoni "freddi", dunque) sono descrivibili come coppie elettrone-buca delocalizzate, su lati opposti dell'interfaccia. L'energia necessaria per separare le cariche e portarle a distanza infinita è effettivamente dell'ordine dell'energia termica, $k_B T$. Oltre che dalla delocalizzazione, la dissociazione è favorita anche dalla formazione di un dipolo interfacciale.

Possiamo dunque rispondere alla domanda posta a titolo di questo contributo. La legge di Coulomb vale anche nei materiali organici e non è in contraddizione con l'efficienza con cui avviene la separazione di carica. Tuttavia, essa deve essere integrata con una descrizione quantomeccanica che ammette la delocalizzazione delle cariche, o la loro localizzazione nei sistemi caratterizzati da un forte disordine (nell'energia o nella sovrapposizione degli orbitali di siti diversi). Piu in generale, il nostro modello dimostra ancora una volta l'utilità di descrizioni che bilanciano la semplicità con il mantenimento di alcuni effetti fisici essenziali, nella descrizione di sistemi e fenomeni complessi.

REFERENCES

(1) Armaroli, N.; Balzani, V. Solar Electricity and Solar Fuels: Status and Perspectives in the Context of the Energy Transition. *Chemistry - A European Journal* **2016**, *22*, 32–57.

(2) Günes, S.; Neugebauer, H.; Sariciftci, N. S. Conjugated Polymer-Based Organic Solar Cells. *Chemical Reviews* **2007**, *107*, 1324–1338.

(3) Nelson, J. Polymer:fullerene bulk heterojunction solar cells. *Materials Today* **2011**, *14*, 462–470.

(4) Lu, L.; Zheng, T.; Wu, Q.; Schneider, A. M.; Zhao, D.; Yu, L. Recent Advances in Bulk Heterojunction Polymer Solar Cells. *Chemical Reviews* **2015**, *115*, 12666–12731.

(5) Park, S. H.; Roy, A.; Beaupré, S.; Cho, S.; Coates, N.; Moon, J. S.; Moses, D.; Leclerc, M.; Lee, K.; Heeger, A. J. Bulk heterojunction solar cells with internal quantum efficiency approaching 100%. *Nature Photonics* **2009**, *3*, 297–302.

(6) Bässler, H.; Köhler, A. "Hot or cold": How do charge transfer states at the donor–acceptor interface of an organic solar cell dissociate? *Physical Chemistry Chemical Physics:* **2015**, *17*, 28451–28462.

(7) Few, S.; Frost, J. M.; Nelson, J. Models of charge pair generation in organic solar cells. *Physical Chemistry Chemical Physics* **2015**, *17*, 2311–2325.

(8) Lagendijk, A.; van Tiggelen, B.; Wiersma, D. S. Fifty years of Anderson localization. *Physics Today* **2009**, *62*, 24–29.

(9) Beljonne, D.; Cornil, J.; Muccioli, L.; Zannoni, C.; Bredas, J.-L.; Castet, F. Electronic Processes at Organic-Organic Interfaces: Insight from Modeling and

Implications for Opto-electronic Devices. *Chemistry of Materials* **2011**, *23*, 591–609.
(10) Savoie, B. M.; Jackson, N. E.; Chen, L. X.; Marks, T. J.; Ratner, M. A. Mesoscopic features of charge generation in organic semiconductors. *Accounts of Chemical Research* **2014**, *47*, 3385–3394.
(11) Raos, G.; Casalegno, M.; Idé, J. An Effective Two-Orbital Quantum Chemical Model for Organic Photovoltaic Materials. *Journal of Chemical Theory and Computation* **2014**, *10*, 364372.
(12) Casalegno, M.; Pastore, R.; Idé, J.; Po, R.; Raos, G. Origin of Charge Separation at Organic Photovoltaic Heterojunctions: A Mesoscale Quantum Mechanical View. *Journal of Physical Chemistry C* **2017**, *121*, 16693–16701.

Istituto Lombardo (Rend. Scienze) 153, 89-114 (2019)

THE PHYSICS OF LEONARDO DA VINCI

Nota del m.e. FERDINANDO BORSA (*)

(Adunanza del 2 maggio 2019)

SUNTO. – Tra le numerosissime attività di Leonardo da Vinci ci sono molti esempi in cui il grande artista si confronta con fenomeni fisici di cui cerca di trovare una spiegazione. Tuttavia lo studio delle leggi della fisica da parte di Leonardo non sembra motivato dal'interesse verso la fisica in se ma in quanto la spiegazione di fenomeni fisici gli risulta importante nella sua attività di artista, ingegnere e osservatore della natura. E' sorprendente che il grande genio abbia enunciato in modo magistrale il paradigma della scienza moderna di Galileo e Newton. In questa nota, dopo aver accennato ad alcune grandi intuizioni di Leonardo fisico si cerca di verificare in quale misura Leonardo abbia applicato la metodologia scientifica cosi ben annunciata allo studio di leggi della fisica. Tra i vari argomenti trattati nella sterminata produzione Leonardesca si è scelto di analizzare gli studi sull'attrito e sulla idrodinamica argomenti sui quali Leonardo si è impegnato con maggior continuità e sistematicità.

***

ABSTRACT. – Among the many extraordinary activities of Leonardo da Vinci there are a few examples when he confronts himself with physical phenomena that he tries to explain. However, the studies of physical laws by Leonardo does not seem to be motivated by the interest in Physics itself but rather by the explanation of phenomena entering in his activity of artist, engineer and observer of nature. Nevertheless it is surprising that the renaissance genius is able to enunciate the paradigm of modern experimental science as later established by Galileo and Newton. In the present note, after having reviewed some of the most spectacular intuitions by Leonardo in the field of physics, we try to verify to which extent he applies the scientific methodology, he so well enunciates, to the study of physical laws. Among the very extensive production of Leonardo we chose to analyze in detail the studies on friction and on hydrodynamics since these are the subjects that he pursued with more continuity and method.

(*) Istituto Lombardo Accademia di Scienze e Lettere, Milano. Dipartimento di Fisica, Università degli Studi di Pavia, Italy. E-mail: ferdinando.borsa@unipv.it

Introduction

The bibliography of Leonardo da Vinci is huge. Suffice to say that Gerolamo Calvi charged in 1905 by the Istituto Lombardo di Scienze e Lettere to list all references to works on Leonardo reported 1500 items. At least as many were added in subsequent times. Most of the studies and biographies deal with the artistic masterpieces of Leonardo but there is also a large number which deal with Leonardo as a scientist. In this second case most of the interest seems to be focused of the genius of Leonardo in describing nature through careful observation and description by the use of drawings of exceptional quality. This is particularly evident in the works of Leonardo in the field of engineering and anatomy. In the first case he does not really invent any new machine but, for the first time, he uses technical drawing to show how the different parts operate and he often introduces modifications and improvements. Many museums around the world display models of Leonardo's machines which were constructed following the original drawings, some of them work some of them do not. None of them was probably constructed at the time of Leonardo. Furthermore the anatomical drawings are of incredible accuracy and if he had published them with the collaboration of Antonio Della Torre, then professor of anatomy at the University of Pavia, he would have anticipated the atlas of Vesalio by some fifty years.

On these premises most of the biographers and scholars of Leonardo conclude that he was a great scientist besides being one of the greatest artist ever lived. But what about Physics? Physics is not a descriptive science nor is engineering. Physics is an experimental science based on a paradigm which requires experiments from which to arrive by inductive reasoning to the formulation of theories in mathematical language, theories which can in turn be used to explain further experiments and natural phenomena with deductive reasoning. Leonardo did not seem to be interested in developing physical theories by themselves but he was forced to deal with physical laws whenever he pursued his interests in art, engineering and observation of nature. He studied optics since light is of fundamental importance in painting, he studied fluids as part of his work as hydraulic engineer, he studied mechanics to learn how machines operate, he studied acoustics as a result of his interest in musical instruments and so on. It is amazing to discover that Leonardo had a very clear idea of what is the paradigm of modern science. Let us see what he has to say:

> *"Nessuna umana investigazione si può dimandare vera scienza, s'essa non passa per le matematiche dimostrazioni…."* ("no human investigation can be considered true science, if it does not go through mathematical demonstrations…")
>
> *"Ma prima farò alcuna esperienza, …, perchè mia intenzione è allegare prima l'esperienza, e poi colla ragione dimostrare, perchè tale esperienza è costretta in tal modo ad operare".* ("But first I will perform some experiments …, because my purpose is to asses first the experience (experiment?) and then with reasoning demonstrate why the experiment is coming out that way")

On the basis of the above premises it is natural to ask oneself if Leonardo did indeed use the modern science method to investigate physical phenomena. It is very difficult to address this question because the interest of Leonardo in specific physical phenomena is scattered over space and time. It is scattered in space because Leonardo writings have been largely dismembered and are found now in different museums and institutions around the world organized in the form of Codes and Manuscripts. It is scattered in time because, although the chronology of the "folios" has been reconstructed thanks to the wonderful work of Carlo Pedretti, Leonardo used to work on a given problem only for a short time, then jump to a different subject, may be to an artistic drawing or a technical one or simply to the note of his expenses all superimposed on the same "folio" then go back to the initial subject days or months later on the same or on different "folios". This difficulty in concentrating on a subject, particularly when it had to do with a physics law which requires concentration, method and persistence is part of a recent investigation by Carlo Catani and Paolo Mazzarello (Brain 2019,142,1842-1846) which is based mostly on Leonardo's problem to complete his art projects. The authors arrive to a striking conclusion which will most likely ignite a controversy in years to come.

There are only a few physicists who looked at Leonardo's contribution in the field of Physics. Among them I like to quote here in chronological order: Gilberto Govi (Leonardo letterato e scienziato, Saggio delle opera di Leonardo da Vinci, Ricordi, Milano 1872 pp. 5-22); Leopold Infeld (Leonardo Da Vinci and the fundamental laws of science, Science and Society, Vol 17, N_0.1, Winter 1953 pp. 26-41); Domenico Argentieri and Sebastiano Timpanaro (in Leonardo, edited by Ralph Steadman, Jonathan Cape Ltd, London 1983); Fritjof Capra

(The science of Leonardo: Inside the mind of the Great genius of the Renaissance, Anchor Books, New York 2007); Luigi Borzacchini, La solitudine di Leonardo. Il genio universale e le origini della scienza moderna. Edizioni Dedalo (2019).

All the above cited works are very captivating and well documented but they are generic on the subject of physics and all seem to miss the point which interests me the most here: did Leonardo apply the paradigm of modern science to his investigations of a particular physical phenomenon to derive a law formulated in mathematical terms ? Only recently this issue has been examined in detail by a Cambridge professor of engineering, Ian M. Hutchings (Leonardo Da Vinci's studies of friction, Wear 360-361 (2016) pp. 51-66). In his paper Hutchings traces all the studies of Leonardo over a span of 20 years dealing with friction and its applications. He demonstrates that Leonardo did extensive experimentation on the subject and that he arrived to the formulation of the correct laws of friction two hundred years before Guillaume Amontons who is credited with the formulation of the laws.

In the present note I would like to discuss first the work of Hutchings to see if Leonardo did indeed perform the experiments that he drew and to which extent he arrived to the mathematical formulation of the laws of friction which actually imply only the understanding of the concept of proportionality of two physical quantities. By using Hutchings' methodology I will try then to analyze the work of Leonardo in the field of hydrodynamics in order to show to which extent he came close to formulate Torricelli 's Theorem, the definition of fluid flow and the principle of continuity. Only in discussing in detail these two specific cases one can get a clear idea of the extent to which Leonardo was able to apply the principles of modern scientific investigation which he had indeed formulated correctly. But first let us see some of the "discoveries" in physics reported by most of the biographers and scholars of Leonardo.

Some intuitions and anticipations of Physics laws by Leonardo

The fame and recognition of Leonardo as a physicist is attributed by the many biographers and scholars of the renaissance genius to the many intuitions he had about the explanation of physical phenomena and laws. It is worth mentioning a few of them to see how he came about to his

almost miraculous anticipations. We could start with the most far-fetched which is mentioned by Infeld in his article. When Leonardo says "*prendi un oggetto riscaldato al color rosso e pesalo nuovamente quando è freddo*" (take an object heated to the red color and weigh it again when is cold) does he anticipate the ponderability of matter expressed by the famous relation between energy and mass, $E = mc^2$, result of Einstein's theory of special relativity? Very unlikely! Nevertheless the scientific curiosity of Leonardo and his fine intellect are shown in the mere fact that he poses the right questions. Let us now see some more concrete cases:

a) The explanation of the blue color of the sky

In the Leicester code Leonardo gives an explanation of the blue color of the sky which is remarkably close to the correct explanation given by Lord Rayleigh in 1871 based of the detailed theory of the scattering of solar light by the molecules in the atmosphere.

"*Dico l'azzurro in che si mostra l'aria non essere suo proprio colore, ma è causato da umidità calda, vaporata in minutissimi e insensibili attimi, la quale piglia dopo se la percussione dei razzi solari e……..*" (I say that the blue which the air shows is not its own color but is caused by the hot humidity, vaporated in tiny and insensitive fractions(?) which receives on itself the percussion of the solar rays and……)

In other folio's Leonardo describes experiments where he shines light on smoke rising in front of a dark background and he obtains so a bluish color of the smoke. Clearly he arrives to the intuition of the explanation of the blue color of the sky by applying analogical reasoning as he does very often in his scientific observations

b) Vibrations, waves and the nature of sound and light

In the study of wave motion Leonardo shows at his best the quality of subtle observer of nature and of scientific reasoning. He observes the circular waves generated by a stone cast in a pool of water and concludes in the Manuscript A *"….e la ragione di questo è che l'acqua, malgrado sembri muoversi, in realtà non lascia la posizione originale, perchè il vuoto creato dal sasso si richiude prontamente. Pertanto il moto prodotto dall'aprirsi e chiudersi veloce dell'acqua ha come conseguenza solo un urto che può essere descritto come un tremore piuttosto che un movimento"* ("….and the reason of this is that the water, although it seems to

move, in reality does not abandon the original position, because the vacuum created by the stone is closed immediately. Therefore the motion generated by the fast opening and closing of the water has the only consequence of an impact which can be better described as a tremor rather than a movement").

This statement may appear trivial to us today but it took 200 years to arrive to a rigorous formulation of transverse waves by Huygens (a colleague of Newton) who used in his *Traitè de la lumiere (1690)* words that seem to have been copied from Leonardo statement. Leonardo went even further. By using analogical reasoning, so dear to him, he states that light and sound must also be transverse "tremors" although the nature of either was totally unknown to him. This indicates that his mind was more driven by his artistic sense than by sound scientific reasoning.

c) Reflection, refraction and dispersion of light

In his treatise on painting, where Leonardo gathers most of his studies on light, one can find several examples of his understanding of the law of reflection always expressed by drawings of light rays impinging on a surface and bouncing back forming an angle equal to the incident angle but he never states clearly the law which implies that the two rays and the normal to the surface have to be in the same plane.

Regarding the phenomenon of refraction Leonardo observes correctly that light rays are bent as they cross the surface separating two transparent media but he never investigates the phenomenon quantitatively. Understandably, with the knowledge of mathematics at his time, he could not arrive at the Snell's law.

In the Leicester code Leonardo describes in words the decomposition of white light in different colors when crossing water anticipating the experiments of Newton with the prism:

> *"……..Ma se tu poni tal bicchiere pieno d'acqua sul piano della finestra in modo che dall'opposta parte lo ferischino li razzi solari, allora tu vedrai li predetti colori generarsi nella impressio fatta dalli razzi solari penetrati per esso bicchiere e terminati sopra il pavimento…".* ("…..But if you set such glass full of water on the plane of the window in such a way that from the opposite side it is hit by the sun's rays, you will see the given colors generated by the impression done by the solar rays penetrated in the glass and terminated on the pavement…"

As can be seen from the folio reproduced in *Fig. 1*, in this rare case he formulates the observation of light dispersion correctly in words while the drawing is grossly wrong! Normally is the other way around with Leonardo.

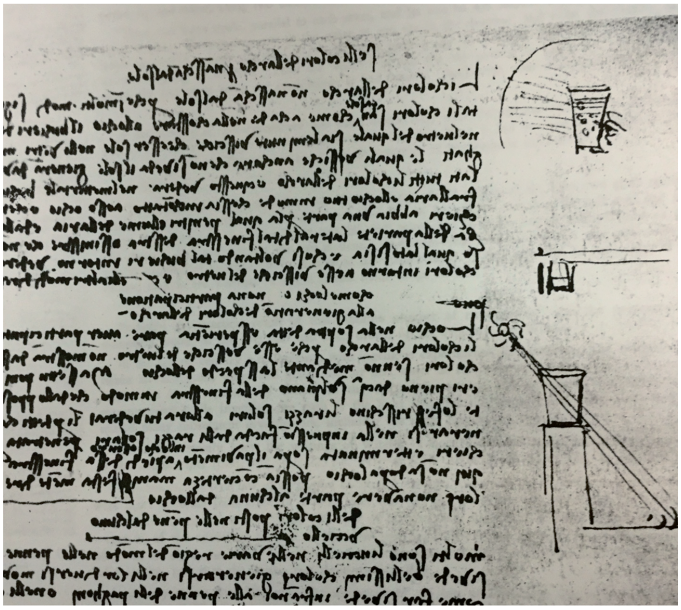


Fig. 1.

d) The three laws of mechanics (Newton's laws)

Classical mechanics is based on Newton's second law, **F =** m **a.** At the time of Leonardo calculus did not exist and thus it was impossible for Leonardo to acquire the concept of acceleration. Without acceleration it was impossible to define the force and the inertial mass. In fact Leonardo never mentions mass (he talks about weight but not as a force) and gives a definition of force which is purely metaphysical ! However, the first law or law of inertia, which is a particular case of the second law when **F = 0** could be discovered by Galileo who had the great intuition that to investigate the motion of bodies one has to eliminate friction from the picture. Some biographer and scholars of Leonardo claim that he anticipated Galileo by formulating the law of inertia in these terms:

> *"Ogni moto naturale e continuo desidera conservare suo corso per la linia del suo principio"* ("Every natural and continuous motion desires to maintain its course along the line of its beginning"
> *"Ogni moto attende al suo mantenimento, ovvero: ogni corpo mosso sempre si move, mentre che la impressione della potenzia del suo motore, in lui si riserva"* ("Every motion aims at his continuation, namely: every body set in motion keeps moving, as long as the impulse of the power of its motor, remains in it")

It is clear to me that these statements do not adumbrate the law of inertia. In the first one there is no mention of the need for having no force on the object while the second is a formulation of the theory of the impetus of Buridan in line with Aristotelian dynamics.

The third law is another matter. The quasi-static version namely to every action there is a reaction equal and opposite does not require any understanding of motion and thus it is no surprise that Leonardo did arrive to the correct formulation of the third law:

In the folio 135 of the Arundel code we read:

> *"Ogni corpo si muove in direzione opposta dal punto da cui è spinto dall'oggetto che lo colpisce….Il corpo colpisce l'oggetto nella stessa misura in cui l'oggetto colpisce il corpo"* ("Every object moves in the opposite direction from the point from which is pushed by the object which strikes it….The body strikes the object in the same way as the object strikes the body"

The more complete formulation namely the conservation of momentum for an isolated system was out of reach for Leonardo. However, it is surprising that he understood and described many cases of collisions both elastic and inelastic in which the conservation of momentum is implicit.

In the Arundel code 263,83v he reports this remarkable statement which is clearly the result of reasoning on experiments performed

> "*..ci sono due tipi di percussione, quando l'oggetto si allontana dal corpo in movimento che lo ha colpito e quando un corpo rimbalza contro un oggetto che lo ha colpito…..Quando i due corpi hanno lo stesso volume e peso il corpo che viene colpito lascia il corpo che lo colpisce al suo posto e prosegue con il moto del primo corpo"* ("…there are two types of collision, when the object goes away from the moving body that has hit it and when a body bounces back from an object that has hit it…When

> the two bodies have the same volume and same weight the body which is hit leaves the body which hit it at his place and continues with the motion of the first body")

All different cases of collisions are summarized in manuscript A folio 8r with beautiful and accurate drawings of the trajectories of the colliding bodies without, however, any description of the experiments which led to those geometrical conclusions.

e) The conservation of energy and the perpetual motion

Leonardo could not have the concept of work and of energy. He struggles to build a machine which can operate continuously without an energy input. After many attempts he abandons the idea and concludes that it is impossible and he compares whoever tries to obtain perpetual motion to the alchemist:

> *"O speculatori dello continuo moto, quanti vani disegni, in simile cerca, avete creato! Accompagnatevi colli cercatori dell'oro"* ("O speculators of continuous motion, how many useless designs, in such a search, you have created! Join the ones who search for gold"

In effect one has to make a distinction between continuous motion and perpetual motion. In the first case the motion can continue indefinitely in a force field which is conservative. The motion of celestial bodies is an example. Thus in principle continuous motion is possible if one can eliminate frictional forces completely. A machine moving continuously in a conservative force field would not generate any work but only measure time. With the discovery of phenomena such as superconductivity and superfluidity a continuous pendulum is in principle feasible. Perpetual motion, on the other hand is impossible even in principle because it violates the principle of conservation of energy. It is not clear if Leonardo was after a machine that would measure time by running continuously without friction or a machine that would produce work without an energy input *i.e.* perpetual motion. In the first case his statement would be wrong while in the second case it would be correct.

Leonardo expressed, however, the intuition of the conservation of energy when he describes the cycle of a living creature. He says:

> *"Il corpo di qualunque cosa, la qual si nutrica, al continuo muore e al continuo rinasce, perché entrare non può nutrimento, se non in quelli lochi, dove il passato nutrimento è spirato……Ma se tu ne rendi tanto quanto se ne distrugge alla giornata, allora tanto rinasce di vita, quanto se ne consuma……"* "The body of whatever thing, that nourishes itself, continuously dies and continuously is born again, because no nourishment can enter, if not in those places where the previous nourishment is finished….But if you give back the same amount that was destroyed, during the day, then so much is born again of life as it is consumed…"

The above statement sounds pretty much as the application of the first principle of thermodynamics to a living creature!

I mentioned some of the most relevant examples of intuitions of Leonardo about physical laws. These have been reported in most of the biographies of the artist and in specific studies about Leonardo as a scientist. As I pointed out along the way there is no systematic study of physics in the above instances, no quantitative experiments, no mathematical formulation of any law. On the basis of those "discoveries" one could claim that Leonardo was the greatest scientific genius and somebody else could claim that he was only a great artist with acute sense of observation and great fantasy (like *e.g.* the French writers Jules Verne and Albert Robida who anticipated in their novels many discoveries of the following century)! This is why it is necessary to focus on a specific physics law that Leonardo studied over the years with some consistency to understand to which extent he applied the paradigm of modern science that he so clearly stated. We thus report in the following the studies of Leonardo on friction and on fluid dynamics.

The laws of friction

I. Hutchings, in his important paper (Wear 360-361 (2016) 51-66), starts by observing that some of the first experiments on the study of friction are drawn on a small notebook- size piece of paper (which I call jokingly "pizzino" *i.e.* the name of the pieces of paper that the mafia bosses used to communicate with their accomplices from the jail):

Hutchings points out,with maybe excessive animosity, that the art historians who examined the "pizzino" (*Fig. 2*) were only interested in

the sketch, on the up right hand corner, of the old woman (Elena of Troy in old age?) and the Petrarca saying "cosa bella mortal passa e non dura" (a mortal beautiful thing does not last).

Starting from this "pizzino" Hutchings goes back and reconstructs all investigations of Leonardo regarding the laws of friction and their applications. These span over a period of about 20 years and they present conclusions often incorrect that were changed later.

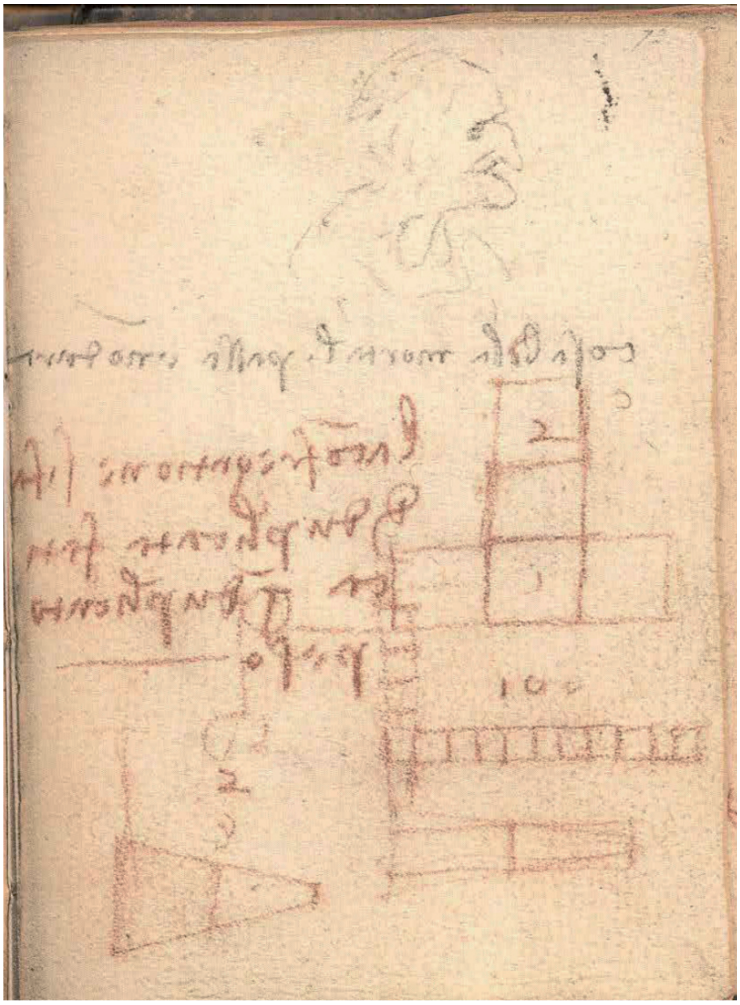


Fig. 2.

The two drawings reproduced (*Figg. 3-4*) (two among the many regarding friction found in Leonardo's folio's) represent possible experiments designed at establishing the amount of the frictional force and its dependence upon the weight of the object experiencing friction and the dependence of the frictional force from the surface area of contact of the object on the plane. In the first drawing the many objects are connected to a sort of handle. This seems to indicate that Leonardo tried to measure the frictional force by the effort necessary to set the object in motion. On the folio there is no mention of quantitative results of the experiments. One finds only this comment

> *"La confragazione si fa in duplicata fatica in duplicato peso"* ("Friction is done in twice the effort for twice the weight")

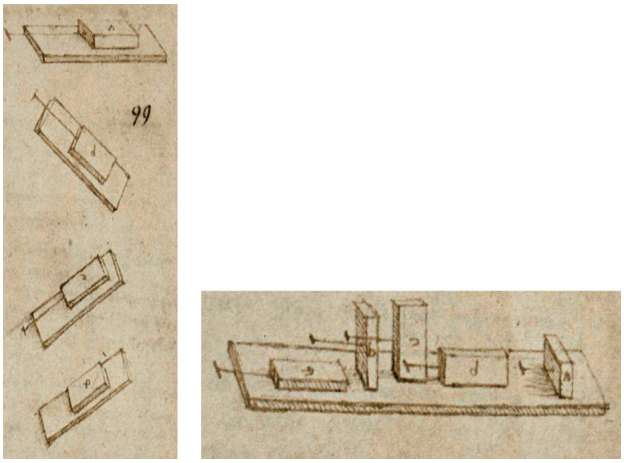


Fig. 3.

The idea is there but there is no attempt to formulate the law in mathematical terms. He does not identify the friction and the muscular force as two equal and opposite forces nor does he distinguish static from dynamic friction. In the second drawing Leonardo sets up a more quantitative experiment. He decides to compare the frictional resistance to the weight acting in opposite direction. But again there are no notes about the results of the experiment. He seems to be content to find out that the same weight displaces the object independently of the area of contact on the plane. Again he is measuring the static friction not the dynamic one.

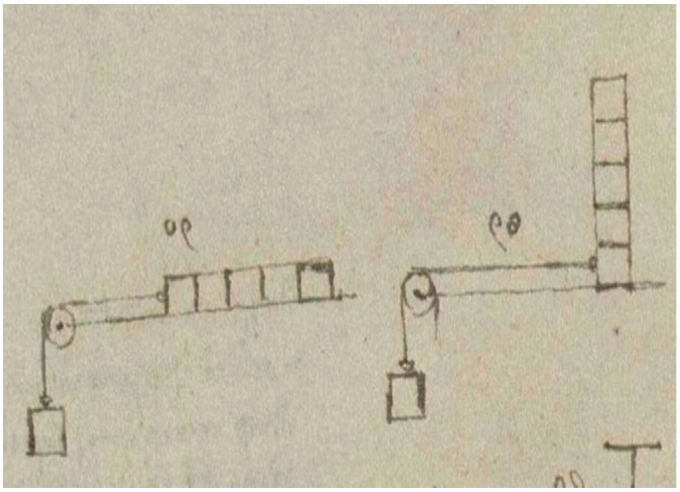


Fig. 4.

It is a matter of great delusion to me to see that Leonardo, who designed extraordinary machines to perform all sort of tasks, did not even try to design an experiment to measure the weight necessary to move an object at constant velocity on an horizontal plane. In that way he would have measured correctly the dynamical friction and as an added bonus he would have perhaps also understood the principle of inertia! In any case he arrives to the correct formulation of the laws of friction which he summarizes in the folio 187 of the atlantic code:

1. *Quelli corpi che son di più pulita superficie hanno più facile confregazione.*
2. *Infra li corpi di ugual lubricità il piùgrave ha più potente resistenzia nella sua confregazione.*
3. *'N un corpo d'ugual lubricità e di lati inequali la confregazione fatta da qualunche d'essi lati non muterà potenzia nel resistere al suo motore.*

In the first he states that friction depends from the conditions of the two surfaces in contact. In the second he states that the frictional resistance depends on the weight of the body. In the third he states that the frictional resistance for the same weight is independent from the area of contact. The discovery of the laws of friction 200 years before Amonton stated the very same laws is a remarkable achievement. The

fact that at that time Luca Pacioli, who taught algebra to Leonardo, did not use yet symbolic representation of known and unknown quantities prevented Leonardo from stating the second law as $\mathbf{F}_A = \mu \mathbf{N}$. However, as described in details by Hutching he uses in many applications values of the coefficient of friction μ which he must have determined from the ratio of frictional force $\mathbf{F}_A$ to weight **N** although it is not clear if the coefficient refers to static or dynamic friction. It is true that Leonardo never records and describes quantitatively the results of his measurements but the judgment of C.Truesdell (The mechanics of Leonardo da Vinci, Springer, Berlin, 1968) who says that Leonardo never did any experiment in his life appears too harsh!

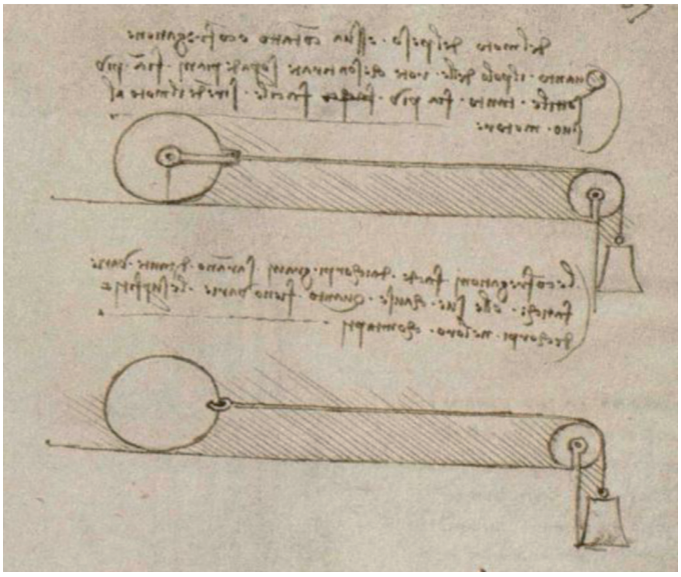


Fig. 5.

Leonardo is interested in applying the laws of friction to the study of the effort necessary to draw a cart on wheels such as the carriages which carry cannons and other materials on the battlefield. Thus first he tackles the problem of the difference between revolving friction and dragging friction. The drawing in *Fig. 5*, found in the Madrid code I65rc.1493–1497, is clearly designed for this purpose. Leonardo understand that the revolving friction is much smaller than the dragging one. He gives a very elaborate and obscure explanation of the revolving friction but he accom-

panies it with a drawing which is self explanatory. In fact he shows that to move a wheel, even heavily loaded, it is sufficient a very small force enough to move the wheel from one equilibrium position to the next. As long as the wheel rolls the friction is static and much larger than the dragging friction. A great anticipation of the ABS (Assisted Braking System)!

In the *Fig. 6* (drawing on the right) he indicates that a wheel loaded by a huge weight (900000) can be displaced by a tiny one (1/900000). The numbers are purely indicative and in arbitrary units! Once more he is more at ease with drawing and geometry than with words and mathematics.

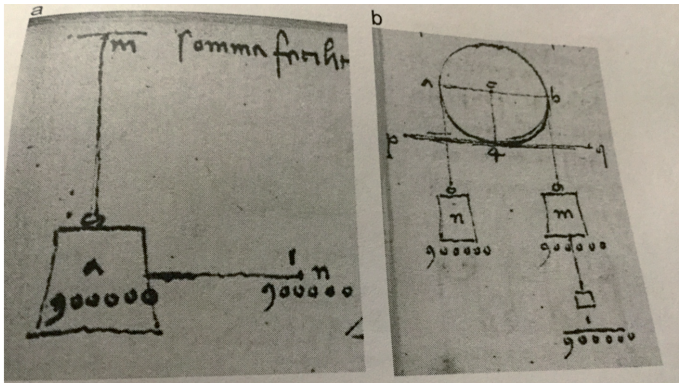


Fig. 6.

At this point Leonardo tries to understand where is the resistance to the movement of a wheel coming from. After several wrong attempts he understands correctly that the friction force acts on the axle of the wheel and that the crucial parameter is the circumference of the axle. The drawing shown in *Fig. 7* (from the Atlantic Code 561r,c.1487–90) demonstrates that he planned experiments to understand the effect. There is no report of quantitative measurements but somehow he states that the friction is less if the circumference of the axle is smaller and he even arrives to the correct conclusion that the force necessary to draw a kart is the frictional force scaled by the ratio of the circumferences of the axle to the one of the wheel. By using the definition of work, the result follows immediately by equating, approximately, the active work to the passive one *i.e.* $F.\ C_W = F_A\ C_A$ from which $F = F_A.\ C_A/C_W = F_A\ d_A/\ d_W$, where C and d are the circumference and the diameter of the wheel and the subscripts $_W$ and $_A$ refer to

the wheel and axle respectively. How Leonardo could arrive to the correct result without the concept of work, without symbolic algebra and even without any systematic and quantitative measurements is one more example of his extraordinary intuition which makes up for his mathematical shortcomings and his lack of quantitative experimentation.

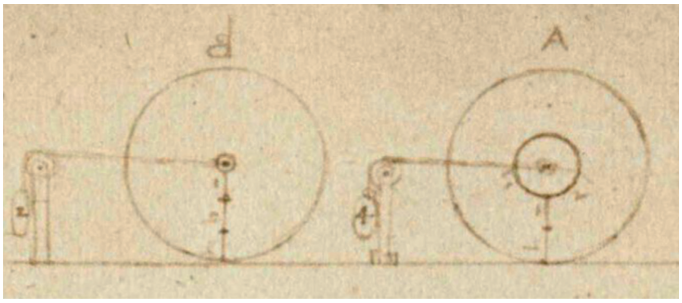


Fig. 7.

In the Madrid code (46r, c. 1503–5), he sets up a scheme to use his discoveries to find out the force necessary to move a given kart for any given load. This is a summary of what he does:

> "He assumes a load of 96 pounds on a wheel having a circumference 24 times the one of the axle. He assumes a friction coefficient of 0.25. Thus the frictional force is 24 pounds and the force to move the wheel is 1 pound. Now he can calculate the force necessary to move that particular wheel for any load. If, for example, the load is 6000 pounds the force necessary to move the wheel is 62.5 pounds. To arrive to this number he uses what at the time was called the "rule of three" which in symbolic algebra is the solution of the proportion 96:1 =6000:x."

The choice of those particular parameters for the wheel and the load make the calculations very simple. One assumes that Leonardo used the method also for more complicate numbers but he does not report any such example demonstrating his uneasiness with mathematics even when only multiplications and divisions are involved.

In the paper by Hutchings there are several more examples of applications of the laws of friction by Leonardo but the ones illustrated here are sufficient to draw conclusions about Leonardo's methods in his physical research:

1. With his drawings he proposes experiments which either he does not execute or, if he does, he does not report quantitative results.

2. He arrives to the formulation of correct laws and conclusions with great intuition and imagination by using analogical more than deductive reasoning and with no use of mathematics.
3. He expresses his thinking better with drawing and geometry than with words and mathematics.

Fluid dynamics

Leonardo investigated the laws of fluid dynamics mostly in connection with his extensive work in the field of hydraulic engineering. He came to a remarkable understanding of the equation of continuity and of the flow capacity for a stationary motion in a rigid pipe. He derived these concepts from the idea of the conservation of mass and understood that the flow is the product of the velocity times the section of the pipes. However, he never expressed these concepts in mathematical terms but he simply stated, correctly, that where the pipe (or river) is narrower the velocity is greater.

In the Leicester Code 6v he says:

"*Ne fiumi di qualunque varieta di largeza e profondita a cha de per necessita che in ogni grado della sua lungheza con equal tempo passa equal quantita d' acqua».*
"In the rivers of whatever size and depth it is necessary that in any point of its path for the same time there is the same amount of water flowing"

In the Manoscript A, 57 r he says:

"Il fiume d equale profondita sara tanto piufugha nella minore largheza che nella maggiore quancto la maggiore largheza avanza la minore …"
"In a river of constant depth the flow will be faster in the regions of less width than in the wider ones by as much as the wider size is bigger than the smaller size…"

In the same folio he continues:

"…Esenplo. Verbigratia se fiaun loco che abbi 3 varie largheze ……"
where he describes in a poetic way the analogy of the flow of water in a river with the flow of a crowd of people in a street of variable width whereby he points out that where the street narrows people are forced to walk faster to keep up with the flow of the crowd.

Leonardo was intrigued by the problem of the velocity with which a liquid spurts from a hole or an opening in a container as a function of the depth of the hole from the free surface of the liquid. This is the problem solved by Torricelli which goes under the name of Torricelli Theorem. The interest of Leonardo for Torricelli Theorem is twofold. On one hand the problem is important when bottling the wine from a barrel and Leonardo did produce wine for a short time in his life when he was given by the duke Ludovico il Moro as a gift a piece of land close to the church of Santa Maria delle Grazie (the vigna di Leonardo has been recently restored in the same place and with the same kind of grapes and is open to the public to be visited). The other application is to the measurement of the amount of water drawn from a gate in the irrigation channels. In the following I will show how Leonardo tackled these problems.

Regarding the problem of bottling the wine he says in the Manuscript I 7r:

> "*…se una botte ha in se il vino alto 4 braccia e gitta il vino lontano da se 4 braccia, quando il vino sarà nel calare disceso all'altezza di 2 braccia della botte gitterà ella il vino per la medesima cannella ancora 4 braccia? Cioè se il calo e l'empito del gittare della cannella diminuisce con equale proporzione o no…"* "if a barrel contains wine up to 4 "braccia" (a unit of the order of a yard) and gushes the wine 4 "braccia" away, when the wine in the barrel is lowered to a level of 2 "braccia" will the wine gush still 4 "braccia" away ? In other words if the level of the wine and the speed of exit from the hole decrease in equal proportion or not…."

Clearly Leonardo does not have the answer so, as he does in many other instances, he poses the problem in the form of a question. Actually the range reached by the wine spurting from the hole depends from the initial velocity and the level of the hole from the ground. This is illustrated in the *Fig. 8* based on the current knowledge of Torricelli's Theorem and parabolic motion.

From what Leonardo says it appears that the barrel is set 1 "braccia" from the ground although he does not says it. In fact for H=1 and h=4 the range or throw G is also 4 "braccia". Of course Leonardo did not know the laws of parabolic motion of a mass in the gravitational field. However, he seems to understand that for a constant height from ground the range G is proportional to the initial horizontal velocity.

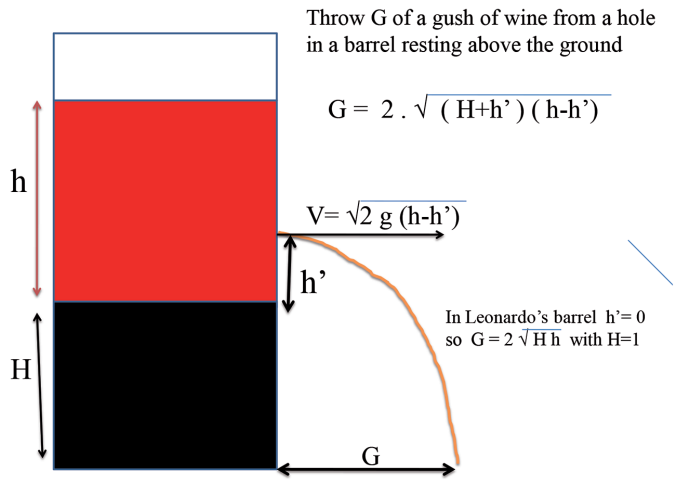


Fig. 8.

Thus he has a way to experiment the velocity (which he calls empito) as a function of the level of the wine above the hole. What clearly sets him off is that the velocity is not proportional to the level h but to the square root of h. I am convinced that he did try to answer to the question that he poses above but got confused by the fact that the throw G when the level decreases from 4 to 2 "braccia" does not decrease from 4 to 2 but from 4 to 2 √2. If he had known the square root (which he should have) and he had used symbolic algebra (which he could not at his time) he would have discovered Torricelli's law!!

His experimentation is described in the drawing from the Madrid Code 1, 124r (*Fig. 9*).

More experiments are described in the Atlantic Code 219r (*Fig. 10*).

The trajectory of the water spurting from holes at different levels are only qualitatively correct (*Fig. 10*). As usual Leonardo does not reports any quantitative and systematic measurements. In some cases the drawings are also incorrect. Thus we can conclude that Leonardo understood that the speed with which the water spurts from a hole in a barrel is higher when the level of the water above the hole is higher but he does not go any further and he does not seem to relate the effect to hydrostatic pressure (a concept familiar to him from Archimedes work).

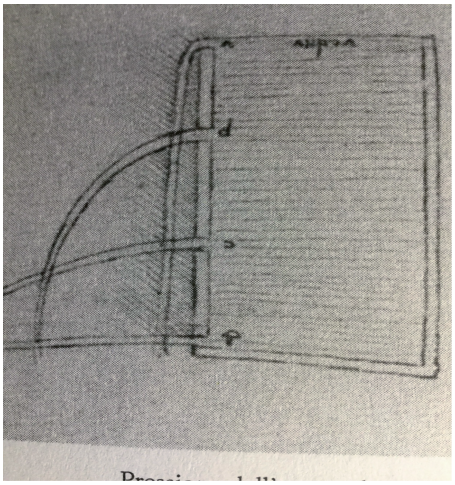


Fig. 9.

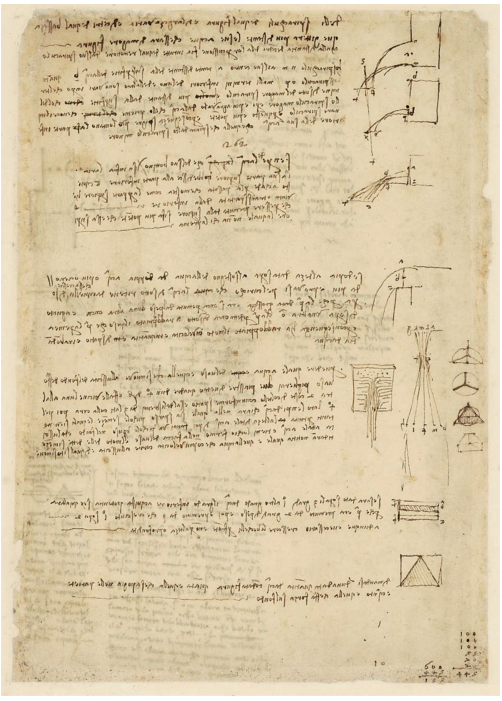


Fig. 10.

Leonardo did not miss to realize that the spout of water out of a container due to hydrostatic pressure afforded an easy way to investigate parabolic motion. He could not understand parabolic motion without the knowledge of Newtons's laws but he could study it and describe it in a phenomenological way. This problem was important to him to understand and design the throw of cannon balls. A device he used for the purpose is described in the drawing below (*Fig. 11*). It is a clever device made up of a pressurized rubber container where he applied spigots that he could orient in different ways to simulate all kind of trajectories.

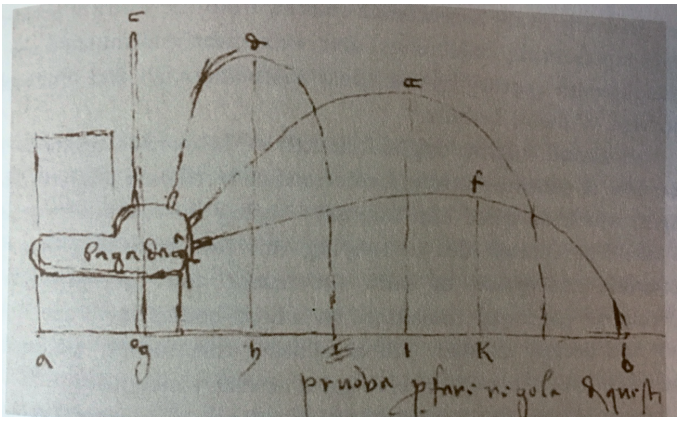
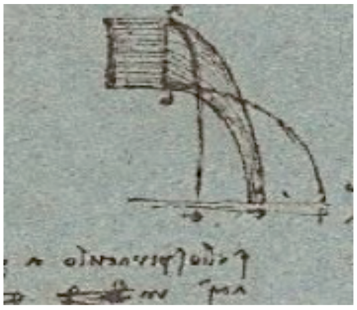


Fig. 11.

His insatiable curiosity pushed him to investigate even more complicate situations like the one described in the drawing below (*Fig. 12*).

A translation of Leonardo's words would not do justice to this remarkable text. What he describes is that when more jets of water with different horizontal velocity merge the range of the throw of the combined jet is intermediate between the ranges of the individual jets. This is the result of the composition of parabolic motion as the combination of horizontal and vertical motion, the conservation of momentum and the different flow from the different holes. He did not have any idea of these concepts (except for the flow) but he did arrive to a qualitative correct result. Did he arrive to this by experimenting or by reasoning? As always with Leonardo it is difficult to answer the question since he does not report the results of his experimentation granted that he performs it.



Codice Atlantico Folio 256 r,
Biblioteca Ambrosiana

Leonardo writes:»*Se lo spiraculo a gitta la sua acqua in e e lo spiraculo b la gitta in d seguira che se li due spiraculi saran contiunati per tutto lo spazio ab, che tutta l acqua di tale spiraculo si gittera in c. E questo accade perché la parte che prima andava dal b in d è poi aggravata dal peso ab che la piega in c e l acqua ae che prima cadeva in e, e poi sostenuta dal moto ab.»*

Fig. 12.

Related to hydrostatic pressure are also Pascal principle and the principle of communicating vessels which were known by the Hellenistic scientists. Leonardo understood both but, as demonstrated by what he says, he draws his conclusions from observation rather than experimentation:

> *"……il mare è la più bassa parte del mondo: nessuna superficie d'acqua può per sua natura essere più bassa che quella del mare. La spera dell'acqua desidera perfetta rotondità e quella parte che supera la sua universal superficie non può durare e con breve tempo si spiana"* "…. the sea is the lowest part of the world: no water surface can by its nature be lower than the one of the sea. The sphere of the water desires perfect roundness and the portion that surmounts its universal surface cannot last and in a short time makes level"

Leonardo's incomplete understanding of Torricelli Theorem does not prevent him from providing the duke Ludovico il Moro a great service in correcting the way water flow was measured. In fact at that time the unit, "oncia", utilized to measure the amount of water dispensed by the irrigation channels was based on the surface of the gate to be open and the time duration of the irrigation. The exit speed of the water was not

taken into account. Leonardo realized that the speed depends on the level of the gate with respect to the level of the water in the channel (as from Torricelli 's law) and thus proposed a change in the definition of the "oncia" to take into account the effect. He did not build himself any device to monitor the flow but the ones built at his time and improved in later centuries are based on Leonardo' observation.

Even in this case one can see that Leonardo was not interested in arriving to a complete and correct statement of the physical laws. Once he had understood enough to solve his practical problems he would be satisfied!

In the Atlantic Code, folio 831 r (*Fig. 13*), it is depicted the system to dispense water in the *Navijlio di san crisstofano dj mjlano facto addi 3 di maggio 1509*.

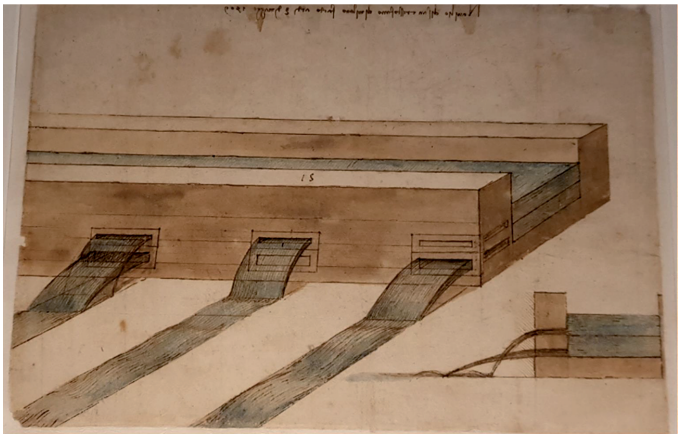


Fig. 13.

Conclusions

Leonardo, although he enunciates correctly the principles on which the modern experimental scientific method is based, does not seem to implement them and thus we cannot say that he anticipated the scientific revolution of Galileo. Part of the problem is his limited understanding of mathematics. Of course many necessary mathematical tools were not available at his time,like symbolic algebra or calculus, but even the tools available were not very familiar to him even after the teachings

of his friend Luca Pacioli. But the main problem is that he uses the experimentation in a limited way. In fact he often draws experiments but he then fails to follow up with the implementation of his drawings and he never gives quantitative results for the experiments granted that he did perform them at all. Furthermore the reasoning that follows up his experimentation is more often analogical than inductive.

He did much better with geometry which is more directly related to his extraordinary ability as an artist and this explains why he was very successful as a descriptive scientist in areas such as anatomy, engineering, geology, botany, cartography. As a visual scientist he is very successful because of his insatiable curiosity and immense capacity in observing nature and phenomena and in describing them with his drawings. In this way he arrives often to incredible intuitions. But an experimental science like physics requires to set up specific experiments to interrogate nature and in this way to materialize the intuitions and this is not generally done by Leonardo.

The great scientific merit of Leonardo is the one of having helped to rediscover the great knowledge of the Hellenistic world and at the same time to having pointed out the necessity to check with its own experience the validity of the previous knowledge without fear of rejecting what appeared to be incorrect. In this respect he stands, as analyzed so well by Luigi Borzacchini, a solitary giant connecting the old scientific world and to the new.

Aknowledgements

I would like to thank A. Rigamonti and P. Mazzarello for stimulating discussions.

REFERENCES

Reference to articles and books dealing specifically with the physics of Leonardo are given in the text. We list below some general references which were consulted in writing this note.

Jean Paul Richter, The notebooks of Leonardo da Vinci, Dover Publications (1970) – original publication 1883.
Luca Beltrami, Scritti di Leonardo, Istituto Editoriale Italiano (1920).
Carlo Pedretti, Leonardo da Vinci, Giunti Gruppo Editoriale, (2000).
Michael White, "Leonardo: the first scientist", St. Martin Press, New York (2000).
Augusto Marinoni, Trascrizioni codice atlantico Vol 1 ,2, 3.
Paolo Galluzzi, Mostre, Conferenze, Interviste, Articoli, Internet.
Domenico Laurenza, Enciclopedia Treccani, ottava appendice, 2013.
Charles Nicholl, Leonardo da Vinci: Flights of the mind, Penguing books, 2004.
Stefan Klein, L'eredità di Leonardo, Bollati Boringhieri, 2017 (2005).
Mario De Micheli, Leonardo da Vinci, l'uomo e la natura, Feltrinelli (1991).
Walter Isaacson, Leonardo da Vinci, Simon and Schuster, 2017.
Carlo Vecce, Leonardo, Salerno editrice (2005).
Edoardo Zanon, Codex on flight, Leonardo 3 srl, (2000).
Giovanni Giovannetti, Indagine su Leonardo, Effigie (2015).

Istituto Lombardo (Rend. Scienze) 153, 115-130 (2019)

SPETTROSCOPIA E IMAGING IN OTTICA DIFFUSA: FOCUS SUL TUMORE DELLA MAMMELLA

PAOLA TARONI (*)

Nota presentata dal m.e. Antonio Pedotti
(Adunanza del 23 maggio 2019)

SUNTO. – Attraverso la misura delle proprietà ottiche (assorbimento e diffusione), la spettroscopia ottica diffusa consente la stima non invasiva della composizione dei tessuti biologici (contenuto di acqua, lipidi e collagene) e dei parametri funzionali del sangue e fornisce informazioni sulla struttura microscopica del tessuto. Può quindi essere efficacemente utilizzata *in vivo* come strumento diagnostico assolutamente non invasivo. Il Dipartimento di Fisica del Politecnico di Milano ha progettato e realizzato un mammografo ottico che sfrutta proprio l'ottica diffusa, operando con luce impulsata a 7 lunghezze d'onda nel rosso e vicino infrarosso (635-1060 nm). Lo strumento è stato utilizzato in uno studio clinico condotto su 200 soggetti, in collaborazione con l'Istituto Europeo di Oncologia: la composizione tissutale derivata otticamente e in particolare il contenuto di collagene nei tessuti si sono rivelati efficaci sia per discriminare tra lesioni mammarie maligne e benigne, che per stimare il rischio di tumore della mammella legato alla densità del tessuto mammario, che è riconosciuta tra i più importanti fattori indipendenti di rischio. Anche sulla base di questi risultati, "SOLUS – Smart optical and ultrasound diagnostics of breast cancer", progetto europeo del Programma Quadro H2020, sta ora lavorando per migliorare la specificità della diagnosi non invasiva del tumore della mammella combinando ottica diffusa con imaging a ultrasuoni.

***

ABSTRACT. – Through the measurement of the optical properties (absorption and scattering), diffuse optical spectroscopy allows one to estimate non-invasively the composition of biological tissues (water, lipid and collagen content) and functional blood parameters. Further, it provides information on the microscopic tissue structure. It can the-

(*) Dipartimento di Fisica, Politecnico di Milano, Italy.
E-mail: paola.taroni@polimi.it

refore be effectively used *in vivo* as an absolutely non-invasive diagnostic tool. The Department of Physics of the Politecnico di Milano has designed and built an optical mammograph that exploits diffused optics, operating with pulsed light at 7 wavelengths in the red and near infrared spectral range (635-1060 nm). The instrument was used in a clinical study on 200 subjects, in collaboration with the European Institute of Oncology: optically derived tissue composition and in particular collagen content in tissues proved to be effective both to discriminate between malignant and benign breast lesions, and to estimate the risk of breast cancer related to the density of breast tissue, which is recognized among the most important independent risk factors. Partly based on those results, "SOLUS - Smart optical and ultrasound diagnostics of breast cancer", a European project in the H2020 Framework Program, is now working to improve the specificity of non-invasive breast cancer diagnosis by combining diffuse optics with ultrasound imaging.

1. L'ottica diffusa e i tessuti biologici

Quando la luce viene iniettata in un tessuto biologico, solo una piccola frazione viene riflessa direttamente dalla superficie. La maggior parte della luce entra nel mezzo e viene assorbita o diffusa al suo interno.

Sia l'assorbimento che la diffusione (scattering) sono fortemente dipendenti dalla lunghezza d'onda: nel visibile si verificano con probabilità comparabile, mentre a lunghezze d'onda maggiori (rosso e vicino infrarosso, near infrared - NIR) gli eventi di scattering diventano dominanti, ma entrambi i processi sono molto meno probabili, consentendo un penetrazione della luce molto maggiore. Di conseguenza, per avere accesso a strati di tessuto più profondi per scopi diagnostici, vengono solitamente scelte lunghezze d'onda comprese tra 600 e 1100 nm.

Attraverso un modello di propagazione della luce nel tessuto si ottiene poi una correlazione tra il modo in cui la luce viene assorbita e diffusa durante la propagazione attraverso un mezzo e le proprietà ottiche del mezzo stesso (essenzialmente la probabilità di assorbimento o scattering per unità di lunghezza, cioè i coefficienti di assorbimento e di scattering m_a e m_s, rispettivamente). In pratica, i fotoni sopravvissuti alla riflessione o trasmissione diffusa (a seconda della geometria di misura) vengono rilevati e forniscono informazioni sul mezzo attraversato.

Il trasporto della luce attraverso mezzi altamente diffusivi può in linea di principio essere descritto utilizzando la teoria analitica basata direttamente sulle equazioni di Maxwell e tenendo conto della natura

ondulatoria della luce. Tuttavia, a causa della complessa struttura del tessuto, l'applicazione della teoria elettromagnetica è estremamente difficile e difficilmente utilizzabile nella pratica. In alternativa, può essere considerata la teoria del trasporto radiativo. Basandosi sulla natura particellare della luce, essa descrive il trasporto di fotoni nel tessuto come la propagazione di un flusso di particelle, come avviene, ad esempio, per i fasci di neutroni. La teoria del trasporto è matematicamente meno rigorosa della teoria elettromagnetica e ulteriori approssimazioni (che portano alla cosiddetta "teoria della diffusione") devono essere introdotte per ottenere una soluzione analitica del problema: gli eventi di scattering devono essere isotropi e fortemente dominanti sugli eventi di assorbimento e diversi eventi di scattering devono verificarsi nel mezzo prima che la descrizione fornita diventi sufficientemente accurata [1]. In particolare, per far coincidere l'ipotesi di scattering isotropo con l'evidenza sperimentale della forte anisotropia tipicamente osservata in un tessuto biologico, viene introdotto il coefficiente di scattering ridotto μ'_s, con $\mu'_s = \mu_s (1 - g)$, dove g è il fattore di anisotropia, cioè il coseno medio dell'angolo di diffusione. Nonostante i suoi limiti, l'approssimazione della diffusione alla teoria del trasporto radiativo si è dimostrata ampiamente utile per descrivere il trasporto della luce nei tessuti biologici. In realtà, soluzioni esatte dell'equazione del trasporto radiativo sono state recentemente introdotte per geometrie specifiche e sono disponibili altri modelli teorici (ad esempio, onde di densità di fotoni [2]. Tuttavia, l'approssimazione della diffusione è ancora l'approccio teorico più utilizzato in campo biomedico, perché le sue ipotesi principali sono soddisfatte nel vicino infrarosso e fornisce una semplice soluzione analitica al problema della propagazione dei fotoni in mezzi altamente diffusivi, come i tessuti biologici.

Una delle applicazioni diagnostiche dell'ottica diffusa che negli anni ha generato maggiore interesse è la diagnosi del tumore della mammella (mammografia ottica). In breve, l'idea è simile a quella alla base della convenzionale mammografia a raggi X. Nel caso ottico, radiazione di bassa potenza, nel vicino infrarosso, viene iniettata nella mammella e la luce rivelata dopo essersi propagata attraverso il tessuto mammario porta informazioni sulle sue proprietà ottiche e su eventuali disomogeneità ottiche, rivelando la presenza e la natura di situazioni patologiche.

Affinché una tale tecnica possa essere utile dal punto di vista diagnostico, è innanzitutto essenziale conoscere le proprietà ottiche di

un tessuto sano e come queste siano influenzate dall'insorgere di una condizione patologica. Tuttavia, il tessuto mammario stesso non è un mezzo otticamente omogeneo e mostra anzi una complessa struttura eterogenea, come prevedibile in base alla sua morfologia. Analogamente a quanto normalmente osservato per altri parametri biologici, le proprietà ottiche del tessuto mammario possono variare sensibilmente da soggetto a soggetto, anche se appartengono alla stessa categoria per quanto riguarda i parametri demografici (ad esempio, hanno la stessa età o lo stesso indice di massa corporea). Inoltre, i cambiamenti fisiologici, come quelli che si verificano durante il ciclo mestruale, la gravidanza o l'allattamento, possono influenzare fortemente le proprietà ottiche del tessuto mammario. Questo, insieme alla variabilità biologica inter- e intra-soggetto sopra menzionata, rende l'indagine del tessuto biologico da un punto di vista ottico piuttosto complessa.

2. Tecniche di misura

Per quanto riguarda le tecniche sperimentali, possono essere previsti tre diversi approcci per stimare le proprietà ottiche di un mezzo altamente diffusivo come un tessuto biologico. La configurazione più semplice implica l'uso di luce in stato stazionario (continuous wave - CW) per eseguire misure dell'attenuazione della luce causata dalla propagazione attraverso il mezzo studiato. L'attenuazione generalmente dipende sia dalle proprietà di assorbimento che di scattering. Perciò una singola misura di attenuazione non consente di separare i contributi dei due fenomeni. Per caratterizzare un mezzo, in generale vengono eseguite più misure a diverse distanze tra sorgente e rivelatore.

Le tecniche che lavorano nel dominio della frequenza misurano l'intensità di luce modulata sinusoidalmente ad almeno una frequenza, tipicamente di poche centinaia di MHz. La propagazione in un mezzo biologico provoca attenuazione (demodulazione) e ritardo di fase della luce riemessa rispetto a quella iniettata. La modulazione e lo sfasamento del segnale rivelato sono legati ai coefficienti ottici del mezzo ed i loro valori, interpretati utilizzando un modello come la teoria della diffusione, consentono la stima sia di μ_a che di μ'_s.

Infine, l'approccio che opera nel dominio del tempo implica

l'uso di brevi impulsi luminosi (con durata tipica dell'ordine di decine di picosecondi). Quando iniettato in un mezzo diffusivo, un impulso luminoso è fortemente attenuato (fino a 5-7 ordini di grandezza nelle applicazioni pratiche), ma anche allargato (fino ad alcuni nanosecondi) e ritardato, principalmente a causa delle proprietà di diffusione del mezzo attraversato. Più in dettaglio, ci aspettiamo che gli eventi di scattering rendano le traiettorie dei fotoni diverse tra loro e molto più lunghe della distanza tra i punti di iniezione e di raccolta, ritardando così in modo significativo i fotoni rivelati e ampliando l'intervallo di tempo nel quale vengono rivelati rispetto alla durata dell'impulso incidente. Gli effetti dei processi di assorbimento sono più evidenti sulla coda dell'impulso, dove si rilevano fotoni che sono rimasti più a lungo nel mezzo, sperimentando così una maggiore probabilità di essere assorbiti. Pertanto l'assorbimento, rimuovendo fotoni a tempi lunghi, aumenta la pendenza negativa della coda dell'impulso. Ritardo e cambiamenti di forma dell'impulso rivelato possono essere correlati ai due coefficienti ottici e consentono la loro stima da una singola misura in riflessione o trasmissione. In questo modo non è necessario utilizzare informazioni derivate dall'ampiezza dell'impulso e si evitano così tutti i problemi pratici legati alla stima assoluta dell'intensità luminosa (problemi che, al contrario, possono influenzare in modo significativo le misurazioni CW).

La dualità degli studi risolti in frequenza e in tempo è evidente, se si considera che la trasformata di Fourier consente di passare da un dominio all'altro. Tuttavia, si dovrebbe tener conto che in linea di principio tutte le frequenze sono contenute in una misura risolta nel tempo e anche in pratica la gamma di frequenze che corrispondono ad una misura nel dominio del tempo è molto ampia e limitata solo dalla risoluzione temporale dell'apparato sperimentale. Pertanto, una singola misura risolta nel tempo garantirà un contenuto informativo più elevato rispetto a un'acquisizione risolta in frequenza eseguita a una singola frequenza o a un limitato numero di frequenze. Per cercare di compensare questo limite, i sistemi più sofisticati che lavorano in frequenza utilizzano un sintetizzatore per eseguire la scansione di un'ampia gamma di frequenze.

In passato la necessità di strumentazione più complessa e costosa ha limitato la diffusione delle tecniche che operano nel dominio del tempo. Tuttavia, recentemente il rapido sviluppo tecnologico in campo optoelettronico ha fortemente promosso l'utilizzo di sistemi risolti nel

tempo, che stanno diventando una scelta comune ed efficace non solo per l'uso di laboratorio, ma anche in applicazioni cliniche.

3. Il mammografo ottico del Politecnico di Milano

3.1 La configurazione dello strumento

Lo strumento è stato progettato per acquisire immagini in proiezione con la mammella in compressione, nella stessa configurazione della convenzionale mammografia a raggi X (*Fig. 1*). Sette laser a diodi impulsati vengono utilizzati come sorgenti luminose che emettono a 635, 680 e 785 nm (visibile, VIS) e a 905, 930, 975 e 1060 nm (vicino infrarosso, NIR), con potenza media di ~1-5 mW, durata dell'impulso di ~150-400 ps e frequenza di ripetizione di 20 MHz. La mammella è compresso tra lastre di vetro temperato parallele, con un grado di compressione molto più modesto che nel caso della mammografia X, dato che la lunghezza della misura è maggiore. Nella versione dello strumento utilizzata per lo studio descritto nel seguito, la luce trasmessa attraverso la mammella veniva raccolta da un fascio di fibre, la cui estremità distale era biforcata e i due bracci guidavano i fotoni a un tubo fotomoltiplicatore (PMT) per la rivelazione delle lunghezze d'onda VIS (R5900U-01-L16, Hamamatsu, Giappone) e a un PMT per le lunghezze d'onda NIR (sensibile fino a 1100 nm, H7422P-60, Hamamatsu, Giappone), rispettivamente. Due schede (SPC130, Becker & Hickl, Germania) per il conteggio dei singoli fotoni correlati nel tempo (TCSPC) venivano utilizzate per l'acquisizione delle 7 curve di trasmittanza risolte nel tempo. L'immagine veniva acquisita mediante una scansione (raster scan) continua, registrando i dati ogni millimetro. Una scansione completa di una vista richiedeva in genere qualche minuto.

Successivamente, per aumentare il livello del segnale raccolto soprattutto alle lunghezze d'onda elevate, è stata rinnovata la catena di acquisizione, sostituendo il fascio di fibre di raccolta della luce e i fotomoltiplicatori con una sonda costituita da 8 fotomoltiplicatori a Silicio (SiPM, S13360-1350PE, Hamamatsu Photonics, Giappone) e le schede di acquisizione con un "time-to-digital converter" (TDC, SC-TDC-1000/08 S, Surface Concept, Germania) a 8 canali.

Una descrizione dettagliata delle due configurazioni dello strumento è riportata in [3] e in [4], rispettivamente.



Fig. 1. Il mammografo ottico sviluppato dal Dipartimento di Fisica del Politecnico di Milano

4. Analisi dei dati

4.1 Modello omogeneo per la caratterizzazione media del tessuto mammario

Le informazioni sulla composizione e la struttura del tessuto sono ottenute direttamente dalle curve risolte nel tempo misurate a 7 lunghezze d'onda. Viene applicata una procedura di interpolazione globale vincolata spettralmente, basata su una soluzione analitica dell'approssimazione di diffusione per una lastra omogenea infinita [5]. I parametri liberi sono le concentrazioni di ossi- e deossi-emoglobina (HbO_2 e Hb, rispettivamente), acqua, lipidi e collagene, insieme ai parametri di scattering (ampiezza a e potenza b).

La legge di Beer viene utilizzata per stimare il coefficiente di assorbimento a ciascuna lunghezza d'onda dalle concentrazioni dei principali costituenti del tessuto.

Conoscendo i parametri a e b, il coefficiente di scattering ridotto viene modellato e ottenuto mediante una semplice approssimazione alla teoria di Mie: $\mu'_s = a\,(\lambda/\lambda_o)^{-b}$, dove poniamo $\lambda_o = 600$ nm ed a è il coefficiente di scattering interpolato $\mu'_s\,(\lambda_o)$ [6].

4.2 Modello perturbativo per la caratterizzazione delle lesioni

Al fine di caratterizzare otticamente le lesioni mammarie, assumiamo che possano essere trattate come perturbazioni di assorbimento localizzate, presenti in un mezzo diffusivo altrimenti omogeneo. Le curve di trasmittanza risolte nel tempo imperturbate e perturbate, rispettivamente $T_0(t)$ e $T(t)$, possono essere collegate sfruttando la legge di Lambert-Beer modificata:

$$T(t)=T_0(t)e^{-\Delta\mu_a l(t)} \tag{1}$$

dove $\lambda(t)$ è la lunghezza del percorso media percorsa nella disomogeneità dai fotoni rilevati al tempo t, mentre $\Delta\mu_a$ rappresenta la variazione di assorbimento all'interno della disomogeneità rispetto all'assorbimento omogeneo imperturbato $\mu_{a,0}$. La lunghezza del percorso $l(t)$ può essere derivata come [7]:

$$l(t) = -\frac{1}{T_0(t)}\frac{\partial T_0(t)}{\partial \mu_a} \tag{2}$$

La curva imperturbata è ottenuta come media su una regione che esclude i bordi e le disomogeneità marcate. La curva perturbata sperimentale T_{pert} è ottenuta da un'area ("area della lesione") centrata sulla corrispondente posizione della disomogeneità (cioè della lesione). L'area della lesione dipende dalla dimensione della lesione stessa. Per diametri di lesione >15 mm, è stata selezionata un'area di lesione di 9×9 mm^2, altrimenti 5×5 mm^2.

Partendo da $\Delta\mu_a$ e conoscendo il coefficiente di estinzione dei principali costituenti del tessuto mammario, dalla legge di Beer si stima la variazione delle concentrazioni ΔC_i tra lesione e tessuto sano (in termini di sangue, acqua, lipidi e collagene).

Questo metodo perturbativo si basa sulla conoscenza a priori del volume e della posizione della disomogeneità. Si è sempre ipotizzata una disomogeneità sferica posta a metà cammino tra sorgente e rivelatore. Per quanto riguarda le dimensioni si è considerata una sfera di volume equivalente in base al diametro massimo della lesione ottenuto dall'istopatologia, quando disponibile, o dalla mammografia o ecografia.

I dettagli dell'implementazione del metodo perturbativo sono riportati in [3].

5. Risultati e discussione

5.1 Misura non invasiva della densità mammaria come fattore di rischio

I principali costituenti del tessuto mammario (cioè acqua, lipidi, proteine strutturali e sangue) mostrano caratteristiche di assorbimento marcate e specifiche nell'intervallo spettrale NIR, come mostrato in *Fig. 2*. Il loro contributo all'assorbimento complessivo del tessuto varia in modo significativo con il tipo di mammella: passando da mammelle estremamente dense (cioè fortemente fibrose) ad adipose, l'assorbimento dell'acqua e del collagene può diminuire di un fattore 2 o anche 3, mentre i lipidi aumentano corrispondentemente. Nel primo caso lo spettro è quindi dominato dall'assorbimento dell'acqua e si possono raggiungere valori molto elevati (μ_a ~ 0,35 cm^{-1}) sul picco dell'acqua a 975 nm, come mostrato in *Fig. 3a*. Al contrario, l'assorbimento dei lipidi a 930 nm è la caratteristica principale nello spettro dei seni adiposi, ma l'assorbimento complessivo rimane relativamente basso a qualsiasi lunghezza d'onda (raggiungendo valori di circa 0,15 cm^{-1} a 930 nm) [8].

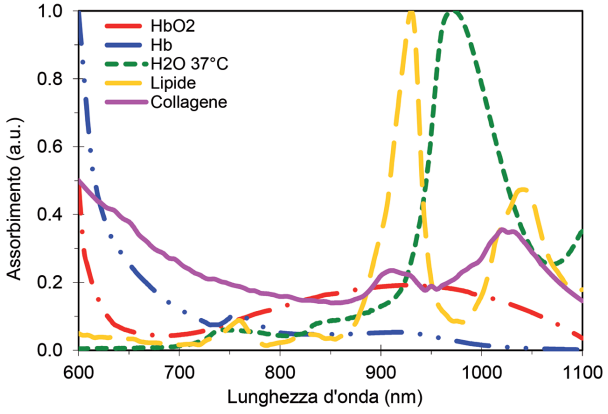


Fig. 2. Spettri di assorbimento normalizzati dei principali assorbitori di tessuto nel vicino infrarosso: deossi- e ossi-emoglobina (Hb e HbO_2, rispettivamente), acqua, lipidi e collagene.

Il coefficiente di scattering ridotto diminuisce progressivamente all'aumentare della lunghezza d'onda, senza alcuna particolare caratteristica spettrale. Tuttavia, diversi tipi di seno hanno proprietà di scattering distinte, legate in particolare alla diversa pendenza dello spettro, con le mammelle fibrose caratterizzate da pendenze più ripide di quelle adipose (*Fig. 3b*). Questo, nell'approssimazione di Mie utilizzata nell'analisi dei dati, corrisponde ad un più elevato valore del parametro *b*.

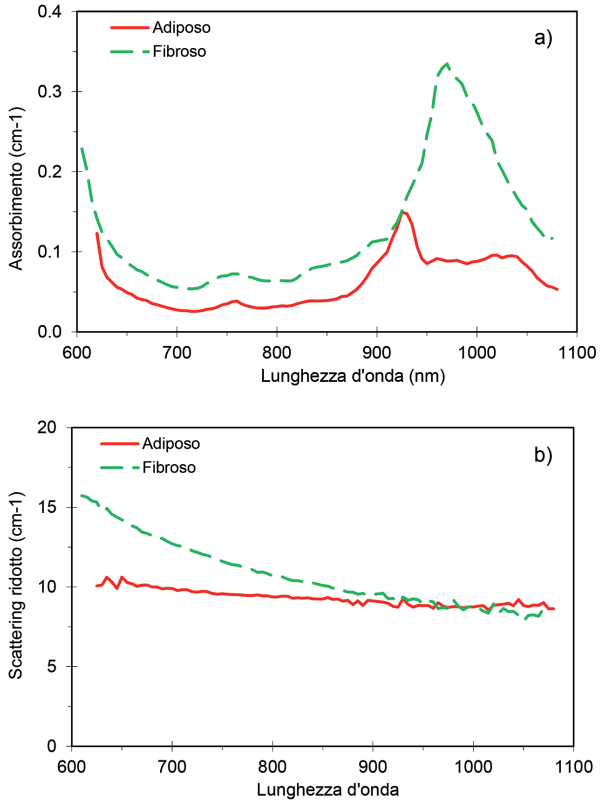


Fig. 3. Spettri tipici di assorbimento (a) e scattering ridotto (b) misurati in vivo da mammelle di diverso tipo: adiposo (solido) e fibroso (tratteggiato).

Più in generale, come decritto precedentemente, le proprietà ottiche stimate possono essere correlate alla composizione del tessuto (interpretando l'assorbimento con la legge di Beer) e alla struttura

microscopica (descrivendo lo scattering con l'approssimazione empirica della teoria di Mie).

In questo modo è stata ottenuta una buona correlazione tra la composizione del tessuto derivata dalle misure ottiche e la densità della mammella. La densità è la frazione di tessuto fibroghiandolare ("denso", perché attenua i raggi X) presente nel tessuto mammario, altrimenti sostanzialmente costituito da tessuto adiposo. A parte l'età, la densità è il più importante fattore di rischio indipendente per lo sviluppo del tumore della mammella. Attualmente viene derivata dall'analisi visuale delle immagini X. Il corrispondente fattore di rischio diventa quindi noto solo al momento della prima mammografia, tipicamente a 50 anni. Sarebbe invece molto utile avere a disposizione un metodo non invasivo, che non utilizzi radiazioni ionizzanti, per valutare la densità molto prima, dato che esistono interventi preventivi per ridurre la densità e il rischio correlato, non solo mediante l'assunzione di farmaci specifici, ma anche agendo sullo stile di vita (ad esempio, sul tipo di alimentazione).

In questa direzione l'ottica diffusa sta fornendo risultati molto promettenti. In particolare, in uno studio su più di 200 soggetti, in collaborazione con l'Istituto Europeo di Oncologia, lo strumento sviluppato dal Politecnico di Milano ha consentito di ottenere una buona correlazione [9,10] tra la concentrazione di collagene e le classi BIRADS (Breast Imaging Reporting and Data System [11]), utilizzate nella pratica clinica per classificare la densità (da BIRADS 1 a BIRADS 4, per densità crescenti). Questo risultato rispecchia la fisiologia dei tessuti, dato che un'elevata densità indica un'elevata frazione fibroghiandolare, più ricca di collagene rispetto al tessuto adiposo. In accordo con questa osservazione è anche la buona correlazione ottenuta tra la pendenza *b* dello spettro di scattering e le classi BIRADS. Il tessuto ricco di collagene è infatti caratterizzato da uno spettro si scattering molto pendente, probabilmente a causa della struttura microscopica del collagene nel tessuto.

6. Discriminazione tra lesioni maligne e benigne

L'utilizzo di lunghezze d'onda elevate (rosso e NIR) garantisce non invasività all'ottica diffusa. Allo stesso tempo, però, essendo molto probabili gli eventi di scattering, i fotoni non procedono secondo traiet-

torie rettilinee e questo limita significativamente la risoluzione spaziale della tecnica. La risoluzione dipende dalle dimensioni della disomogeneità ottica che si deve rivelare, dalla sua profondità e dal suo contrasto ottico, ma difficilmente può essere migliore di alcuni millimetri e spesso, in condizioni reali di misura, è inferiore [12]. L'ottica diffusa non è quindi adatta alla diagnosi basata sulle caratteristiche morfologiche. La capacità di stimare la composizione del tessuto, insieme a importanti parametri sanguigni (volume totale di emoglobina *HbT* e livello di ossigenazione SO_2) possono però essere di aiuto nella discriminazione tra lesioni maligne e benigne. Lo stesso studio già citato nella sezione precedente ha anche consentito di confrontare la composizione di lesioni maligne e benigne, derivata otticamente.

Tutti i tipi di lesioni analizzate avevano mediamente un maggiore contenuto di emoglobina e di collagene rispetto al tessuto sano nella stessa mammella, ma la differenza è risultata più marcata per le lesioni maligne [13]. Un algoritmo di "machine learning" per il "boosting" adattativo è stato poi applicato utilizzando come variabili per la discriminazione la composizione del tessuto ottenuta dalle misure ottiche, insieme ad informazioni note dall'anamnesi delle pazienti e che rappresentano fattori di rischio (ad esempio, età, numero di gravidanze, terapia ormonale sostitutiva, ecc.). In questo modo le lesioni maligne sono state discriminate da quelle benigne con sensibilità = 87.7 ± 1.98%, and specificità = 78.7 ± 3.0% [14].

Il risultato è abbastanza promettente, anche se si deve tenere conto del fatto che è stato ottenuto da un limitato numero di casi (N = 84).

7. Approccio multimodale alla diagnosi del tumore della mammella: il progetto SOLUS

L'ecografia (ultrasuoni - US) è la tecnica di prima scelta per valutare le lesioni non palpabili visibili alla mammografia e le lesioni palpabili e anche per guidare la biopsia mammaria. Le caratteristiche morfologiche sono state classificate per distinguere tra lesioni maligne e benigne. Attualmente, però, i risultati diagnostici degli US convenzionali (in modalità B) non sono del tutto soddisfacenti e per una diagnosi definitiva si ricorre alla biopsia, che in molti casi si dimostra poi inutile (falso positivo).

Recentemente, è stata introdotta una tecnica a ultrasuoni più avanzata, detta elastografia (shear wave elastography - SWE).

L'elastografia fornisce una misura quantitativa e riproducibile della rigidezza dei tessuti e sfrutta la maggiore presenza di matrice extracellulare nel tessuto neoplastico come marker di malignità. Una recente meta-analisi ha valutato l'efficacia dell'elastografia per la diagnosi del tumore della mammella: la specificità degli US convenzionali è risultata del 55%, mentre l'aggiunta dell'elastografia aumenta la specificità all'80%, lasciando la sensibilità sostanzialmente inalterata al 95% [15]. Tuttavia, è necessario un ulteriore miglioramento della specificità per ottenere una riduzione significativa del tasso di falsi positivi e generare un impatto reale sia sulle pazienti (evitando un inutile esame invasivo) che sui sistemi sanitari (riducendo la spesa per le biopsie).

Il progetto SOLUS – "Smart Optical and Ultrasound Diagnostics of Breast Cancer" [16], finanziato nell'ambito del programma quadro europeo H2020, ha proprio lo scopo di sviluppare un sistema di imaging multimodale per migliorare in modo non invasivo la specificità della diagnosi del tumore della mammella, discriminando in particolare le lesioni che sono al limite tra benigne e maligne.

Lo sviluppo della sonda multimodale SOLUS è partito da uno strumento per imaging a US (Aixplorer® Mach30, commercializzato da Supersonic Imagine, partner del progetto), che consente di acquisire immagini in modalità B ed effettuare la misura della rigidezza mediante SWE. Per poter valutare la composizione del tessuto nello stesso volume analizzato dagli US, attorno al trasduttore US sono stati aggiunti otto punti di misura ottica che consentono di effettuare una tomografia ottica diffusa a 8 lunghezze d'onda nel dominio del tempo. Per ottenere questo risultato, è stato necessario combinare elevate performance e piccole dimensioni. Sono stati quindi sviluppati tutti gli elementi del sistema: laser impulsati a picosecondi a 8 lunghezze d'onda (630-1060 nm), rivelatore SiPM di area elevata (per ottenere elevata sensibilità) con gate temporale (per poter gestire segnali che variano su vari ordini di grandezza al variare della distanza tra sorgente e rivelatore) ed elettronica di acquisizione di segnali nel dominio del tempo. L'insieme di tutti questi elementi è racchiuso in un volume con sezione di circa 1 cm^2 e lunghezza di circa 4 cm, ordini di grandezza più piccolo del volume occupato dagli strumenti finora disponibili per eseguire lo stesso tipo di misure con performance analoghe.

La sonda è attualmente in fase di caratterizzazione. Appena questa sarà conclusa, inizierà la valutazione clinica del sistema SOLUS per la discriminazione tra lesioni maligne e benigne.

BIBLIOGRAFIA

1. Contini D, Martelli F, Zaccanti G, Photon migration through a turbid slab described by a model based on diffusion approximation. I. Theory, *Appl Opt.*, 1997: 36:4587–4599.
2. Fishkin JB, Gratton E, Propagation of photon-density waves in strongly scattering media containing an absorbing semi-infinite plane bounded by a straight edge. *J Opt Soc Am A.*, 1993: 10: 127.
3. Taroni P, Pifferi A, Salvagnini E, Spinelli L, Torricelli A, Cubeddu R, Seven-wavelength time-resolved optical mammography extending beyond 1000 nm for breast collagen quantification, *Opt Express*, 2009: 17(18):15932–15946.
4. Ferocino E, Martinenghi E, Mora AD, Pifferi A, Cubeddu R, Taroni P, High throughput detection chain for time domain optical mammography, *Biomed Opt Express.*, 2018: 9: 755-770.
5. D'Andrea C, Spinelli L, Bassi A, Giusto A, Contini D, Swartling J, Torricelli A, Cubeddu R, Time-resolved spectrally constrained method for the quantification of chromophore concentrations and scattering parameters in diffusing media, *Opt. Express,* 2006. 14: 1888–1898.
6. Nguyen J, Hayakawa CK, Mourant JR, Spanier J, Perturbation Monte Carlo methods for tissue structure alterations, *Biomed Opt Express*, 2013: 4: 1946–1963.
7. Del Bianco S, Martelli F, Zaccanti G, Penetration depth of light re-emitted by a diffusive medium: theoretical and experimental investigation, *Phys Med Biol*, 2002: 47: 4131–4144.
8. Taroni P, Quarto G, Pifferi A, Abbate F, Balestreri N, Menna S, Cassano E, Cubeddu R, Breast tissue composition and its dependence on demographic risk factors for breast cancer: Non-invasive assessment by Time Domain diffuse optical spectroscopy, *PLoS One*, 2015: 10: e0128941.
9. Taroni P, Quarto G, Pifferi A, Ieva F, Paganoni AM, Abbate F, Balestreri N, Cassano E, Cubeddu R, Optical identification of subjects at high risk for developing breast cancer, *J Biomed Opt*, 2013: 18: 060507.
10. Taroni P, Pifferi A, Quarto G, Spinelli L, Torricelli A, Abbate F, Balestreri N, Cassano E, Cubeddu R, Noninvasive assessment of breast cancer risk using time-resolved diffuse optical spectroscopy. *J Biomed Opt*, 2010: 15: 060501.
11. BI-RADS® – Mammography 2013 - American College of Radiology.
12. Cubeddu R, Pifferi A, Taroni P, Torricelli A, Valentini G, Spatial resolution of imaging with diffusing light: Edge spread function measurements on a realistic tissue phantom, *Med Phys*, 1999: 26: 462-471.
13. Quarto G, Spinelli L, Pifferi A, Torricelli A, Cubeddu R, Abbate F, Balestreri N, Menna S, Cassano E, Taroni P, Estimate of tissue composition in malignant and benign breast lesions by time-domain optical mammography. *Biomed Opt Express*, 2014: 5: 3684-3698.
14. Taroni P, Paganoni AM, Ieva F, Pifferi A, Quarto G, Abbate F, et al. Non-invasive optical estimate of tissue composition to differentiate malignant from benign breast lesions: A pilot study. *Sci Rep*, 2017: 7: 40683.
15. Liu B, Zheng Y, Huang G, Lin M, Shan Q, Lu Y, Tian W, Xie X, Breast Lesions:

Quantitative Diagnosis Using Ultrasound Shear Wave Elastography-A Systematic Review and Meta-Analysis. *Ultrasound Med Biol*, 2016: 42: 835–847.
16. SOLUS - Smart Optical and Ultrasound Diagnostics of Breast Cancer [Internet]. Progetto H2020 No. 731877. http://solus-project.eu/

SELF-ASSEMBLY DI DNA SU SUPERFICIE: UNO STUDIO COMBINATO NANOLITOGRAFIA/SPETTROSCOPIA OTTICA

ORNELLA CAVALLERI (*), *et al.*[#]

Nota presentata dal m.e. Giorgio Benedek
(Adunanza del 24 ottobre 2019)

SUNTO. – Il *self-assembly* di molecole organiche e biomolecole su superfici inorganiche è un processo complesso che deriva dal bilancio di diverse interazioni: molecola/molecola, molecola/superficie, molecola/solvente. L'interesse per lo studio del *self-assembly* molecolare deriva dalla possibilità di realizzare superfici funzionalizzate che presentano le proprietà funzionali proprie delle (bio) molecole depositate. I SAM (Self-Assembled Monolayers) trovano infatti ampia applicazione nel biosensing. Tra i SAM più studiati figurano i SAM di organosolfuri su oro. Nel seguito viene presentato un approccio che combina la microscopia a scansione di sonda e la spettroellissometria per lo studio delle proprietà morfologiche, meccaniche ed ottiche di SAM costituiti da filamenti di DNA, un sistema ampiamente studiato poiché, sfruttando l'ibridazione tra filamenti di DNA complementari, è alla base della realizzazione dei microarray di DNA.

***

ABSTRACT. – The *self-assembly* of organic molecules and biomolecules on inorganic surfaces is a complex process which results from the interplay between different interactions: molecule/molecule, molecule/surface, and molecule/solvent interactions. The interest in the study of molecular *self-assembly* derives from the possibility to obtain functionalized surfaces that exhibit the functional properties of the deposited (bio) molecules. Self-Assembled Monolayers (SAMs) are indeed widely used in biosensing.

(*) Dipartimento di Fisica, Università degli Studi di Genova, Italy.
E-mail: cavalleri@fisica.unige.it

La nota è il risultato di un lavoro di ricerca condotto in collaborazione con: Giulia Pinto e Maurizio Canepa.

SAMs of organosulphur compounds on gold are among the most widely studied SAMs. In the following, a combined approach that couples scanning probe microscopy and spectroellipsometry is discussed and applied to the study of the morphological, mechanical and optical properties of SAMs of DNA single strands. DNA SAMs are a widely investigated system since they can be exploited for the design of DNA microarrays based on the hybridization between complementary DNA strands.

1. Introduzione

A partire dagli anni ’80 con i primi lavori di Nuzzo, Allara e Whitesides sul *self-assembly* di alcantioli di superfici d’oro[1–3], lo studio del *self-assembly* di molecole organiche su superfici ha conosciuto un enorme sviluppo. Negli anni ’80-’90, sono stati portati avanti studi di carattere fondamentale indirizzati a comprendere quali siano le interazioni coinvolte nella formazione dei SAM e come queste determinino l’ordinamento molecolare all’interno dei SAM. Schematizzata la molecola come composta da un gruppo di ancoraggio alla superficie, una catena centrale e un gruppo funzionale terminale (*Fig. 1*), tali studi hanno messo in evidenza come la struttura dei SAM derivi essenzialmente dal bilancio di tre tipi di interazioni: le interazioni gruppo di ancoraggio/superficie, responsabili del legame, in genere covalente, del SAM al substrato, le interazioni inter-catena, essenzialmente di tipo van der Waals nel caso di semplici catene alchiliche, e le interazioni tra i gruppi funzionali terminali, la cui natura dipende dalle proprietà del gruppo terminale stesso. Il complesso bilancio di queste interazioni, modulato da fattori dipendenti dalle molecole, quali la lunghezza della catena alchilica o la natura del gruppo terminale (polarità, carica, ingombro sterico, ecc), e da fattori esterni, quali la struttura cristallografica della superficie, il tipo di solvente utilizzato per la deposizione o la temperatura, determina la densità e il grado di ordine molecolare del SAM [4]. In funzione della coppia gruppo di ancoraggio/superficie, si possono identificare due classi principali di SAM: i SAM formati da molecole con il gruppo tiolo (-SH) come gruppo di ancoraggio, che presenta un’elevata affinità per le superfici di metalli quali Au, Ag, Pt, Cu o Hg, e i SAM formati su superfici di ossidi da molecole aventi diversi gruppi di ancoraggio, tra cui silani, fosfonati o carbossilati [5]. Nel seguito ci soffermeremo sui SAM di organosolfuri su oro.

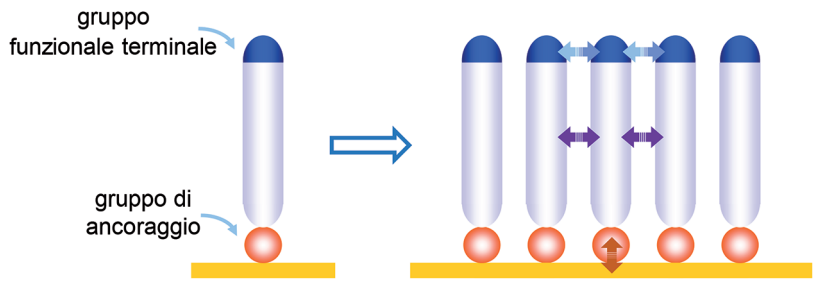


Fig. 1. Dalla molecola su superficie al SAM.

Dopo i primi studi strutturali condotti su SAM formati da molecole relativamente semplici, l'attenzione si è via via spostata verso lo studio di SAM formati da molecole con gruppi funzionali più complessi in grado di dare origine alla formazione di SAM con proprietà funzionali più interessanti da un punto di vista applicativo, dal sensing di ioni metallici [6] alla rilevazione di marker tumorali [7]. Va sottolineato che al crescere delle proprietà funzionali del gruppo terminale è associato in genere un aumento dell'ingombro sterico e della complessità delle interazioni intermolecolari che guidano il processo di *self-assembly*. Nel seguito discuteremo la deposizione, per *self-assembly*, di filamenti di DNA. SAM di singoli filamenti di oligonucleotidi sono alla base della realizzazione di microarray di DNA, ampiamente impiegati nel biosensing. La caratteristica chiave di questi dispositivi è il riconoscimento altamente specifico derivante dall'accoppiamento tra basi complemetari che può essere sfruttato per la rilevazione selettiva di diversi analiti, dagli oligonucleotidi fino alle cellule [8].

L'ibridazione tra sequenze complementari può essere sfruttata per il biosensing su due livelli. Da un lato, può essere utilizzata per il riconoscimento diretto di sequenze target specifiche, con applicazioni in biologia molecolare e nella diagnostica, poiché alcune patologie possono essere identificate sulla base della rilevazione di specifiche sequenze di DNA o RNA. Esempi ne sono l'identificazione di *S. Aureus* [9], di sequenze di mi-RNA correlate a tumori o malattie cardiovascolari [10,11] o di sequenze di RNA di virus [12,13]. D'altra parte, l'ibridazione può anche essere sfruttata per l'immobilizzazione su superficie di molecole sensori che mantengano le loro proprietà selettive e funzionali anche dopo l'immobilizzazione: in questo caso, l'ibridazione tra la sequenza di DNA ancorata alla superficie e la sua sequenza complementare, opportuna-

mente funzionalizzata con la molecola sensore, permette di ancorare quest'ultima alla superficie. In questo modo, le molecole sensore (ad esempio anticorpi) possono essere ancorate alla superficie, pur rimanendo disaccoppiate da essa grazie al SAM di DNA, mantenendo quindi la propria funzionalità e selettività verso specifici marcatori molecolari [14–16].

La rilevanza dei SAM di DNA per lo sviluppo di biosensori ha motivato numerosissimi studi nell'arco degli ultimi 20 anni, riguardanti il *self-assembly* di singoli filamenti di DNA tiolati e della loro ibridazione [17–20]. Per lo sviluppo di sensori robusti ed affidabili, è cruciale infatti procedere attraverso un'attenta caratterizzazione del monostrato di DNA. Rispetto al *self-assembly* di molecole più semplici come gli alcantioli, il chemisorbimento di filamenti di DNA tiolati su oro è un processo più complesso, dal momento che i filamenti di DNA non sono molecole neutre, ma hanno una carica netta negativa dovuta alla presenza dei gruppi fosfato e presentano un maggiore ingombro sterico, parametri entrambi modulabili dalla forza ionica della soluzione [21,22]. Tra i metodi sperimentali impiegati per studiare il *self-assembly* di DNA su oro figurano metodi elettrochimici [23,24], così come metodi spettroscopici, dalla spettroscopia ottica [25], a quella infrarossa [26] e a raggi X [27].

2. Nanolitografia combinata alla spettroellissometria per l'analisi di SAM di DNA

Nel seguito vengono presentati i risultati di un approccio metodologico combinato nanolitografia/spettroscopia ottica per lo studio delle proprietà morfologiche, meccaniche ed ottiche di SAM di DNA.

Un aspetto finora meno studiato dei SAM di DNA è quello che riguarda le loro proprietà meccaniche. La valutazione delle proprietà meccaniche di film ultrasottili come i SAM è un obiettivo complesso, poiché è necessaria un'elevata sensibilità per poter disaccoppiare il contributo del film da quello del substrato. In effetti, la valutazione delle proprietà meccaniche di film organici sottili ed ultrasottili è un aspetto importante per lo sviluppo e l'ottimizzazione dei biosensori, in quanto è stato mostrato che l'interazione tra substrati e campioni biologici, in particolare cellule, è fortemente influenzata dalla rigidità del substrato [28].

Come di seguito riportato, l'utilizzo della microscopia a scansione di sonda, accoppiata a metodi nanolitografici, permette di studiare le proprietà morfologiche e meccaniche di SAM di DNA ed è sufficiente-

mente sensibile da discriminare tra SAM di DNA depositati a diverse forze ioniche.

Per correlare proprietà strutturali e proprietà ottiche dei SAM, gli esperimenti di nanolitografia sono stati accoppiati a misure di spettroellissometria (SE). L'ellissometria spettroscopica è una tecnica ottica affidabile, rapida, altamente sensibile e non perturbativa, che misura la variazione di polarizzazione di un raggio luminoso in seguito alla riflessione da una superficie, fornendo informazioni sullo spessore e le proprietà ottiche di film sottili e ultrasottili. Per film ultrasottili, come i SAM di DNA, spessore e indice di rifrazione sono parametri correlati. Tale correlazione può essere risolta mediante un'analisi combinata sfruttando i dati di spessore ottenuti dalla nanolitografia [29,30].

A partire dai primi lavori di Liu [31,32], la nanolitografia a scansione di sonda (nanoshaving e nanografting) è stata impiegata con successo per produrre interfacce organiche/inorganiche patternate per l'immobilizzazione selettiva di nano-oggetti (dalle nanoparticelle alle biomolecole) [33,34] e per sviluppare piattaforme sensibili per il riconoscimento molecolare [35,36].

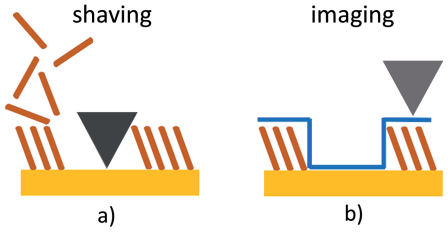


Fig. 2. Rappresentazione schematica dell'utilizzo dell'AFM in modalità shaving (a) e in modalità imaging (b).

Gli esperimenti di nanoshaving possono essere vantaggiosamente impiegati anche per misurare lo spessore degli strati molecolari a livello sub-nanometrico, permettendo così di monitorare le variazioni di spessore del film in seguito all'adsorbimento selettivo di analiti. In un esperimento di nanoshaving la punta del microscopio a forza atomica (Atomic Force Microscope, AFM) viene utilizzata come uno "scalpello" per rimuovere selettivamente le molecole in una regione del monostrato (*Fig. 2*). Modulando la forza applicata, è possibile passare da condizioni di soft imaging (forza bassa) a condizioni di shaving (forza alta). In genere gli esperimenti di shaving vengono effettuati in modalità con-

tatto sia nella fase di imaging sia in quella di rimozione selettiva delle molecole. Una variante metodologica interessante consiste nel lavorare in modalità quantitative imaging (QI) nella fase di imaging. Il QI (https://www.jpk.com/products/atomic-force-microscopy/qi-mode) è una modalità di imaging basata sulla spettroscopia di forza: attraverso l'acquisizione di un ampio set di curve forza-distanza, consente di ricostruire la topografia del campione attraverso una mappa della posizione z della punta a forza applicata fissata. Poiché la punta viene avvicinata ed allontanata dalla superficie punto per punto durante la scansione, le forze laterali sono ridotte al minimo e si evitano problemi di trascinamento. Inoltre, la combinazione di imaging e spettroscopia di forza permette di ottenere dati quantitativi non solo relativi alle misure di spessore del film, ma anche sulle sue proprietà meccaniche. In particolare, l'analisi delle curve di forza consente la valutazione locale di parametri meccanici come l'adesione e il modulo di Young [37].

In *Fig. 3* vengono riportati i risultati di due esperimenti di shaving condotti su SAM di DNA tiolato su oro preparati a diversa forza ionica, bassa forza ionica (1 mM NaCl, *Fig. 1a-c*) ed alta forza ionica (1 M NaCl, *Fig. 1d-f*). Si tratta di un oligonucleotide formato da 22 basi avente un linker tiolato costituito da una catena di esantiolo. Si rimanda a Pinto *et al.* [38] per i dettagli sperimentali. Entrambi shaving ed imaging sono stati eseguiti nello stesso buffer utilizzato per il *self-assembly* del DNA. Mentre lo shaving è stato condotto in modalità hard tapping, l'imaging è stato effettuato sia in soft contact sia in QI.

Lo spessore del SAM può essere valutato dall'analisi delle immagini misurando la differenza di altezza tra la regione esterna corrispondente al SAM e la regione centrale, corrispondente all'oro, dove il SAM è stato rimosso. In *Fig. 3a,d* sono riportate le immagini in soft contact, mentre in *Fig. 3b,e* sono riportate le immagini in QI relative a shaving su SAM preparati a bassa forza ionica (sezione superiore, *Fig. 3a-c*) e ad alta forza ionica (sezione inferiore, *Fig. 3d-f*). Dai profili z delle immagini si può misurare lo spessore del SAM (*Fig. 3c,f*).

La sovrapposizione delle curve continue (relative alle immagini acquisite in soft contact) e tratteggiate (relative alle immagini acquisite in modalità QI) indica che la stima dell'altezza ottenuta in modalità contatto è leggermente inferiore al valore derivante dalle immagini in QI. Questa differenza si può spiegare tenendo conto delle differenze tra i due tipi di imaging: in modalità contatto, l'interazione continua tra il campione e la punta può provocare forze di trascinamento che possono

perturbare il monostrato, determinando una sottostima dell'altezza. È quindi necessario lavorare minimizzando per quanto possibile la forza di interazione (tipicamente 0.1 nN). In modalità QI, l'interazione intermittente punta-campione evita gli effetti di trascinamento e si traduce in una modalità operativa meno perturbativa. Un'analisi comparativa delle immagini ottenute in modalità soft contact e QI può pertanto aumentare l'affidabilità della misura dell'altezza del SAM. Vale la pena notare che l'accordo relativo tra i profili z ottenuti nelle modalità contatto e QI è leggermente migliore per i SAM preparati ad alta forza ionica, indicando che questi film sono meno vulnerabili alla perturbazione della punta AFM. Questa osservazione è in accordo con il fatto che i SAM depositati ad alta forza ionica sono più spessi, quindi verosimilmente più densi e compatti rispetto ai SAM preparati a bassa concentrazione di sale. Da un'analisi statistica dei profili z ottenuti su diversi campioni, risultano spessori di (1.0 ± 0.3) nm e (3.3 ± 0.5) nm per i SAM depositati, rispettivamente, a bassa e ad elevata forza ionica.

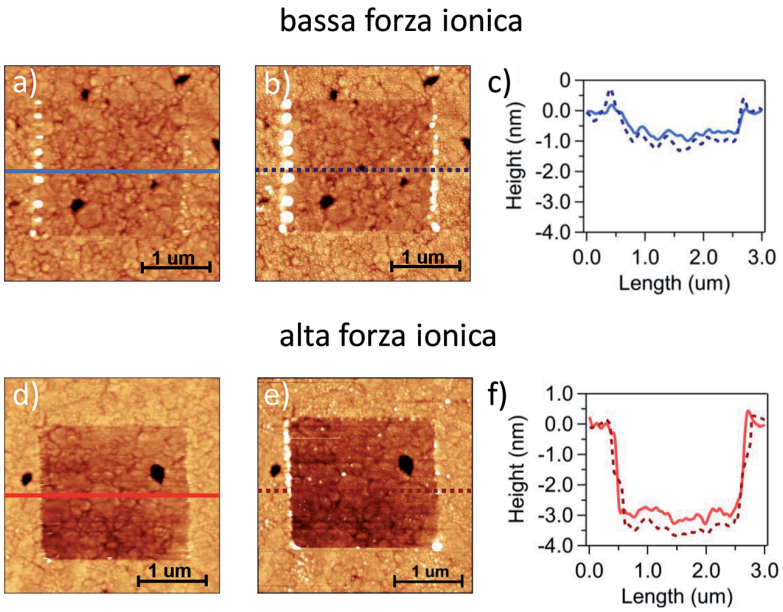


Fig. 3. Esperimenti di shaving su SAM di DNA preparati a bassa forza ionica (a-c) e ad alta forza ionica (d-f). Immagini AFM acquisite in modalità soft contact (a,d) e QI (b,e) (scala z: 9 nm). (c,f) profili z relativi alle immagini in contatto (linee continue) e in QI (linee tratteggiate). Modificata da [38].

Oltre all'informazione morfologica sullo spessore, l'analisi AFM può fornire informazioni anche sulle proprietà meccaniche del SAM. La *Fig. 4* riporta le mappe della deflessione laterale e del modulo di Young relativi agli esperimenti di shaving mostrati in *Fig. 3*. La deflessione laterale, misurata operando con l'AFM in modalità contatto, fornisce informazioni sulle proprietà tribologiche locali tra punta e campione. Ci si attende quindi di misurare diversi valori di deflessione laterale se la punta si muove su una regione dove è presente il SAM o se si muove sul substrato di oro. L'immagine in deflessione laterale (*Fig. 4a*) mostra una differenza molto bassa tra i valori di deflessione laterale nel caso dei SAM di DNA preparati a bassa forza ionica, mentre il contrasto è decisamente maggiore per i SAM preparati ad elevata forza ionica (*Fig. 4d*). Tali risultati possono essere interpretati sulla base delle misure di altezza soprariportate. I SAM preparati a bassa forza ionica hanno uno spessore molto ridotto, questo indica una densità molecolare molto bassa, verosimilmente legata ad una disposizione delle molecole quasi parallela alla superficie. Con un film di così bassa densità risulta difficile disaccoppiare le proprietà tribologiche del film da quelle del substrato sottostante. Al contrario, nel caso dei SAM depositati a forza ionica più elevata, che risultano più spessi e compatti, le immagini in deflessione laterale permettono di osservare una netta differenza tra SAM e substrato.

Ulteriori informazioni sulle proprietà meccaniche del campione possono essere ottenute dall'analisi in modalità QI. In particolare, l'analisi dei dati QI consente di valutare il modulo di Young del campione. Le *Fig. 4b,e* mostrano le mappe 2D dei valori del modulo di Young relativi agli esperimenti, rispettivamente, a bassa ed alta forza ionica. Come previsto, il substrato d'oro presenta un modulo di Young decisamente più elevato rispetto al film di DNA nel caso di SAM preparati ad alta forza ionica (*Fig. 4e*) mentre a bassa concentrazione di sale (*Fig. 4b*) non vi è una differenza significativa tra substrato e SAM. Quando presente, la differenza nei valori dei moduli di Young del film e del substrato può essere meglio apprezzata analizzando gli istogrammi delle mappe dei moduli di Young riportate nelle *Fig. 4c,f*. L'istogramma in *Fig. 4c* può essere riprodotto con un singolo profilo gaussiano centrato a 100 MPa, il che significa che nel caso di SAM depositati a bassa forza ionica non è possibile rilevare alcuna differenza significativa tra il modulo di Young del substrato nudo e del substrato coperto dal SAM. Viceversa, ad alta forza ionica, devono essere utilizzati due profili gaussiani per riprodurre l'inviluppo dell'istogramma (*Fig. 4f*). Dall'analisi

dei dati risulta una differenza di circa 40 Mpa tra il centro del profilo gaussiano che corrisponde al SAM (curva marrone, centrata a 57 MPa) e il centro del profilo gaussiano che corrisponde al substrato (curva gialla, centrata a 100 Mpa). I valori del modulo di Young del biofilm risultano, come previsto, inferiori a quelli del substrato d'oro.

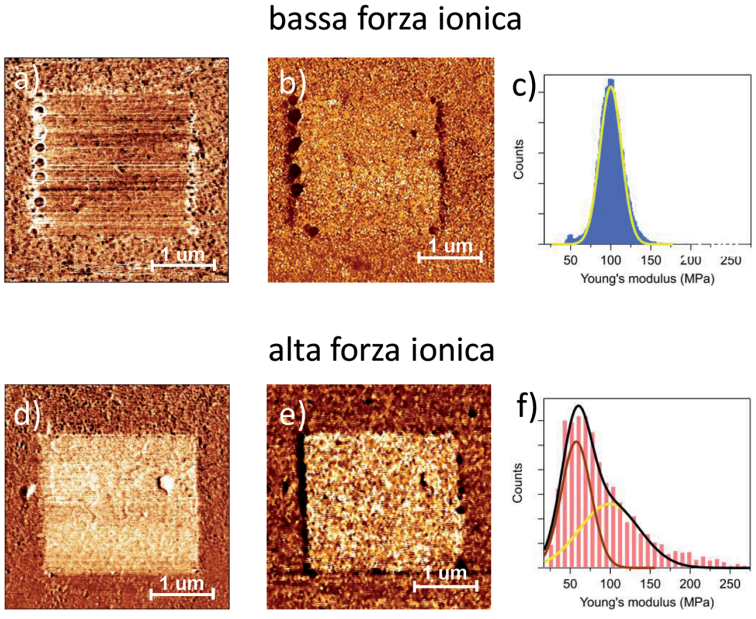


Fig. 4. Esperimenti di shaving su SAM di DNA preparati a bassa forza ionica (a-c) e ad alta forza ionica (d-f). (a,d) Immagini in deflessione laterale acquisite in modalità contatto (scala: 20 mV). (b,e) Mappe del modulo di Young ottenute dall'analisi delle curve di forza in QI (scala: 150 Mpa) e (c,f) relativi istogrammi. Modificata da [38].

Gli esperimenti di nanoshaving indicano che la forza ionica del tampone influenza fortemente la deposizione del DNA. A bassa forza ionica, a causa della repulsione tra i filamenti di DNA, carichi negativamente, si ottiene un SAM sottile, di bassa densità verosimilmente formato da molecole quasi sdraiate che assumono una conformazione prevalentemente allungata. Viceversa, ad alta forza ionica, la carica del DNA risulta schermata con una conseguente conformazione più ripiegata e una ridotta repulsione tra filamenti adiacenti. In queste condizioni si ottiene un SAM più spesso e con una copertura maggiore formata

da molecole quasi “in piedi”. Questa interpretazione è in accordo con l’osservazione, su SAM di densità molecolare fissata, di filamenti allungati a bassa forza ionica e filamenti ripiegati ad alta forza ionica [39].

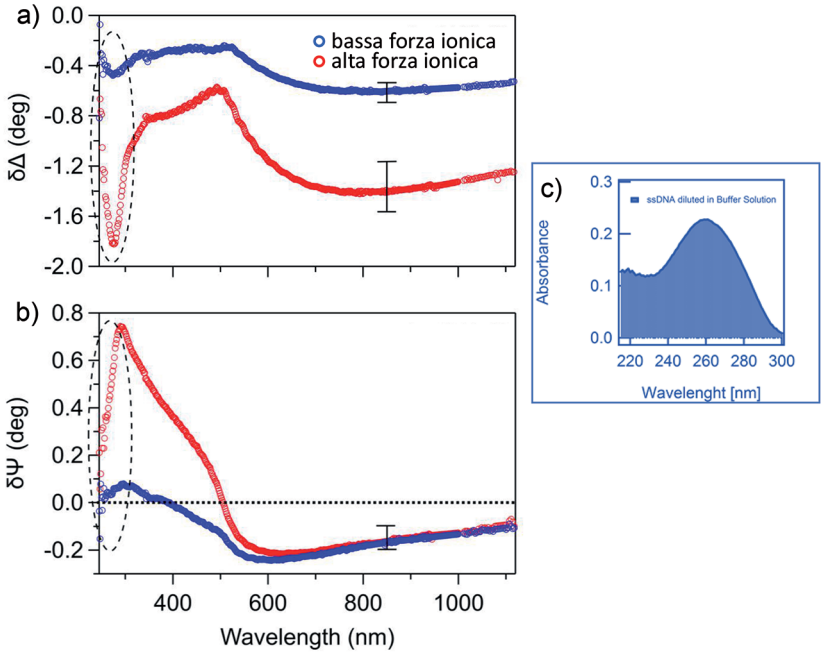


Fig. 5. (a,b) Spettri ellissometrici differenza relativi a SAM depositati a bassa forza ionica (curve blu) e ad alta forza ionica (curve rosse). Le regioni evidenziate indicano i minimi associati all’assorbimento molecolare del DNA. (c) Assorbimento UV-Vis di una soluzione di DNA. Modificata da [38].

Una conferma dei risultati AFM viene dall’analisi spettroellissometrica (SE). In *Fig. 5* sono riportati gli spettri differenza ottenuti sottraendo agli spettri acquisiti dopo la deposizione del SAM gli spettri del solo substrato. L’analisi degli spettri differenza consente di mettere in evidenza il contributo del SAM. Per una discussione più dettagliata di questi dati si rimanda alla ref. [38]. Una prima caratteristica spettrale interessante delle curve SE è la presenza di un minimo profondo intorno a 270 nm in δΔ (*Fig. 5a*) e di una corrispondente decrescita di δΨ intorno a 290 nm (*Fig. 5b*). Questi minimi sono ben definiti ad alta forza ionica (curve rosse), ma sono presenti, seppure più deboli, anche

a bassa forza ionica (curve blu). Considerando che l'assorbimento UV-Vis del DNA in soluzione presenta una banda di assorbimento situata a 260 nm (*Fig. 5c*) e che i film trasparenti presentano un comportamento molto diverso in questa regione [40], è ragionevole attribuire questa struttura spettrale all'assorbimento del DNA. Va inoltre osservato che i valori di $|\delta\Delta|$ nel NIR (lontano dalle risonanze molecolari) e i valori $\delta\Psi$ nel vicino UV sono decisamente maggiori per i SAM depositati ad elevata forza ionica. Come discusso in precedenti lavori [30,41], queste caratteristiche sono associate allo spessore ottico del film, una grandezza correlata all'indice di rifrazione e allo spessore del film, che risulta più elevata per i SAM depositati ad elevata forza ionica. I dati SE confermano quindi che la forza ionica del buffer gioca un ruolo significativo nella formazione del SAM. Va notato, inoltre, che per i film ultrasottili, come i monostrati analizzati in questo studio, l'indice di rifrazione e lo spessore sono parametri altamente correlati. Gli esperimenti di nanoshaving possono essere sfruttati per risolvere questa correlazione: assumendo come valori per lo spessore del SAM i risultati ottenuti dall'AFM, è infatti possibile ricavare una stima dell'indice di rifrazione attraverso un confronto tra dati ellissometrici e curve simulate ottenute a partire da un opportuno modello ottico del sistema. Questo tipo di analisi, per i cui dettagli si rimanda alla ref. [38], fornisce, per i SAM depositati ad elevata forza ionica, valori dell'indice di rifrazione intorno a 1.45 per λ=650 nm, valore in ragionevole accordo con la formazione di SAM compatti.

3. Conclusioni

L'accoppiamento di nanolitografia AFM e spettroellissometria ha permesso di caratterizzare le proprietà morfologiche, meccaniche ed ottiche di SAM di DNA e di valutarne la dipendenza dalla forza ionica della soluzione utilizzata per la deposizione dei SAM stessi. I SAM depositati ad elevata forza ionica risultano più spessi, densi e meno deformabili rispetto a quelli depositati a bassa forza ionica. In particolare, sui SAM densi, l'utilizzo dell'AFM in modalità QI ha consentito di rilevare una differenza nel modulo di Young tra SAM e substrato. Inoltre, l'analisi degli spettri ellissometrici ha permesso di rilevare l'assorbimento molecolare del DNA a livello di singolo monostrato.

BIBLIOGRAFIA

1. Bain, C.D.; Whitesides, G.M. Molecular-Level Control over Surface Order in Self-Assembled Monolayer Films of Thiols on Gold. *Science* **1988**, *240*, 62–63, doi:10.1126/science.240.4848.62.
2. Nuzzo, R.G.; Zegarski, B.R.; Dubois, L.H. Fundamental studies of the chemisorption of organosulfur compounds on gold(111). Implications for molecular *self-assembly* on gold surfaces. *J. Am. Chem. Soc.* **1987**, *109*, 733–740, doi:10.1021/ja00237a017.
3. Bain, C.D.; Troughton, E.B.; Tao, Y.T.; Evall, J.; Whitesides, G.M.; Nuzzo, R.G. Formation of monolayer films by the spontaneous assembly of organic thiols from solution onto gold. *J. Am. Chem. Soc.* **1989**, *111*, 321–335, doi:10.1021/ja00183a049.
4. Love, J.C.; Estroff, L.A.; Kriebel, J.K.; Nuzzo, R.G.; Whitesides, G.M. Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology. *Chem. Rev.* **2005**, *105*, 1103–1170, doi:10.1021/cr0300789.
5. Pujari, S.P.; Scheres, L.; Marcelis, A.T.M.; Zuilhof, H. Covalent Surface Modification of Oxide Surfaces. *Angew. Chem. Int. Ed.* **2014**, *53*, 6322–6356, doi:10. 1002/anie.201306709.
6. Flink, S.; Veggel, F.C.J.M. van; Reinhoudt, D.N. Sensor Functionalities in Self-Assembled Monolayers. *Advanced Materials* **2000**, *12*, 1315–1328, doi:10.1002/1521-4095(200009)12:18<1315::AID-ADMA1315>3.0.CO;2-K.
7. Ibau, C.; Md Arshad, M.K.; Gopinath, S.C.B.; Nuzaihan M.N, M.; M. Fathil, M.F.; Estrela, P. Gold interdigitated triple-microelectrodes for label-free prognosticative aptasensing of prostate cancer biomarker in serum. *Biosensors and Bioelectronics* **2019**, *136*, 118–127, doi:10.1016/j.bios.2019.04.048.
8. Yu, S.; Chen, T.; Zhang, Q.; Zhou, M.; Zhu, X. Application of DNA nanodevices for biosensing. *Analyst* **2020**, *145*, 3481–3489, doi:10.1039/D0AN00159G.
9. MacAskill, A.; Crawford, D.; Graham, D.; Faulds, K. DNA Sequence Detection Using Surface-Enhanced Resonance Raman Spectroscopy in a Homogeneous Multiplexed Assay. *Anal. Chem.* **2009**, *81*, 8134–8140, doi:10.1021/ac901361b.
10. Qian, S.; Lin, M.; Ji, W.; Yuan, H.; Zhang, Y.; Jing, Z.; Zhao, J. Boronic Acid Functionalized Au Nanoparticles for Selective MicroRNA Signal Amplification in Fiber-Optic Surface Plasmon Resonance Sensing System. *ACS Sensors* **2018**, 7.
11. Caneira, C.R.F.; Soares, R.R.G.; Pinto, I.F.; Mueller-Landau, H.S.; Azevedo, A.M.; Chu, V.; Conde, J.P. Development of a rapid bead-based microfluidic platform for DNA hybridization using single- and multi-mode interactions for probe immobilization. *Sensors and Actuators B: Chemical* **2019**, *286*, 328–336, doi:10.1016/j.snb.2019.01.133.
12. Teengam, P.; Siangproh, W.; Tuantranont, A.; Vilaivan, T.; Chailapakul, O.; Henry, C.S. Multiplex Paper-Based Colorimetric DNA Sensor Using Pyrrolidinyl Peptide Nucleic Acid-Induced AgNPs Aggregation for Detecting MERS-CoV, MTB, and HPV Oligonucleotides. *Anal. Chem.* **2017**, *89*, 5428–5435, doi:10.1021/acs.analchem.7b00255.
13. Kim, H.; Park, M.; Hwang, J.; Kim, J.H.; Chung, D.-R.; Lee, K.; Kang, M.

Development of Label-Free Colorimetric Assay for MERS-CoV Using Gold Nanoparticles. *ACS Sens.* **2019**, *4*, 1306–1312, doi:10.1021/acssensors.9b00175.
14. Wold, E.D.; McBride, R.; Axup, J.Y.; Kazane, S.A.; Smider, V.V. Antibody Microarrays Utilizing Site-Specific Antibody–Oligonucleotide Conjugates. *Bioconjugate Chem.* **2015**, *26*, 807–811, doi:10.1021/acs.bioconjchem.5b00111.
15. Yang, Y.; Wang, S.; Zhou, Z.; Zhang, R.; Shen, H.; Song, J.; Su, P.; Yang, Y. Enhanced reusability and activity: DNA directed immobilization of enzyme on polydopamine modified magnetic nanoparticles. *Biochemical Engineering Journal* **2018**, *137*, 108–115, doi:10.1016/j.bej.2018.05.019.
16. Wang, H.; Liu, M.; Bai, W.; Sun, H.; Li, Y.; Deng, H. A convenient electrogenerated chemiluminescence biosensing method for selective detection of 5-hydroxymethylcytosine in genomic DNA. *Sensors and Actuators B: Chemical* **2019**, *284*, 236–242, doi:10.1016/j.snb.2018.12.132.
17. Herne, T.M.; Tarlov, M.J. Characterization of DNA Probes Immobilized on Gold Surfaces. *Journal of the American Chemical Society* **1997**, *119*, 8916–8920, doi:10.1021/ja9719586.
18. Liu, D.; Bruckbauer, A.; Abell, C.; Balasubramanian, S.; Kang, D.-J.; Klenerman, D.; Zhou, D. A Reversible pH-Driven DNA Nanoswitch Array. *Journal of the American Chemical Society* **2006**, *128*, 2067–2071, doi:10.1021/ja0568300.
19. Guiducci, C.; Stagni, C.; Fischetti, A.; Mastromatteo, U.; Benini, L.; Riccoricco, B. Microelectrodes on a Silicon Chip for Label-Free Capacitive DNA Sensing. *IEEE Sensors J.* **2006**, *6*, 1084–1093, doi:10.1109/JSEN.2006.877979.
20. Drozd, M.; Pietrzak, M.D.; Malinowska, E. SPRi-Based Biosensing Platforms for Detection of Specific DNA Sequences Using Thiolate and Dithiocarbamate Assemblies. *Front Chem* **2018**, *6*, doi:10.3389/fchem.2018.00173.
21. Li, Z.; Niu, T.; Zhang, Z.; Feng, G.; Bi, S. Effect of monovalent cations (Li+, Na+, K+, Cs+) on *self-assembly* of thiol-modified double-stranded and single-stranded DNA on gold electrode. *Analyst* **2012**, *137*, 1680–1691, doi:10.1039/c2an15716k.
22. Guilbaud, S.; Salomé, L.; Destainville, N.; Manghi, M.; Tardin, C. Dependence of DNA Persistence Length on Ionic Strength and Ion Type. *Phys. Rev. Lett.* **2019**, *122*, 028102, doi:10.1103/PhysRevLett.122.028102.
23. Moses, S.; Brewer, S.H.; Lowe, L.B.; Lappi, S.E.; Gilvey, L.B.G.; Sauthier, M.; Tenent, R.C.; Feldheim, D.L.; Franzen, S. Characterization of Single- and Double-Stranded DNA on Gold Surfaces. *Langmuir* **2004**, *20*, 11134–11140, doi:10.1021/la0492815.
24. Keighley, S.D.; Li, P.; Estrela, P.; Migliorato, P. Optimization of DNA immobilization on gold electrodes for label-free detection by electrochemical impedance spectroscopy. *Biosensors and Bioelectronics* **2008**, *23*, 1291–1297, doi:10.1016/j.bios.2007.11.012.
25. Wolf, L.K.; Gao, Y.; Georgiadis, R.M. Sequence-Dependent DNA Immobilization: Specific versus Nonspecific Contributions. *Langmuir* **2004**, *20*, 3357–3361, doi:10.1021/la036125+.
26. Petrovykh, D.Y.; Kimura-Suda, H.; Whitman, L.J.; Tarlov, M.J. Quantitative Analysis and Characterization of DNA Immobilized on Gold. *Journal of the American Chemical Society* **2003**, *125*, 5219–5226, doi:10.1021/ja029450c.

27. Howell, C.; Hamoudi, H.; Zharnikov, M. Thymine/adenine diblock-oligonucleotide monolayers and hybrid brushes on gold: a spectroscopic study. *Biointerphases* **2013**, *8*, 6, doi:10.1186/1559-4106-8-6.
28. Charrier, E.E.; Pogoda, K.; Wells, R.G.; Janmey, P.A. Control of cell morphology and differentiation by substrates with independently tunable elasticity and viscous dissipation. *Nat Commun* **2018**, *9*, 449, doi:10.1038/s41467-018-02906-9.
29. Solano, I.; Parisse, P.; Gramazio, F.; Cavalleri, O.; Bracco, G.; Castronovo, M.; Casalis, L.; Canepa, M. Spectroscopic ellipsometry meets AFM nanolithography: about hydration of bio-inert oligo(ethylene glycol)-terminated self assembled monolayers on gold. *Physical Chemistry Chemical Physics* **2015**, *17*, 28774–28781, doi:10.1039/C5CP04028K.
30. Pinto, G.; Parisse, P.; Solano, I.; Canepa, P.; Canepa, M.; Casalis, L.; Cavalleri, O. Functionalizing gold with single strand DNA: novel insight into optical properties *via* combined spectroscopic ellipsometry and nanolithography measurements. *Soft Matter* **2019**, *15*, 2463–2468, doi:10.1039/C8SM02589D.
31. Xu, S.; Liu, G.-Y. Nanometer-Scale Fabrication by Simultaneous Nanoshaving and Molecular *Self-Assembly*. *Langmuir* **1997**, *13*, 127–129, doi:10.1021/la962029f.
32. Liu, M.; Amro, N.A.; Chow, C.S.; Liu, G. Production of Nanostructures of DNA on Surfaces. *Nano Letters* **2002**, *2*, 863–867, doi:10.1021/nl025626x.
33. Liu, S.; Maoz, R.; Schmid, G.; Sagiv, J. Template Guided *Self-Assembly* of [Au55] Clusters on Nanolithographically Defined Monolayer Patterns. *Nano Lett.* **2002**, *2*, 1055–1060, doi:10.1021/nl025659c.
34. Unruh, D.A.; Malduin, C.; Pastine, S.J.; Rolandi, M.; Fréchet, J.M.J. Bifunctional Patterning of Mixed Monolayer Surfaces Using Scanning Probe Lithography for Multiplexed Directed Assembly. *Journal of the American Chemical Society* **2010**, *132*, 6890–6891, doi:10.1021/ja101627e.
35. Ambrosetti, E.; Paoletti, P.; Bosco, A.; Parisse, P.; Scaini, D.; Tagliabue, E.; de Marco, A.; Casalis, L. Quantification of Circulating Cancer Biomarkers via Sensitive Topographic Measurements on Single Binder Nanoarrays. *ACS Omega* **2017**, *2*, 2618–2629, doi:10.1021/acsomega.7b00284.
36. Solano, I.; Parisse, P.; Gramazio, F.; Ianeselli, L.; Medagli, B.; Cavalleri, O.; Casalis, L.; Canepa, M. Atomic Force Microscopy and Spectroscopic Ellipsometry combined analysis of Small Ubiquitin-like Modifier adsorption on functional monolayers. *Applied Surface Science* **2017**, *421*, 722–727, doi:10.1016/j.apsusc.2016.10.195.
37. Dufrêne, Y.F.; Martínez-Martín, D.; Medalsy, I.; Alsteens, D.; Müller, D.J. Multiparametric imaging of biological systems by force-distance curve–based AFM. *Nat Methods* **2013**, *10*, 847–854, doi:10.1038/nmeth.2602.
38. Pinto, G.; Canepa, P.; Canale, C.; Canepa, M.; Cavalleri, O. Morphological and Mechanical Characterization of DNA SAMs Combining Nanolithography with AFM and Optical Methods. *Materials* **2020**, *13*, 2888, doi:10.3390/ma13132888.
39. Gil, P.S.; Lacks, D.J.; Parisse, P.; Casalis, L.; Nkoua Ngavouka, M.D. Single-stranded DNA oligomer brush structure is dominated by intramolecular interactions mediated by the ion environment. *Soft Matter* **2018**, *14*, 9675–9680, doi:10.1039/C8SM01743C.
40. Prato, M.; Moroni, R.; Bisio, F.; Rolandi, R.; Mattera, L.; Cavalleri, O.; Canepa, M.

Optical Characterization of Thiolate Self-Assembled Monolayers on Au(111). *The Journal of Physical Chemistry C* **2008**, *112*, 3899–3906, doi:10.1021/jp711194s.

41. Parisse, P.; Solano, I.; Magnozzi, M.; Bisio, F.; Casalis, L.; Cavalleri, O.; Canepa, M. Thickness and Beyond. Exploiting Spectroscopic Ellipsometry and Atomic Force Nanolithography for the Investigation of Ultrathin Interfaces of Biologic Interest. In *Ellipsometry of Functional Organic Surfaces and Films*; Hinrichs, K., Eichhorn, K.-J., Eds.; Springer International Publishing: Cham, 2018; Vol. 52, pp. 63–93 ISBN 978-3-319-75894-7.

Istituto Lombardo (Rend. Scienze) 153, 147-190 (2019)

COMPRENDERE LE BASI MOLECOLARI DELLA DISABILITÀ INTELLETTIVA ATTRAVERSO IL MODELLO DEI NEURONI DI CELLULE STAMINALI PLURIPOTENTI INDOTTE

Nota del s.c. LIDIA LARIZZA (*)

(Adunanza del 24 ottobre 2019)

SUNTO. – Il punto di partenza di questa nota è il Premio Nobel per la Medicina assegnato nel 2012 all'inglese Sir John Gurdon e al giapponeses Shinya Yamanaka per la scoperta della "riprogrammazione cellulare", il processo per cui una cellula differenziata (matura) di un organismo può in determinate condizioni acquisire proprietà di cellula staminale pluripotente in grado di sviluppare tutte le cellule e i tessuti dell'organismo di provenienza. Oltre a rivoluzionare la nostra visione della biologia dello sviluppo tale scoperta ha permesso di generare *in vitro* cellule staminali pluripotenti (iPSC) da qualsiasi individuo, sano o malato e poi differenziarle come desiderato permettendo di modellare malattie che colpiscono tessuti e sistemi prima inaccessibili, come ad esempio il sistema nervoso ("brain in a dish"). In generale i modelli iPSC-derivati di malattie umane, genetiche o su base genetica, consentono di esplorarne i meccanismi patogenetici complementandosi ai modelli *in vivo* finora principalmente utilizzati (topo, ratto, zebrafish, drosophila) e costituiscono una piattaforma per lo screening di migliaia di piccole molecole o farmaci candidati per una data patologia selezionandoli in base alla capacità di far revertire alterazioni morfologiche, molecolari e funzionali registrate nelle cellule differenziate paziente-specifiche. Inoltre i modelli iPSC-derivati promettono potenzialità di terapia rigenerativa cellulo-mediata per malattie ad alta prevalenza benchè alcune problematiche quali i numeri di cellule necessarie e il loro corretto posizionamento nell'organo danneggiato rendano laboriosi gli sforzi in corso.

Abbiamo generato il primo modello neuronale iPSC-derivato per pazienti con Sindrome di Rubinstein-Taybi, un disordine mendeliano del neurosviluppo caratterizzato da dismorfismi facciali, ritardo della crescita e del linguaggio, displasia scheletrica,

(*) IRCCS Istituto Auxologico Italiano, Milano, Italy.
E-mail: l.larizza@auxologico.it ; lidia.larizza@unimi.it

malformazioni di vario tipo e deficit cognitivo spesso associato a disordine del comportamento. La sindrome risulta da mutazioni di uno dei geni omologhi *CREBBP* ed *EP300* che codificano proteine che agiscono come "epigenetic writers" essendo coattivatori della trascrizione genica con attività acetiltransferasica diretta sia a istoni della cromatina che a proteine non istoniche. Benchè le mutazioni in questi geni siano irreversibili, le modificazioni epigenetiche determinate dal loro difetto sono reversibili aprendo potenziali opzioni terapeutiche postnatali volte ad attenuare il deficit cognitivo dei pazienti. Sono stati modellati sei pazienti RSTS, 4 *CREBBP*- e 2 *EP300*-mutati in parallelo a 4 individui sani: viene illustrato il work flow dalla riprogrammazione delle cellule mononucleate del sangue periferico alla generazione di iPSC, la loro caratterizzazione morfologica, immunoistochimica, citofluorimetrica, il loro potenziale a generare i tre foglietti embrionali e il differenziamento di linee iPSC validate per la stabilità genomica agli stadi di rosette neurali, progenitori neurali e neuroni corticali maturi. Benché i neuroni RSTS non presentino alla caratterizzazione immunoistochimica differenze evidenti nell' espressione di marcatori stadio-specifici rispetto a neuroni di controllo, l'analisi a due timepoints, ha rivelato biomarcatori distintivi. Allo stadio di neuroni in differenziamento (42 giorni), ancora a bassa densità in coltura e quindi passibili di analisi morfologiche precise, i neuroni di 4 su 5 pazienti RSTS presentano ridotta lunghezza e aumentato numero di "branches" (dendriti) rispetto a neuroni di controllo: la differenza è statisticamente significativa e l'unica eccezione è rappresentata dalla paziente con la disabilità intellettiva più lieve. Neuroni postmitotici (>70 giorni) RSTS caratterizzati con la tecnica di "voltage clamp" nostrano l'espressione di correnti di sodio e potassio analogamente ai controlli, e si dimostrano in grado di generare con la tecnica del "current clamp", seppur in percentuale inferiore al controllo, potenziali d'azione in risposta a stimoli depolarizzanti. Tuttavia i neuroni di tutti i pazienti risultano "ipoeccitabili" rispetto a neuroni di controllo, hanno cioè bisogno di uno stimolo di intensità superiore per raggiungere la soglia necessaria per generare potenziali d'azione.

In relazione all'identificazione nei neuroni RSTS di biomarcatori morfologici ed elettrofisiologici che li differenziano dai neuroni di individui sani, sono stati avviati esperimenti di trattamento dei neuroni del paziente RSTS con il fenotipo neurocognitivo più grave (P149) con un inibitore delle istondeacetilasi, la tricostatina A, dimostratosi in grado di migliorare il fenotipo neurologico in modelli murini *cbp+/-* . Risultati preliminari con TSA 20 nM supplementata a due linee iPSC-.derivate di progenitori neurali di P149 hanno mostrato significativa, anche se parziale correzione di parametri morfologici alterati nei neuroni non trattati. Analogamente la somministrazione cronica per 5 settimane di TSA 0,2 nM alle colture di progenitori neurali di P149 ha mostrato significativo miglioramento delle proprietà elettrofisiologiche dei neuroni maturi (>70 giorni) di entrambe le linee di P149, ivi compresa la capacità di generare potenziali d'azione che non é significativamente diversa tra il controllo e il paziente trattato. La validazione di questi risultati mediante ulteriori sperimentazioni *in vitro* su neuroni di pazienti e controlli permetterà di integrare le modificazioni epigenetiche causate dal difetto di CBP/p300 nel contesto appropriato di omeostasi neuronale con un ulteriore passo in avanti verso trials clinici su pazienti RSTS in età pediatrica.

***

ABSTRACT. – In 2012 the Nobel Prize for Medicine was awarded to Sir John Gurdon and Shinya Yamanaka for their joined discovery that mature, specialized cells can be repro-

grammed to become immature pluripotent cells capable of developing into all different tissues of the body. This seminal breakthrough changed our vision of developmental biology and made possible to generate in vitro from somatic cells of any healthy or affected individual Induced Pluripotent Stem Cells (iPSCs) and differentiate them into virtually all types of cells. Scientists around the world were provided with the tool of modelling human diseases, even those involving inaccessible tissues such as the nervous system, to study the pathological process. iPSCs also revealed an enormous potential for pharmaceutical and clinical applications as on these cells a myriad of small molecules or candidate drugs can be tested some of which hopefully will become new effective medicines for intractable diseases. We have generated the first iPSC-derived neuronal model for the Rubinstein-Taybi syndrome (RSTS), a mendelian neurodevelopmental disorder, characterized by facial dysmorphisms, growth and speech delay, skeletal dysplasia and intellectual disability often associated to behavior disorder. RSTS is caused by mutation in the homologous *CREBBP* and *EP300* genes encoding proteins acting as chromatin regulators through their intertwined activity of transcriptional co-activators and acetyltransferases (KATs) on histone and non histone proteins. I summarize herein the workflow used to generate iPSC-derived neurons (iNeurons) from six RSTS patients with a variable cognitive impairment in parallel to iNeurons from four healthy controls. Immunohistochemical characterization of samples at the stages of iPSCs, neural rosettes, neural progenitors and post-mitotic cortical neurons did not reveal gross alterations in the expression of stage-specific differentiation markers across patients' or between patients' and controls' neuronal cultures. Conversely, altered morphology of patients' differentiating neurons, showing reduced branch length and increased branch number and hypoexcitability of differentiated neurons emerged as relevant "disease" biomarkers. Both the anomalous neuronal morphology and the impaired electrophysiological performance varied across iNeurons from different RSTS patients. possibly reflecting cognitive and behavioural impairment of the donor patients. To validate the identified morpho-functional markers we performed further studies on iNeurons from the RSTS patient with the most severe intellectual disability by using trichostatin A (TSA), an epidrug that by inhibiting histonedeacetylases (HDAcs), the KATs opposing enzymes, has been successfully applied in different *cbp+/cbp-* mouse models with improvement of some neurological abnormalities. Short acute and chronic TSA treatment of two independent neural progenitor lines from the selected RSTS patient determined in both replicas a consistent reversal of a few morphological abnormalities and a significant rescue of the defective electrophysiological performance, highlighting the potential postnatal intervention of epidrugs to ameliorate the cognitive impairment of RSTS patients.

La riprogrammazione cellulare

Il punto di partenza di questa nota è la scoperta della **riprogrammazione cellulare**, il fenomeno per cui cellule differenziate (mature) di qualsiasi tessuto di un organismo possono ritornare ad uno stadio indifferenziato immaturo "pluripotente" in grado di sviluppare tutti i tessuti di quell'organismo (*Fig. 1*). Questa scoperta ha rivoluzionato la

nostra comprensione di come le cellule e gli organismi si sviluppano, ha avviato applicazioni importanti per la comprensione dell'eziologia di patologie genetiche e su base genetica e ha dischiuso ricadute terapeutiche.

Per questa scoperta nel 2012 il Premio Nobel per la Fisiologia e la Medicina è stato assegnato congiuntamente all'inglese **Sir John Gurdon** e al giapponese **Shinya Yamanaka**: Gurdon ha avuto l'intuizione geniale, che sfatava il dogma contemporaneo della biologia dello sviluppo, che il differenziamento cellulare **non** è irreversibile; Yamanaka più di 40 anni dopo ne ha fornito la dimostrazione sperimentale in un contesto scientifico notevolmente avanzato grazie alla ricerca sulle cellule staminali embrionali caratterizzate dall'inglese Martin Evans, Nobel per la Medicina nel 2007 (*Fig. 2*).

Nel dettaglio, nel 1962 John Gurdon scoprì che la specializzazione cellulare è reversibile: in un classico esperimento ha sostituito il nucleo di una cellula uovo di rana (*Xenopus laevis*), enucleata con radiazioni, con quello di una cellula dell'epitelio intestinale di rana. La cellula uovo così "modificata" si è sviluppata in un girino e quindi in una rana normale, dimostrando che il DNA della cellula matura dell'epitelio intestinale conserva tutta l'informazione necessaria a sviluppare tutte le cellule di quell'organismo (*Gurdon, 1962*). La scoperta seminale di Gurdon è stata inizialmente accolta con scetticismo, ma iniziò ad essere accettata perché le tecniche di "trasferimento nucleare" aprirono la strada alla clonazione dei mammiferi, a partire dalla pecora Dolly nel '97 da parte del gruppo dell'inglese Jan Wilnut, a quella di numerosi altri mammiferi, soprattutto bovini, fino ai primati. Dolly fu la prima clonazione riuscita di un mammifero adulto. Come tutti gli animali clonati Dolly è nata in provetta dal trasferimento del nucleo di una cellula di ghiandola mammaria di pecora nella cellula uovo enucleata di un'altra pecora: una volta che il normale sviluppo è stato confermato, l'embrione, che alla fine sarebbe diventato Dolly, è stato trasferito in una madre surrogata. Tecnicamente quindi Dolly ha avuto tre madri!

Nonostante l'estesa ricaduta scientifica e biotecnologica degli esperimenti di Gurdon rimanevano ancora riserve sulle manipolazioni con pipette per la rimozione del nucleo e la sua introduzione in altre cellule: sarebbe mai stato possibile riprogrammare una cellula "intatta" in una cellula staminale pluripotente?

Nel 2006 Shinia Yamanaka è stato in grado di dare una risposta a questa domanda nel contesto della sua ricerca focalizzata sulle cellule staminali pluripotenti isolate dall'embrione e coltivate in laboratorio. Tali cellule erano state inizialmente caratterizzate da M. Evans e il gruppo di Yamanaka cercava di identificare i geni in grado di mantenerle nello stato di pluripotenza. A seguito dell'identificazione di parecchi di tali geni testava quale fosse in grado di riprogrammare cellule mature a cellule staminali pluripotenti. Introducendo diverse combinazioni geniche in cellule mature del tessuto connettivo di topo (fibroblasti) e analizzando i risultati al microscopio, trovò alla fine una combinazione di soli quattro geni (OCT4, SOX2, KLF4 e c-MYC) che induceva la riprogrammazione dei fibroblasti a cellule staminali. Le cellule staminali pluripotenti indotte (iPSC) erano poi in grado di differenziarsi in tutti i tipi di cellule mature del topo, dai fibroblasti ai neuroni alle cellule dell'epitelio intestinale (*Takahashi & Yamanaka, 2006*). La dimostrazione l'anno succesivo che cellule mature "intatte" possono essere riprogrammate per i fibroblasti umani *(Takahashi et al., 2007*) è stata immediatamente recepita come scoperta con grandi potenzialità nella medicina clinica.

Le scoperte di Gurdon e Yamanaka hanno profondamente modificato i concetti cardine della biologia dello sviluppo quali il differenziamento terminale considerato processo "irreversibile". Le iPSC sono simili alle cellule staminali embrionali (derivanti da embrioni allo stadio di blastocisti) nella morfologia, proliferazione, antigeni di superficie e geni di pluripotenza. Come le cellule staminali embrionali sono in grado di differenziarsi nei tipi cellulari dei tre foglietti (ecto, meso e endodermico) *in vitro* e di generare teratomi *in vivo*. Hanno però il grande vantaggio di non creare tutti i problemi etico-legali delle cellule staminali embrionali per la loro derivazione da qualsiasi tipo di cellule somatiche di individui sani e malati.

Modelli *in vitro* di malattie umane

Riprogrammando cellule umane gli scienziati hanno creato nuove opportunità per studiare malattie generando modelli *in vitro* malattia- e paziente-specifici a partire da vari tipi di cellule "periferiche" (sangue, cute, urina, etc). Ciò si è rivelato particolarmente utile per disordini di

sistemi ed organi difficilmente accessibili quali ad esempio il Sistema Nervoso Centrale, per il quale il tessuto umano fonte della maggior parte della sperimentazione é stato il cervello "post mortem" che, a parte le problematiche di accessibilità, etiche e politiche, non permette lo studio del periodo fetale dello sviluppo cerebrale, mentre è noto che la maggior parte della formazione del cervello umano avviene in epoca prenatale. Naturalmente modelli animali hanno supplito a tale carenza, ma solo parzialmente per le differente organizzazione del cervello nel topo e nell'uomo.

Modellare malattie umane *in vitro* ha permesso di decodificarne i **meccanismi eziopatogenetici**: nel caso dei disordini del neurosviluppo (*Zhao & Bhattacharyya, 2018*) in cui rientra la sindrome di Rubinstein-Taybi che tratterò a breve, e di malattie neurodegenerative il confronto dei neuroni *in vitro* di pazienti e individui sani a tempi diversi di differenziamento ha evidenziato "biomarcatori" morfologici, elettrofisiologici e molecolari utilizzabili per la terapia. I modelli *in vitro* di cellule differenziate da iPSC offrono infatti la piattaforma per test ad alta processività di migliaia di prodotti naturali e chimici o per lo screening di farmaci già noti come candidati per una data patologia: sia nel primo che nel secondo caso le alterazioni delle cellule paziente-specifiche devono revertire dopo l'esposizione al farmaco *in vitro* per portare a studi preclinici *in vivo* su modelli animali e trials clinici su pazienti (*Avior et al., 2016*)(*Fig. 3*).

Le iPSC hanno alimentato grandi speranze anche nel campo della **medicina rigenerativa**: la possibilità di indurre il differenziamento nella maggior parte dei tipi cellulari di un organismo (neuroni, cellule pancreatiche, cardiache, epatiche etc) può essere infatti sfruttata nella rigenerazione/riparazione di tessuti o organi danneggiati (**terapia rigenerativa cellulo-mediata**). Tuttavia, benché molti sforzi siano stati fatti e siano tuttora in corso per malattie ad alta prevalenza, quali ad esempio la maculopatia degenerativa dovuta all'età, il Parkinson, il diabete, le lesioni traumatiche del midollo spinale, i costi dovuti al numero astronomico di cellule da differenziare *in vitro* paziente-specifiche, che come tali quindi non creerebbero alcun rigetto immunitario, ma potrebbero acquisire mutazioni nel corso del lungo periodo di coltura e soprattutto il posizionamento corretto nell'organo danneggiato, non hanno portato finora a risultati incoraggianti, nonostante trials clinici di fase 1 e 2 siano in corso per le patologie sopramenzionate (*Trounson & DeWitt, 2016*).

Generazione del modello neuronale per la sindrome con disabilità intellettiva Rubinstein-Taybi

Abbiamo generato il modello iPSC neuronale per una sindrome genetica a trasmissione autosomico dominante, la sindrome di Rubinstein-Taybi (RSTS), caratterizzata da ritardo di crescita e psicomotorio, dismorfismi facciali, displasia scheletrica e disabilità intellettiva, definita come difetto di funzioni cognitive e di comportamento che si verifica durante lo sviluppo (*Fig. 4*) (*Hennekam, 2006; Larizza & Finelli, 2019*). L'incidenza é 1:125.000 nati, ma sale a 1:300 in pazienti istituzionalizzati con ritardo mentale. RSTS è causata da mutazioni o nel gene *CREBBP* o nel suo omologo *EP300* (*Spena et al., 2015; Negri et al., 2016*) che codificano le proteine multifunzionali CBP e p300, coattivatori della trascrizione genica con attività K (lisina) acetiltransferasica (KAT) diretta agli istoni della cromatina e a proteine nucleari non istoniche (*Figg. 5*, *6*, *7*) (*Sheikh e Akhtar, 2019*). La cromatina è il complesso del DNA che si avvolge attorno a ottameri di istoni: la sequenza delle basi del DNA attraverso il codice genetico detta **l'informazione genetica** che è stabile e irreversibile, mentre modificazioni dei gruppi metilici della citosina (una delle quattro basi del DNA), modificazioni post-traduzionali degli istoni, quali acetilazione, metilazione etc. e modificazioni opposte (deacetilazione, demetilazione etc), varianti istoniche e piccoli RNA non codificanti costituiscono **l'informazione epigenetica**, che è dinamica e reversibile (*Figg. 8* e *9*). Le modificazioni epigenetiche regolano la struttura tridimensionale della cromatina rendendola più o meno adatta alla trascrizione genica e sono quindi alla base dei programmi genetici tessuto-specifici. Le proteine CBP e p300 introducendo il gruppo acetilico donato dall'acetilcoenzima A nelle code di lisina di istoni e fattori di trascrizione agiscono come "epigenetic writers": il marchio epigenetico viene rimosso da enzimi con attività opposta, le iston-deacetilasi che sono "epigenetic erasers". Esiste una bilancia tra acetilazione e deacetilazione che è molto importante, specie per l'omeostasi neuronale: pertanto mutazioni di *CREBBP* o *EP300* e difetto di CBP o p300 spostano l'equilibrio a favore della deacetilazione e uno stato di cromatina "chiusa" non compatibile alla trascrizione. RSTS è una cromatinopatia: patologia in cui è difettoso per mutazione il gene per una

proteina componente dell'apparato epigenetico della cellula. I principali attori della macchina epigenetica sono oltre ai "writers" e gli "erasers" i "readers", proteine in grado di leggere il marchio epigenetico e i "remodellers", enzimi che usano l'energia dell'ATP per rimodellare la cromatina favorendo interazioni DNA-proteine (*Fig. 10*). Il gruppo dei disordini dell'apparato epigenetico, definito nel 2015 (*Bijornsson, 2015*) e in progressiva espansione (*Fahrner et al., 2019*) comprende sindromi mendeliane che condividono disabilità intellettiva e difetti della crescita (*Fig. 11*) sulle quali si è concentrata l'attenzione per la potenziale applicazione postnatale di terapie dirette all' attenuazione del deficit cognitivo. Difatti, benché la mutazione di un gene "epigenetico" sia irreversibile, i cambiamenti epigenetici che ne conseguono possono essere contrastati adottando farmaci che ripristinino lo stato di equilibrio alterato.

Al fine di acquisire una migliore comprensione dei meccanismi alla base del deficit cognitivo di pazienti RSTS abbiamo generato *in vitro* un modello di neuroni iPSC-derivati riprogrammando cellule del sangue periferico di 6 pazienti RSTS, 4 *CREBBP*- e 2 *EP300*-mutati manifestanti un diverso grado disabilità intellettiva associato in tre a disordine del comportamento: in parallelo sono state riprogrammate cellule del sangue periferico di 4 individui sani (*Fig. 12*). Il workflow della riprogrammazione e del differenziamento neuronale è dettagliato in *Fig. 13*.

Dalle iPSC paziente-specifiche ai neuroni: caratterizzazione immunoistochimica

Come mostrato nel workflow (*Fig. 13*) le iPSC sono generate dalle cellule mononucleate del sangue periferico di pazienti RSTS e individui sani utilizzando dopo 9 giorni di arricchimento delle cellule da riprogrammare (eritroblasti) il metodo del virus Sendai (SeV) raccomandato perché non integrativo (il virus non si integra nel genoma delle cellule riprogrammate dalle quali progressivamente scompare). Particelle di SeV trasducono nelle cellule bersaglio RNA per i 4 fattori di Yamanaka; le cellule trasdotte sono quindi piastrate su fibroblasti embrionali di topo (MEFs) irradiati in terreno per cellule staminali

embrionali. A partire dal 20 giorno le colonie emergenti sono prelevate e trasferite in nuove piastre con "feeder layer" per essere poi testate per la morfologia, l'espressione qualitativa (mediante immunofluorescenza) e quantitativa (mediante citofluorimetria) degli antigeni di staminalità (OCT4, SSEA4, TRA-1-60) e la capacità di differenziarsi nei tre foglietti dell'embrione, ectoderma, mesoderma e endoderma (*Fig. 14*). Le colonie iPSC (almeno 3 per ciascun donatore) sono caratterizzate mediante analisi del cariotipo, ibridazione genomica comparativa su array (aCGH) e sequenziamento Sanger del DNA (*Fig. 15*). I cloni iPSC che superano il controllo della stabilità genomica usando come riferimento le cellule somatiche (sangue periferico) del donatore sono poi differenziati in neuroni corticali secondo il protocollo in monostrato che utilizza "Neurobasal Medium" supplementato con noggina. Dopo circa due settimane il differenziamento neurale è marcato dalla comparsa delle tipiche "rosette" che sono molto simili a sezioni trasversali del tubo neurale durante lo sviluppo del sistema nervoso *in vivo* (*Fig. 16*). Allo stadio di rosette la maggior parte delle cellule di pazienti RSTS così come di individui sani mostra positività all'immunofluorescenza per i marcatori neuroectodermici Nestina e PAX6 (*Fig. 17*), confermata dal rilievo degli stessi marcatori mediante analisi di espressione con RT-PCR (Reverse Transcription PCR). Con taglio manuale e trasferimento in nuove piastre le rosette sono poi indotte a differenziarsi in Progenitori Neurali (NPCs) (35esimo giorno) mediante sostituzione al 28esimo giorno del terreno di base con "Neural Differentiation Medium". I progenitori neurali proliferano e sono ulteriormente indotti al differenziamento terminale, per cui solo aliquote di NPCs al 35 giorno sono stoccabili in azoto liquido e riscongelabili per futuri esperimenti. A 42 giorni i neuroni sia di controlli che di pazienti esprimono i marcatori panneuronali MAP2 (microtubule-associated protein 2) e TUJ1 (betaIII-tubulin) (*Fig. 18*). Sui neuroni a questo stadio è effettuata l'analisi morfologica (vedi oltre). Dopo prolungato mantenimento in coltura (>70 giorni) i neuroni di pazienti RSTS sviluppano come quelli di controlli una intricata rete di cellule arricchita in neuroni corticali, positivi per l'espressione di TBR-1, regolatore master dello sviluppo corticale, CUX-1 necessario per la formazione dei dendriti (*Fig. 19*), e il marcatore pre-sinaptico sinaptofisina (SYP) (*Fig. 20*). La maggior parte dei neu-

roni maturi di pazienti RSTS come di controlli sono positivi per il marcatore glutamatergico vGLUT (*Fig. 21*), tipico dei neuroni eccitatori, mentre pochi lo sono per il marker dei neuroni inibitori GAD65/67 (*Fig. 21*). Sotto il profilo qualitativo non sono evidenti differenze significative nel differenziamento *in vitro* di neuroni RSTS-specifici rispetto a neuroni di controllo. Sui neuroni posmitotici (>70 giorni) è effettuata l'analisi elettrofisiologica (vedi oltre) e saggi molecolari che possono comprendere l'analisi dell'intero trascrittoma.

ANALISI DELLA MORFOLOGIA DI NEURONI (42 GIORNI) DI RSTS E CONTROLLI

Ci siamo avvalsi dello stadio di neuroni giovani (42-45 giorni) ancora in coltura a bassa densità per stabilire se alcune caratteristiche morfologiche, quali la lunghezza e il numero di dendriti o "branches" differiscono tra neuroni RSTS e di controlli. A tal fine abbiamo utilizzando la Plugin "Analyse Skeleton" che permette di elaborare la cellula la cui identità neuronale è accertata dalla positività in immunofluorescenza per MAP2 e TUJ1(*Fig. 22*) in un'immagine binaria e di ottenerne lo "skeleton" per poter stimare precisamente il numero e la lunghezza dei branches in un campione di neuroni di pazienti e controlli. L'analisi di un numero di 30-50 cellule di vetrini trattati con gli anticorpi specifici di 3 controlli e 5 pazienti ha messo in luce in 4 su 5 pazienti una lunghezza media dei branches significativamente ridotta ed un numero medio dei branches significativamente aumentato rispetto ai controlli (*Fig. 23*): il paziente che non si diversifica dai controlli è quello con disabilità intellettiva più lieve (P46) portatore di una mutazione missenso di *CREBBP* che porta a sostituzione di un residuo aminoacidico nel dominio KAT causando attività acetiltrasferasica defettiva, ma non aploinsufficienza, come confermato dal Western Blot su estratti di iPSC di P46 che rileva con un anticorpo per la porzione-N-terminale di CBP una proteina "full length" in quantità analoga ai controlli (*Alari et al., 2019*). L'analisi di correlazione tra i due parametri morfologici valutati nei neuroni di pazienti e controlli conferma la tendenza opposta nei pazienti RSTS, mentre la correlazione è debole nei controlli (*Fig. 23*).

Elettrofisiologia dei neuroni corticali maturi

Neuroni maturi (>70 giorni) di un controllo e 4 pazienti RSTS sono stati analizzati mediante "patch-clamp recordings" per testare la loro maturità ed eccitabilità. Gli esperimenti hanno mostrato che i neuroni dei pazienti esprimono correnti di sodio e di potassio analogamente a neuroni maturi normali, come evidenziato dalle tracce rappresentative di un controllo e 4 pazienti (*Fig. 24*, primo pannello). Anche le correnti di potassio presentano ampiezze simili nei pazienti 46, 207 e 65, mentre risultano lievemente ma significativamente aumentate in P34 (*Fig. 24* secondo pannello). Esperimenti relativi al "clamp" delle correnti hanno confermato che i neuroni dei pazienti sono in grado di generare potenziali di azione in risposta a stimoli depolarizzanti, benché quelli di P34 e P207 in percentuale minore (*Fig. 24*, terzo pannello). Inoltre i neuroni di tutti i pazienti analizzati sono risultati "ipoeccitabili" rispetto a quelli dei controlli, come mostrato dai rapporti input-output (*Fig. 24* quarto pannello).

Le conclusioni di questo studio sono illustrate nella *Fig. 25* che schematizza i risultati raggiunti, ovvero: la generazione e caratterizzazione di un modello neuronale iPSC-derivato per la sindrome RSTS, sia causata da mutazioni di *CREBBP*, che di *EP300* e l' identificazione di biomarcatori che distinguono i neuroni dei pazienti da quelli di individui sani consistenti in alterazioni morfologiche (numero e lunghezza dei dendriti) dei neuroni giovani e alterazioni elettrofisiologiche indicative di una più lenta risposta agli stimoli (ipoeccitabilità) dei neuroni maturi. Maggiori dettagli del percorso sperimentale e degli approcci metodologici impiegati sono contenuti in *Alari et al., 2018a, Alari et al., 2018b* e *Alari et al., 2019.*

Trattamento dei neuroni di pazienti RSTS *in vitro* con potenziali agenti terapeutici

E' stato sottolineato che la sindrome RSTS è annoverata tra i disordini dell'apparato epigenetico per i quali è possibile una terapia postnatale che porti ad attenuazione della disabilità intellettiva. La disponibilità del modello neuronale *in vitro* permette di utilizzare pic-

cole molecole candidate a contrastare il difetto derivante dal deficit di acetilazione che sono gli inibitori delle iston-deacetilasi (HDACi) (*Fig. 26*). Esiste un'ampia gamma di tali composti molti dei quali utilizzati con successo nella cura di diverse patologie genetiche, su base genetica e tumorali (*Ganai, 2016*): tra questi i derivati degli acidi idrossamici SAHA/Vorinostat e tricostatina A (TSA) sono stata già testati *in vivo* in diversi modelli di topo *cbp+/cbp-* con miglioramento di vari difetti neurologici (*Alarcon et al., 2004; Korzus, 2004*), e dal nostro gruppo *in vitro* su linee linfoblastoidi di pazienti *CREBBP*-mutati in cui hanno determinato incremento del livello di acetilazione degli istoni della cromatina (*Lopez-Atalaya J , 2012*).

Abbiamo selezionato tra i pazienti RSTS modellati *in vitro* P149, che presenta il fenotipo neurologico più grave, sulla base anche di malformazioni cerebrali (ipoplasia del corpo calloso e ipoplasia cerebellare) e il più accentuato deficit cognitivo (*Fig. 12*) per effettuare esperimenti con tricostatina A (TSA) e valutare una possibile reversione delle alterazioni morfologiche ed elettrofisiologiche documentate nei neuroni giovani (45 giorni) e maturi (>70 giorni) dei pazienti RSTS. Per il primo timepoint (trattamento acuto) la supplementazione del terreno di coltura ha utilizzato per una settimana TSA 20nM, dose indicata da saggi di tossicità come la più alta che non influenza la percentuale di cellule vive sia nelle linee di controllo che nelle linee del paziente. Per il trattamento cronico (almeno 5 settimane) si è impiegata TSA 0,2nM.

Risultati preliminari su neuroni differenziati dai progenitori derivati da due cloni iPSC (6 e 7) di P149 mostrano rispetto a quelli ottenuti dai progenitori di controlli ridotta dimensione dei nuclei (*Fig. 27*) e diminuita lunghezza media dei branches (dendriti): questi parametri morfologici sono parzialmente, ma significativamente corretti dal trattamento con TSA.

ll trattamento a lungo termine (5 settimane) dei progenitori neurali di un controllo e di P149 con TSA 0,2nM fino allo stadio di neuroni maturi (>75 giorni) ha analogamente mostrato miglioramento dei parametri elettrofisiologici rispetto ai neuroni del paziente non trattati: quantificando la percentuale di neuroni in grado di generare almeno un potenziale d'azione con stimoli depolarizzanti si osserva differenza significativa tra il controllo e P149 non trattato così come tra P149 non

trattato e trattato con TSA, mentre non c' è differenza significativa tra controllo e P149 trattato con TSA (*Fig. 28*).

Questi risultati, i primi ad essere ottenuti su "neuroni" di pazienti RSTS, necessitano di convalida con esperimenti complementari di analisi citofluorimetrica dei marcatori di maturità neuronale del paziente 149 rispetto al controllo ai due time points indagati. Benché preliminari, essi tuttavia suggeriscono potenziali benefici di terapie postnatali per pazienti con questo disordine epigenetico, essendo dimostrata nel modello murino la rilevanza del deficit di CBP non solo durante lo sviluppo del cervello ma anche per le attività cognitive nel corso della vita (*Korzus , 2004*). L'adozione della terapia con "epidrugs" per i pazienti RSTS necessita inoltre ottimizzazioni relative alla specificità di substrato dei diversi tipi di HDACi e alla loro tossicità in trattamenti cronici. Pur con queste riserve tuttavia solo sperimentazioni *in vitro* sulle cellule neuronali di pazienti e controlli potranno integrare le modifiche di acetilazione degli istoni e proteine non istoniche causate dal difetto di CBP/p300 nel contesto appropriato di omeostasi neuronale.

Ringraziamenti

Si ringraziano i pazienti e le loro famiglie per la partecipazione a questo studio e l'Associazione Italiana per la Sindrome di Rubinstein-Taybi "Una Vita Speciale". Il lavoro di ricerca è stato supportato dal progetto ERA-NET NEURON JTC2015 ChromiSyn.

BIBLIOGRAFIA

Gurdon, JB. The developmental capacity of nuclei taken from intestinal epithelial cells of feeding tadpoles. J Embryol Exp Morph 10, 622-640, 1962.

Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663-676, 2006.

Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131, 861-872, 2007.

Zhao X. & Bhattacharyya A. Human Models Are Needed for Studying Human Neurodevelopmental Disorders. Am J Hum Genet 103, 829-857, 6, 2018.

Avior Y et al . Pluripotent stem cells in disease modelling and drug discovery. Nat Rev Mol Cell Biol. 17(3):170-182, 2016 doi:10.1038/nrm.2015.27

Trounson A & DeWitt ND. Pluripotent stem cells progressing to the clinic. Nat Rev Mol Cell Biol. 17(3):194-200, 2016. doi:10.1038/nrm.2016.10

Hennekam, R.C. Rubinstein-Taybi syndrome. Eur. J. Hum. Genet. 14(9), 981-985, 2006. doi: 10.1038/sj.ejhg.5201594.

Larizza L, Finelli P. Developmental disorders with intellectual disability driven by chromatin dysregulation: Clinical overlaps and molecular mechanisms. Clin Genet. 95(2):231-240., 2019 doi: 10.1111/cge.13365

Spena, S., Milani, D., Rusconi, D., Negri, G., Colapietro, P., Elcioglu, N., et al. Insights into genotype-phenotype correlations from CREBBP point mutation screening in a cohort of 46 Rubinstein-Taybi syndrome patients. Clin. Genet. 88(5), 431-440., 2015 doi: 10.1111/cge.12537.

Negri, G., Magini, P., Milani, D., Colapietro, P., Rusconi, D., Scarano, E., , Bonati MT, Priolo M, Crippa M, Mazzanti L, Wischmeijer A, Tamburrino F, Pippucci T, Finelli P, Larizza L, Gervasini C. From Whole Gene Deletion to Point Mutations of EP300-Positive Rubinstein-Taybi Patients: New Insights into the Mutational Spectrum and Peculiar Clinical Hallmarks. Hum. Mutat. 37(2), 175-183, 2016 doi: 10.1002/humu.22922.

Sheikh, B.N., and Akhtar, A.. The many lives of KATs - detectors, integrators and modulators of the cellular environment. Nat. Rev. Genet. 20(1), 7-23, 2019. doi: 10.1038/s41576-018-0072-4.

Bjornsson, H.T.. The Mendelian disorders of the epigenetic machinery. Genome Res. 25(10), 1473-1481, 2015 doi: 10.1101/gr.190629.115.

Fahrner, J. A., & Bjornsson, H. T. Mendelian disorders of the epigenetic machinery: postnatal malleability and therapeutic prospects. Human Molecular Genetics, 28(R2), R254-R264, 2019 doi:10.1093/hmg/ddz174

Alari V, Russo S, Rovina D, Garzo M, Crippa M, Calzari L, Scalera C, Concolino D, Castiglioni E, Giardino D, Prosperi E, Finelli P, Gervasini C, Gowran A, Larizza L Generation of three iPSC lines (IAIi002, IAIi004, IAIi003) from Rubinstein-Taybi syndrome 1 patients carrying CREBBP non sense c.4435G>T, p.(Gly1479*) andc.3474G>A, p.(Trp1158*) and missense c.4627G>T, p.(Asp1543Tyr) mutations. Stem Cell Res.;40:101553., 2019. doi: 10.1016/j.scr.2019.101553.

Alari V, Russo S, Terragni B, Ajmone PF, Sironi A, Catusi I, Calzari L, Concolino D, Marotta R, Milani D, Giardino D, Mantegazza M, Gervasini C, Finelli P, Larizza L. iPSC-derived neurons of CREBBP- and EP300-mutated Rubinstein-Taybi syndrome patients show morphological alterations and hypoexcitability. Stem Cell Res. 30:130-140, 2018. doi: 10.1016/j.scr.2018.05.019.

Alari V, Russo S, Rovina D, Gowran A, Garzo M, Crippa M, Mazzanti L, Scalera C, Prosperi E, Giardino D, Gervasini C, Finelli P, Pompilio G, Larizza L. Generation of the Rubinstein-Taybi syndrome type 2 patient-derived induced pluripotent stem cell line (IAIi001-A) carrying the EP300 exon 23 stop mutation c.3829A > T, p.(Lys1277*). Stem Cell Res. 30:175-179, 2018 doi: 10.1016/j.scr.2018.06.009.

Ganai SA, Ramadoss M, Mahadevan V. Histone Deacetylase (HDAC) Inhibitors - emerging roles in neuronal memory, learning, synaptic plasticity and neural rege-

neration. Curr Neuropharmacol.14(1):55-71, 2016 doi:10.2174/1570159x13666151021111609

Alarcón JM, Malleret G, Touzani K, et al. Chromatin acetylation, memory, and LTP are impaired in CBP+/- mice: a model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. Neuron. 42(6):947-959., 2004 doi:10.1016/j.neuron.2004.05.021

Korzus E, Rosenfeld MG, Mayford M. CBP histone acetyltransferase activity is a critical component of memory consolidation. Neuron.42(6):961-972, 2004. doi:10.1016/j.neuron.2004.06.002

Lopez-Atalaya, J.P., Gervasini, C., Mottadelli, F., Spena, S., Piccione, M., Scarano, G., Selicorni A, Barco A, Larizza L.. Histone acetylation deficits in lymphoblastoid cell lines from patients with Rubinstein-Taybi syndrome. J. Med. Genet. 49(1), 66-74, 2012 doi: 10.1136/jmedgenet-2011-100354.

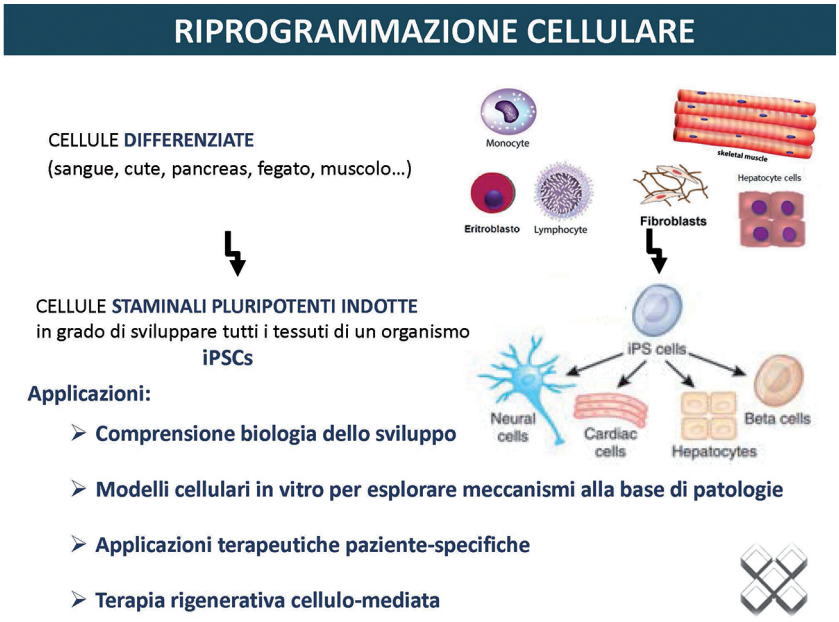


Fig. 1. Schema e applicazioni della riprogrammazione cellulare, il processo per cui cellule differenziate mature possono essere "indotte" *a diventare cellule staminali pluripotenti (iPSC) in grado di differenziarsi in tutti i tessuti dell'organismo.*

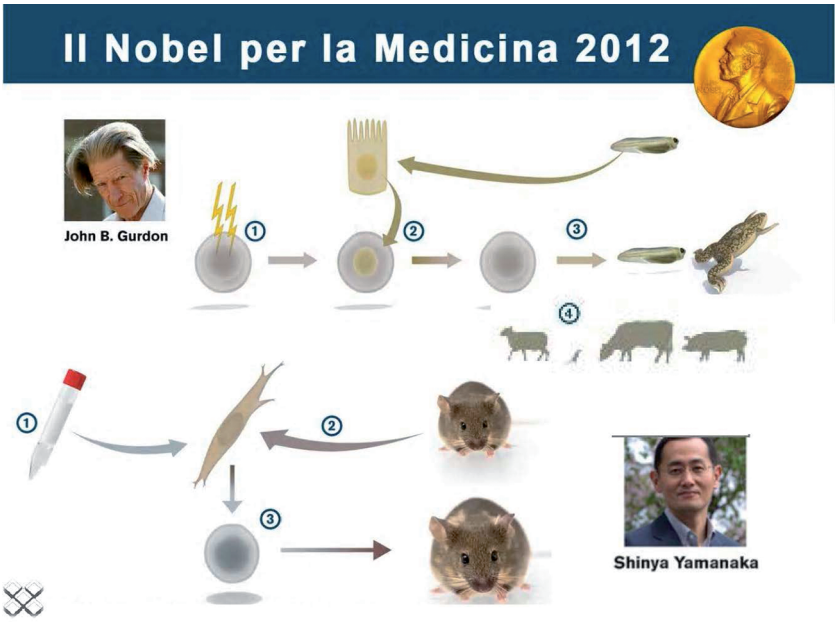


Fig. 2. Il Nobel per la medicina del 2012 condiviso dall'inglese Sir John Gurdon e dal giapponese Shynia Yamanaka. Gli esperimenti di Gurdon di "trasferimento nucleare" (1962) dimostrano che il nucleo di una cellula differenziata dell'epitelio intestinale di rana trasferito nella cellula uovo enucleata di rana può generare una cellula pluripotente da cui si sviluppa l'intero organismo della rana. Questi esperimenti hanno aperto la strada alla clonazione dei mammiferi, dalla pecora Dolly ai primati, ma sono stati accolti con riserve per via delle manipolazioni necessarie per indurre la riprogrammazione cellulare. Più di 40 anni dopo il gruppo di Yamanaka dimostra che una cellula differenziata del tessuto connettivo del topo (fibroblasto) "intatta" può essere riprogrammata in provetta a cellula pluripotente con una combinazione di soli 4 geni (identificati tra quelli in grado di mantenere lo stato di pluripotenza di cellule staminali embrionali) ed è poi in grado di generare l'intero organismo del topo.

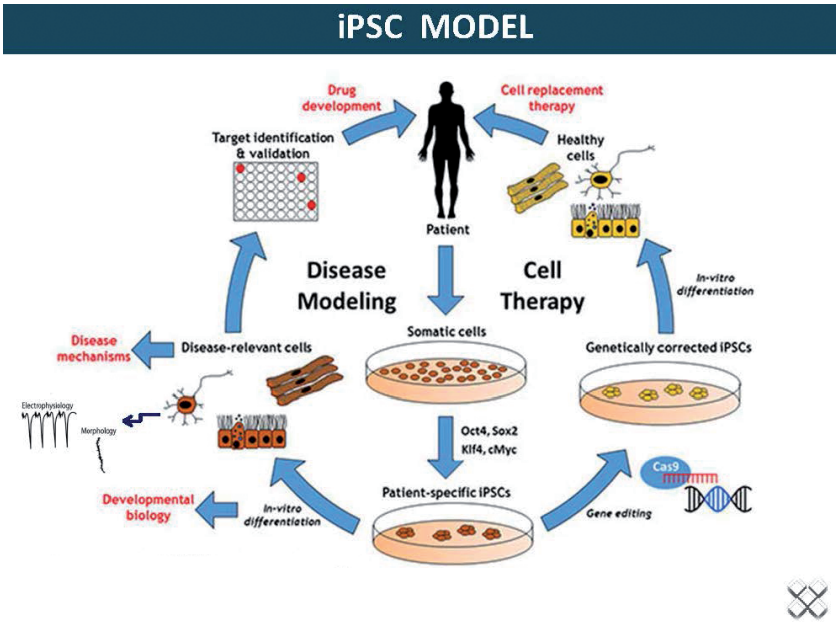


Fig. 3. Il modello dello cellule staminali pluripotenti indotte (iPSC): da cellule somatiche di un paziente con una patologia di interesse si generano tramite riprogrammazione con i fattori di Yamanaka iPSC che vengono fatte differenziare in vitro come desiderato (pannello a sinistra). Nel caso di differenziamento a neuroni si effettuano vari saggi, quali l'analisi morfologica e l'elettrofisiologia per caratterizzarli rispetto a neuroni ottenuti da individui sani. Si possono anche testare piccole molecole (farmaci) per valutare la possibile reversione al fenotipo di neuroni normali. Dalle iPSC (pannello a destra) è possibile con tecniche di "editing" correggere le mutazioni paziente-specifiche e ottenere cellule differenziate normali da utilizzare per terapia sostitutiva paziente-specifica.



Fig. 4. La sindrome di Rubinstein-Taybi (RSTS): patologia Mendeliana rara a trasmissione autosomico dominante, comprendente due entità cliniche, RSTS1 e RSTS2, causate da mutazioni nei geni omologhi CREBBP ed EP300. Sono elencati i principali segni clinici: tra questi la disabilità intellettiva, presente in tutti (RSTS1) o quasi tutti (RSTS2) i pazienti. I meccanismi alla base del deficit cognitivo dei pazienti RSTS possono essere studiati attraverso il modello di iPSC paziente-specifiche differenziate in vitro a neuroni corticali.

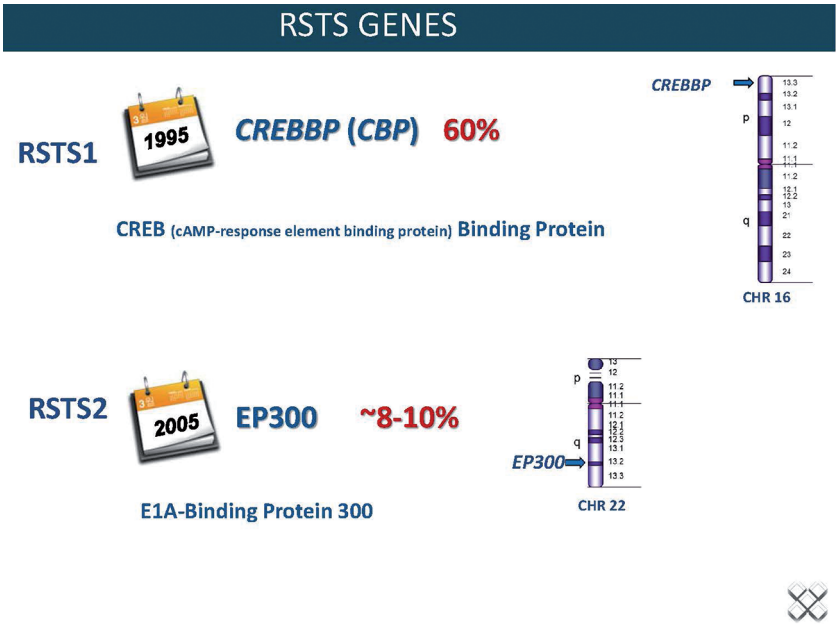


Fig. 5. I geni della Sindrome di Rubinstein-Taybi: CREBBP ed EP300: localizzazione cromosomica e percentuali di pazienti con diagnosi clinica giustificate dalle loro mutazioni.

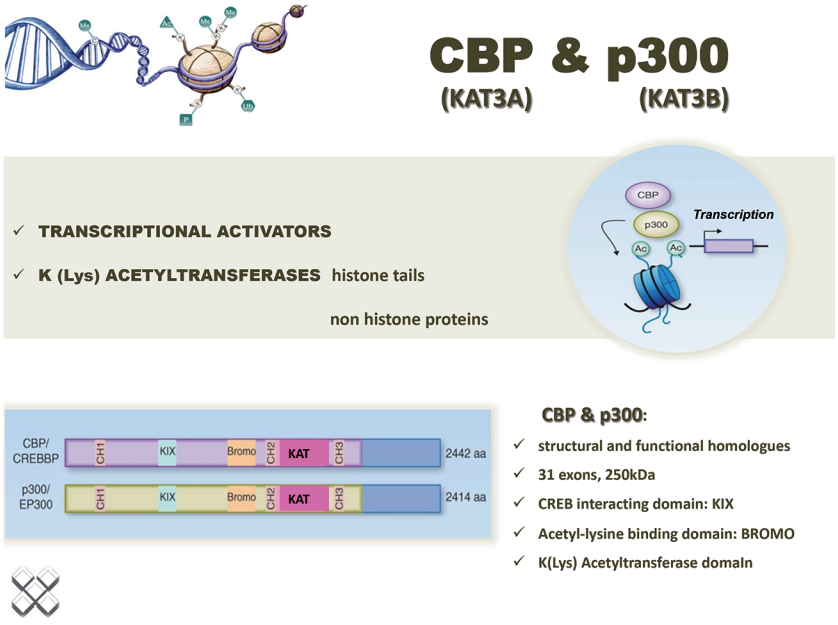


Fig. 6. Le proteine CBP e p300: coattivatori della trascrizione genica con attività acetiltrasferasica diretta verso la lisina (K) sia delle code degli istoni della cromatina che di proteine non istoniche. Lo schema della loro struttura con i vari domini (al centro il dominio KAT) evidenzia l'omologia di sequenza e rende conto delle funzioni condivise. Rilevante il ruolo nello sviluppo del sistema nervoso e nella plasticità neurale.

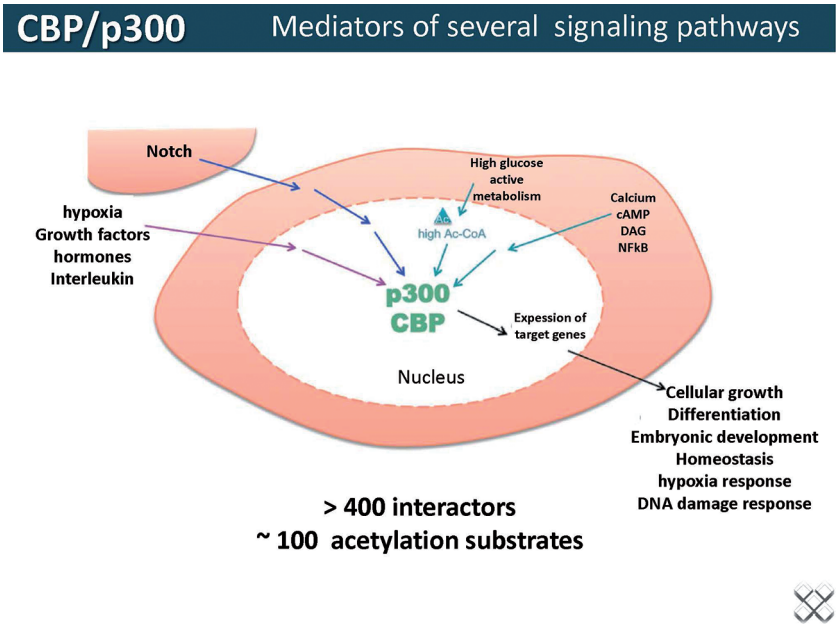


Fig. 7. Le proteine nucleari CBP e p300: mediatori di un numero elevatissimo di pathways di signaling extra-, inter- e intracellulari grazie alla loro capacità di legare una moltitudine di proteine diverse (ad oggi riconosciuti >400 interattori) usando i loro domini di interazione fungendo anche da "scaffold" tra interattori diversi.

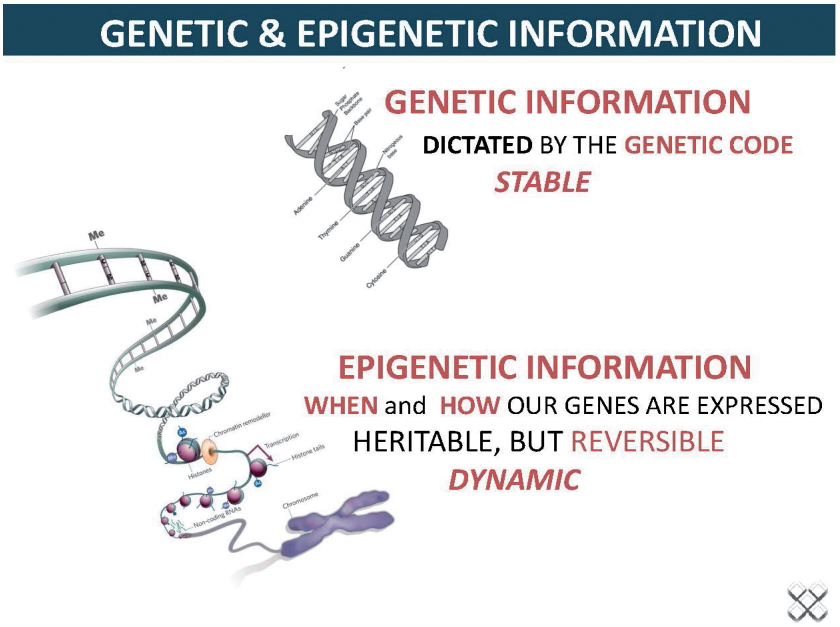


Fig. 8. Informazione Genetica ed Epigenetica. La prima è dettata dalla sequenza delle basi del DNA, è stabile e irreversibile, mentre l'informazione epigenetica, rappresentata da metilazione di una delle 4 basi del DNA (citosina) e da modificazioni post-traduzionali degli istoni della cromatina, è ereditabile, ma dinamica e reversibile. Le modificazioni epigenetiche controllano quando *e* come *i nostri geni sono espressi.*

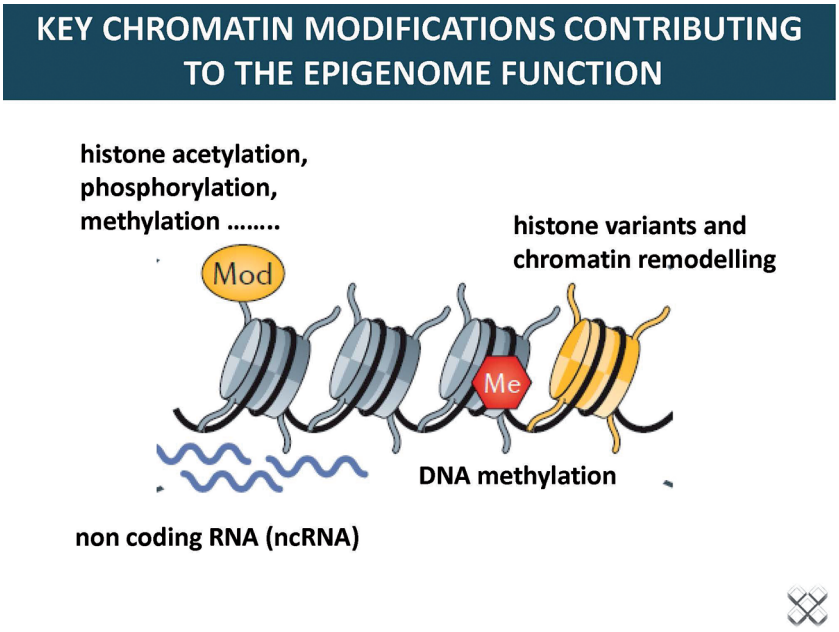


Fig. 9. Principali modificazioni epigenetiche: metilazione del DNA, modificazioni degli istoni, varianti istoniche e piccoli RNA non codificanti. Tali modificazioni definiscono la struttura tridimensionale della cromatina (il DNA complessato a proteine istoniche nel nucleo delle cellule eucariotiche) fungendo da ponte molecolare tra il controllo della trascrizione genica e l'ambiente.

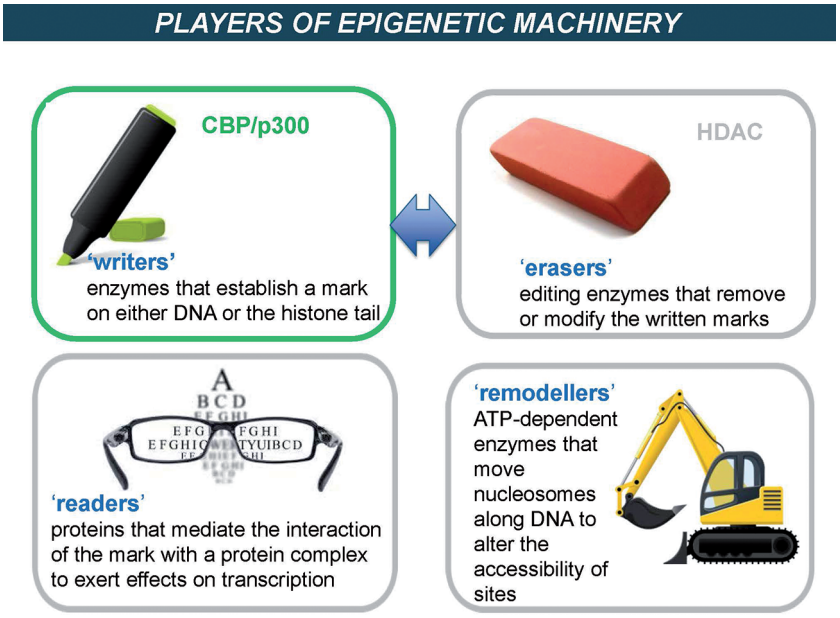


Fig. 10. I principali players dell'apparato epigenetico: "writers", "erasers", "readers" e "remodellers". E' evidenziata l'azione enzimatica opposta di writers ed erasers, qui esemplificata per CBP/p300 e istondeacetilasi (HDAC): in condizioni normali esiste una bilancia precisa tra queste attività enzimatiche, mentre a seguito di un difetto di CBP o p300 , l'equilibrio si sposta a favore delle HDAC.

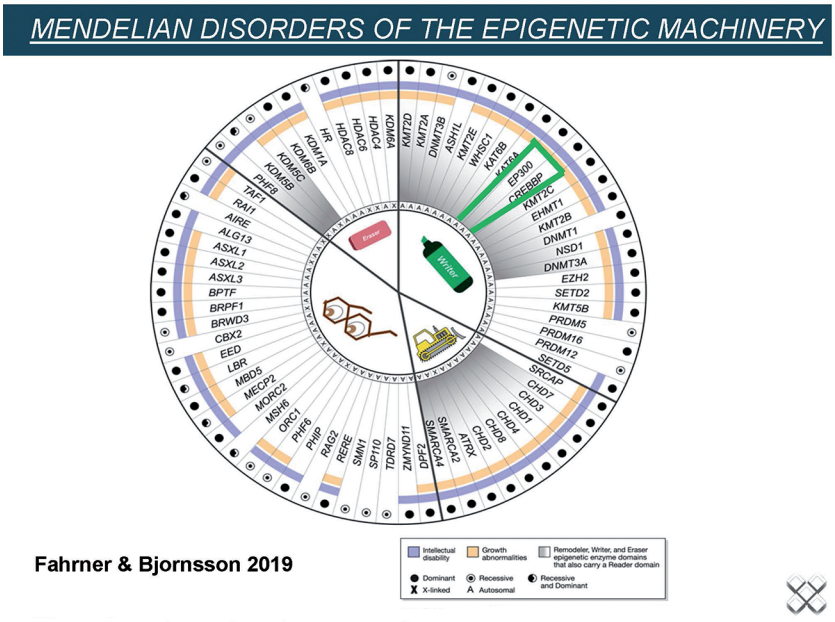


Fig. 11. Disordini mendeliani dell'apparato epigenetico, causati da mutazioni in geni che codificano proteine con domini "epigenetici" classificati nelle categorie di "writers", "erasers", "readers" e "remodellers" (icone al centro). *I domini enzimatici (writer, eraser, remodeller) sono mutualmente esclusivi in ogni data proteina, ma molti coesistono con un dominio reader* (sfumatura grigia). *La maggior parte dei geni causa la patologia in condizione eterozigote* (circolo pieno). *La disabilità intellettiva è condivisa dalla maggior parte di questi disordini* (cerchio viola) *così come i difetti della crescita* (cerchio arancione). *La sindrome di Rubinstein-Taybi (1 e 2)* (riquadro verde) *risulta da mutazioni nei geni CREBBP o EP300 codificanti per proteine con attività acetiltransferasica (writers) provviste anche di un dominio "reader". (Ristampato con permesso. Human Molecular Genetics,* Volume 28, Issue R2, 15 October 2019, Pages R254–R264, https://doi.org/10.1093/hmg/ddz174).

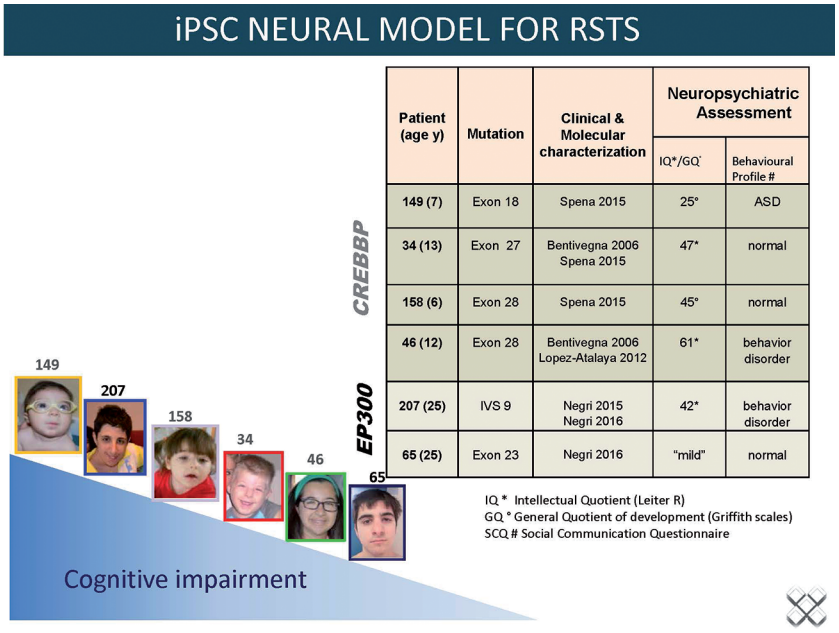


Fig. 12. Modello di neuroni iPSC-derivati per la sindrome di Rubinstein-Taybi: la riprogrammazione a iPSC e il differenziamento a neuroni corticali è stato effettuato per 6 pazienti RSTS, 4 CREBBP- e 2 EP300-mutati qui ordinati secondo il livello di deficit cognitivo dal più grave al più lieve. Sono riportati l'età di ciascun paziente al momento della riprogrammazione, e la valutazione neuropsichiatrica.

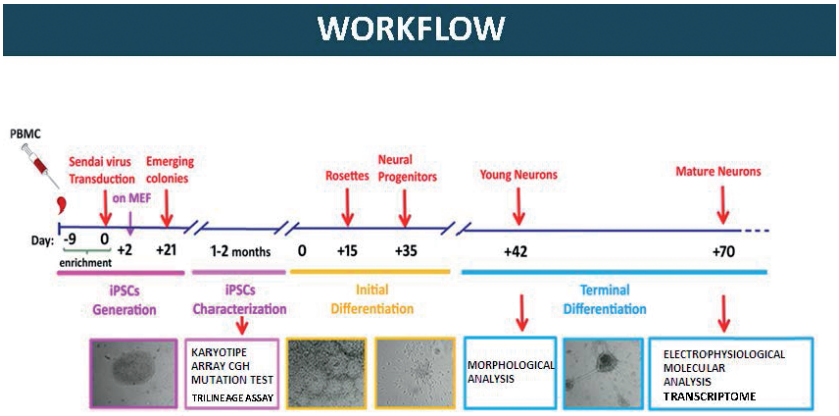


Fig. 13. Generazione di neuroni iPSC-derivati da pazienti RSTS e individui sani: workflow. Vedi testo per dettagli.

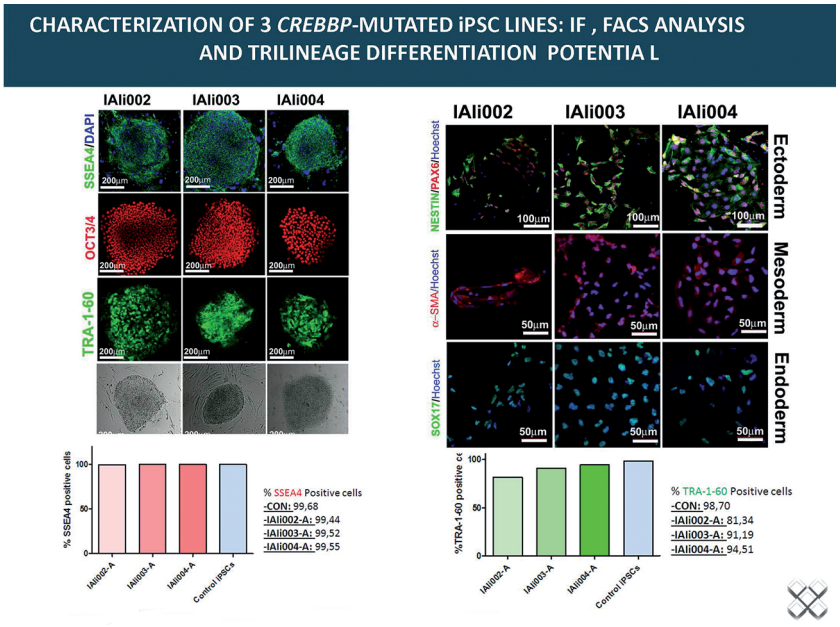


Fig. 14. Caratterizzazione di tre linee iPSC generate da pazienti CREBBP*-mutati.* Pannello a sinistra: *espressione positiva degli antigeni di staminalità (SSEA4, OCT3/4 e TRA-1-60) valutata qualitativamente in immunofluorescenza al microscopio confocale, e quantitativamente al FACs (SSEA4 e TRA-1-60) e morfologia distintiva delle colonie al microscopio rovesciato.* Pannello a destra: *potenziale differenziativo delle colonie iPSC lungo l'asse ectodermico (Tra-1-60 +), mesodermico (alfa-SMA+) e endodermico (SOX17+).*

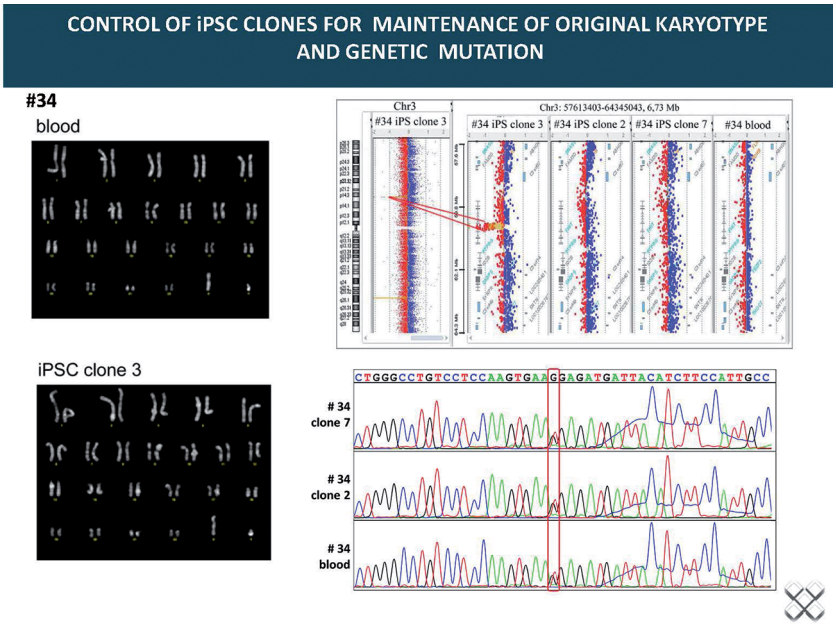


Fig. 15. Caratterizzazione della stabilità genomica di tre cloni iPSC di un paziente CREBBP-mutato. Pannello a sinistra: cariotipo in bande Q dei linfociti del sangue periferico del paziente 34 (in alto) e del clone 3 delle iPSC generate dallo stesso paziente, che non mostra apparenti anomalie (in basso). Pannello a destra: in alto profilo del cromosoma 3 mediante ibridazione genomica comparativa su array (array-CGH) che mostra nel clone iPSC 3 del paziente una rara delezione in corrispondenza della regione 3p14.2 (indicata dalle linee rosse convergenti). Questa variante non è presente nei cloni iPSC 1 e 7 e nel sangue periferico del paziente ed è presumibilmente insorta in vitro: il clone iPSC 3 viene pertanto scartato. In basso: il sequenziamento Sanger del DNA genomico dal sangue periferico e i cloni iPSC 7 e 2 del paziente evidenzia la mutazione in eterozigosi c.4435G>T nell'esone 27 del gene CREBBP (riquadro rettangolo rosso).

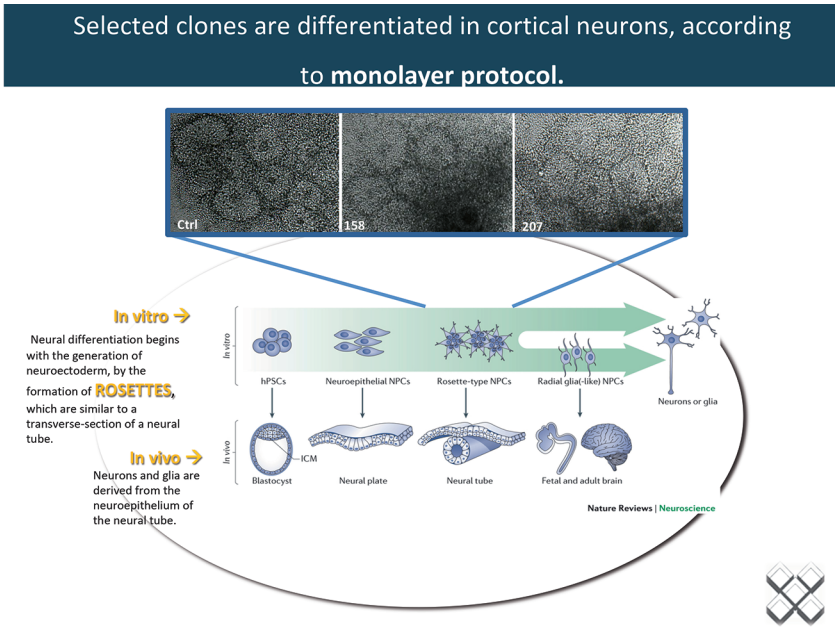


Fig. 16. Differenziamento neurale in vitro *e* in vivo. *A due settimane dall'inizio del differenziamento neurale secondo il protocollo in monostrato compaiono le rosette mostrate al microscopio a contrasto di fase per un controllo e due pazienti RSTS. Si noti la loro somiglianza alla sezione trasversale del tubo neurale che si origina* in vivo *dalla piastra neurale.*

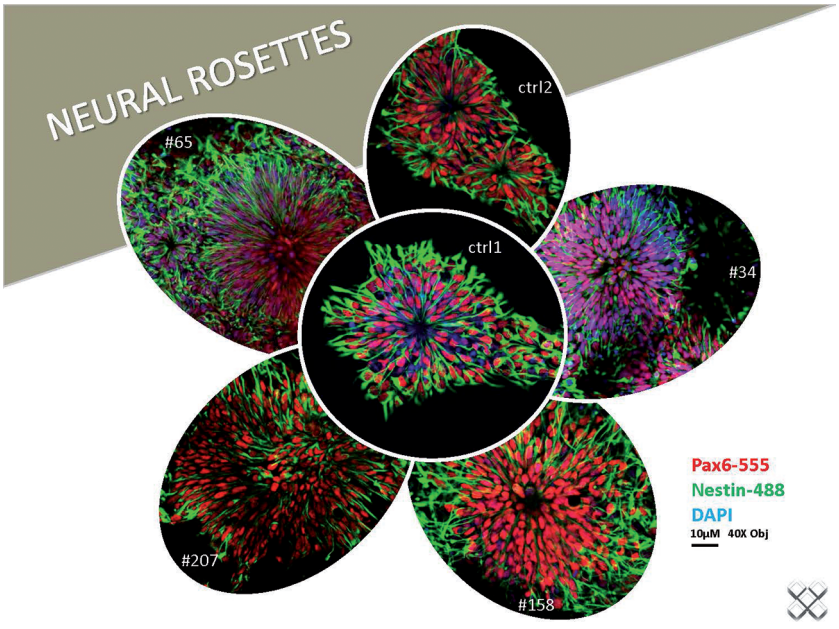


Fig. 17. Caratterizzazione mediante immunofluorescenza di rosette (20 giorni) di controlli e pazienti RSTS. Positività per i marcatori neuroectodermici Nestina e PAX6. Controcolorazione dei nuclei con DAPI.

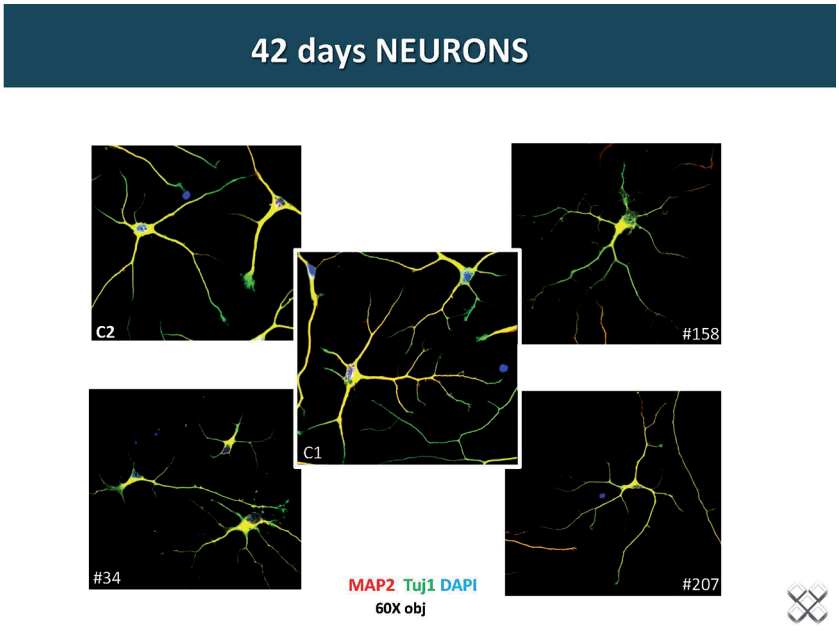


Fig. 18. Caratterizzazione mediante immunofluorescenza di neuroni a 42 giorni di controlli e pazienti RSTS. Positività per i marcatori pan neuronali MAP2 (Microtubule-associated protein 2) e TUJ1 (Tubulina betaIII). Controcolorazione dei nuclei con DAPI.

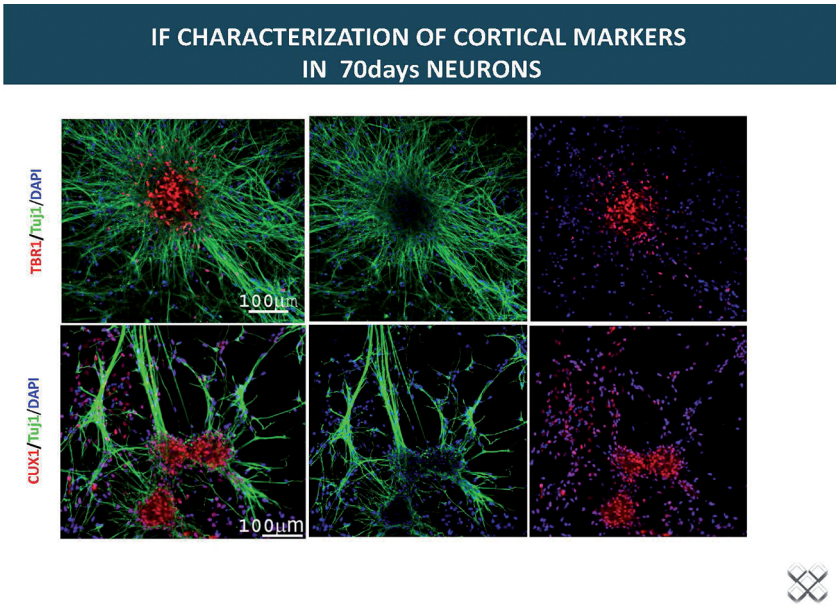


Fig. 19. Immunofluorescenza di neuroni postmitotici (>70 giorni). Positività per TUJ1 e i marcatori corticali TBR1 e CUX1.

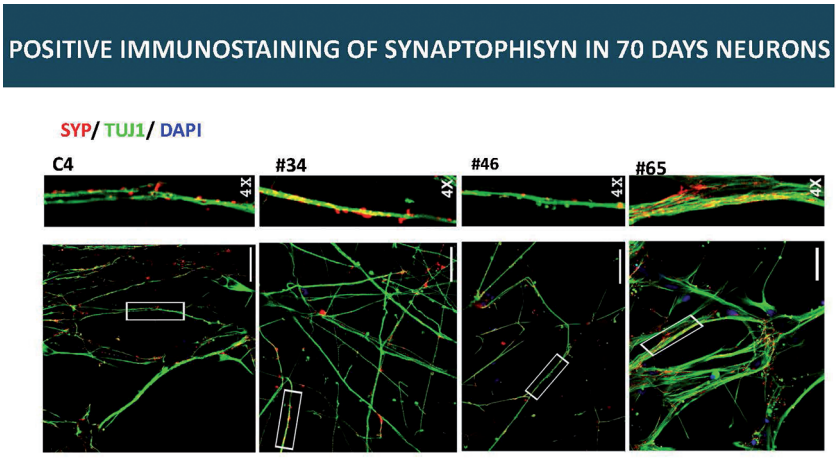


Fig. 20. Immunofluorescenza di neuroni (>70 giorni) di un controllo e 3 pazienti RSTS. Positività per TUJ1 e il marcatore pre-sinaptico sinaptofisina (SYP). I campi incorniciati sono mostrati a ingrandimento 4x sopra le rispettive foto.

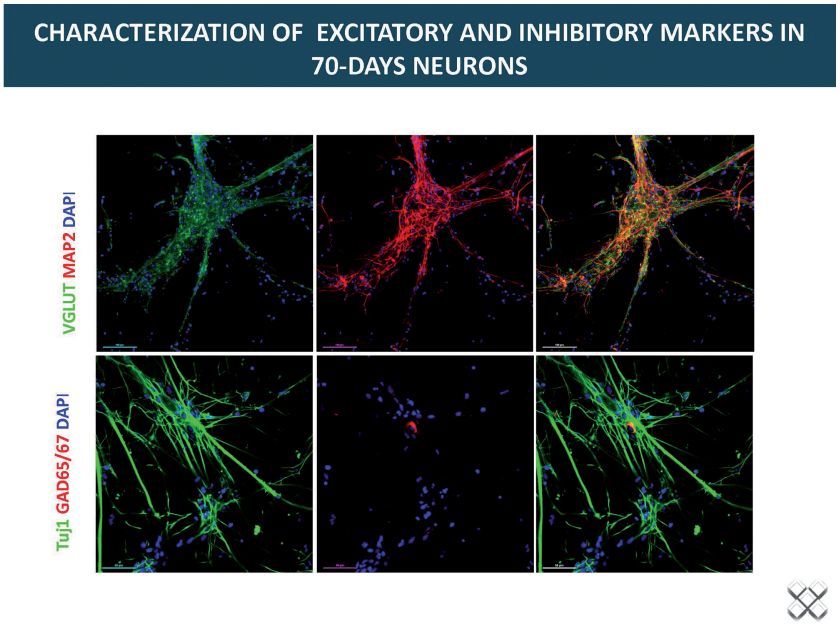


Fig. 21. Immunofluorescenza di neuroni postmitotici (>70 giorni) per vGLUT1/2, marker dei neuroni eccitatori e MAP2 e per GAD65/67, marker dei neuroni inibitori e TUJ1. Sia nei neuroni di controllo che in quelli dei pazienti RSTS i neuroni positivi per vGLUT 1/2, concentrati nei nodi ad alta densità cellulare sono molto più numerosi dei neuroni positivi per i marker inibitori.

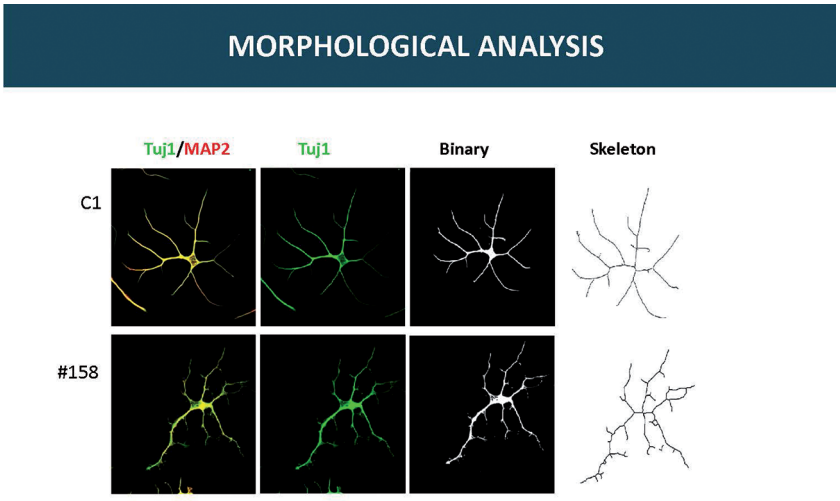


Fig. 22. Analisi morfologica dei neuroni a 42 giorni. Parametri morfologici, quali la lunghezza e il numero dei "branches" sono analizzati in neuroni giovani, presenti nelle piastre a bassa densità, utilizzando la Plugin "Analyze skeleton" (Image J). Come mostrato per neuroni di un controllo e di un paziente RSTS la stessa cellula singola la cui identità neuronale è confermata dalla positività per MAP2 e TUJ1 è elaborata come immagine binaria e ridotta a "sketeton" in modo da permettere di contare precisamente la lunghezza delle ramificazioni (branches) e il loro numero. L'analisi di un numero medio di 30-50 cellule di 3 campioni di controllo e 5 di pazienti RSTS ha permesso di graficare questi parametri.

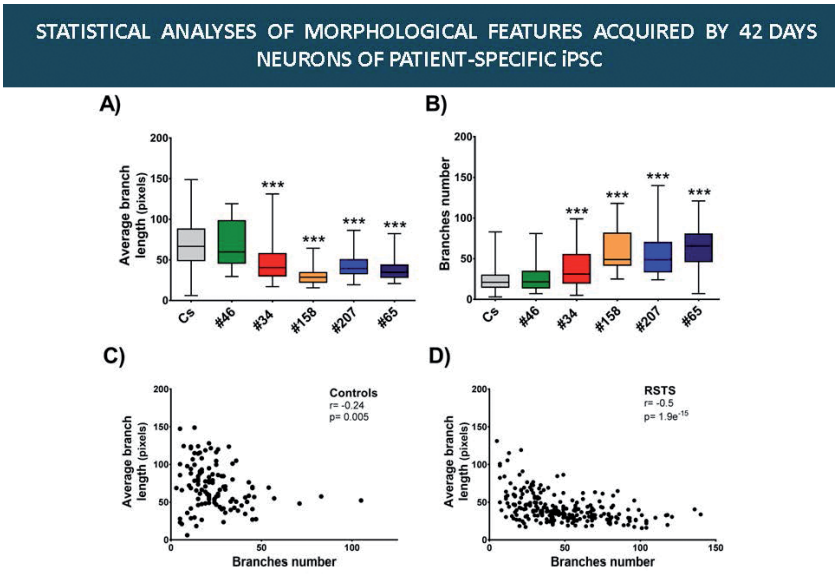


Fig. 23. Analisi statistica di parametri morfologici ottenuti analizzando neuroni di 42 giorni ottenuti da iPSC di 5 pazienti RSTS. A) I box plots mostrano che i neuroni dei pazienti RSTS con l'eccezione di P46 hanno una lunghezza media dei branches significativamente ridotta rispetto a quella di un pool di neuroni di controllo. B) i box plots evidenziano un distribuzione simile ma opposta per il numero medio dei branches con i neuroni di P46 simili ai controlli e quelli degli altri pazienti con numero medio di branches significativamente maggiore dei controlli. La tendenza opposta di questi due parametri morfologici nei pazienti RSTS è confermata dall'analisi di correlazione (C, D): la lunghezza e il numero medio dei branches sono negativamente correlati nei pazienti, mentre la correlazione è debole nei controlli.

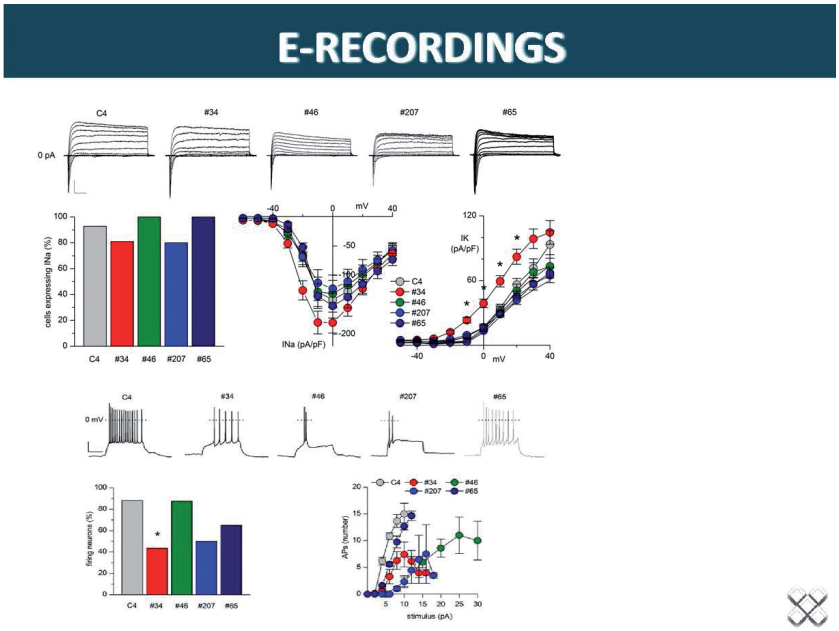


Fig. 24. Caratterizzazione elettrofisiologica di neuroni di controllo e di pazienti RSTS dopo >70 giorni di differenziamento. Pannelli superiori: Come si evidenzia dalle tracce rappresentative delle correnti ioniche totali indotte da stimoli depolarizzanti i neuroni di pazienti RSTS si comportano in modo simile a neuroni maturi: le correnti di sodio sono di ampiezza paragonabile , quelle di potassio hanno un'ampiezza accresciuta rispetto a neuroni del controllo solo in P34. Pannelli inferiori: rispetto a neuroni di controllo una ridotta percentuale di neuroni di P34 e P207 é in grado di generare potenziali d'azione ; inoltre i neuroni di tutti i pazienti investigati si dimostrano meno eccitabili , come dimostrato dai rapporti "input-output", richiedendo uno stimolo di intensità superiore per raggiungere la soglia necessaria per generare potenziali d'azione.

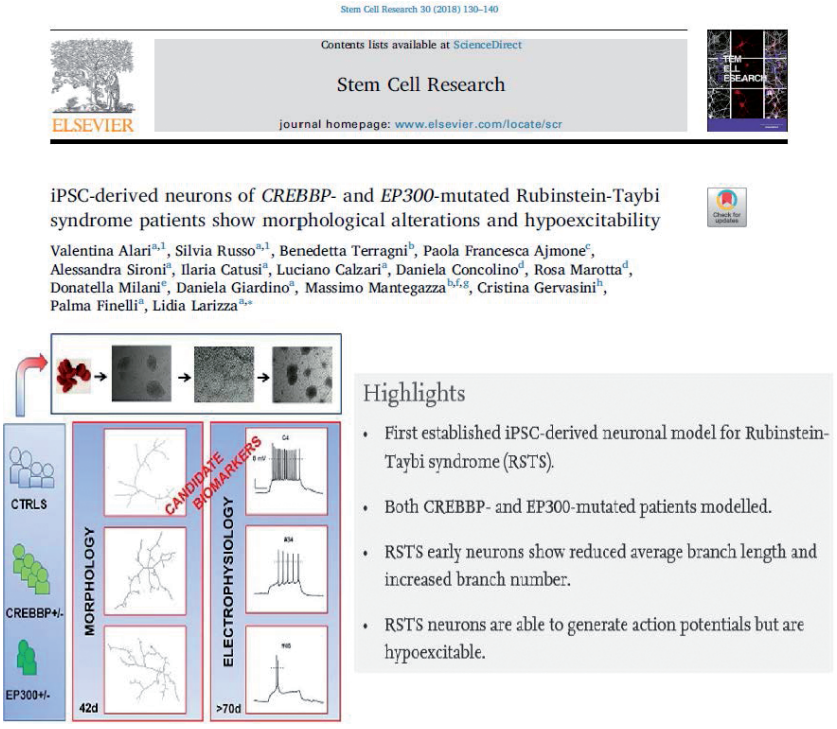


Fig. 25. Conclusioni dello studio: Generazione del primo modello neuronale iPSC-derivato per la sindrome di Rubinstein-Taybi. I neuroni RSTS in differenziamento presentano ridotta lunghezza e aumentato numero dei "branches". I neuroni RSTS postmitotici sono in grado di generare potenziali d'azione, ma sono ipoeccitabili.

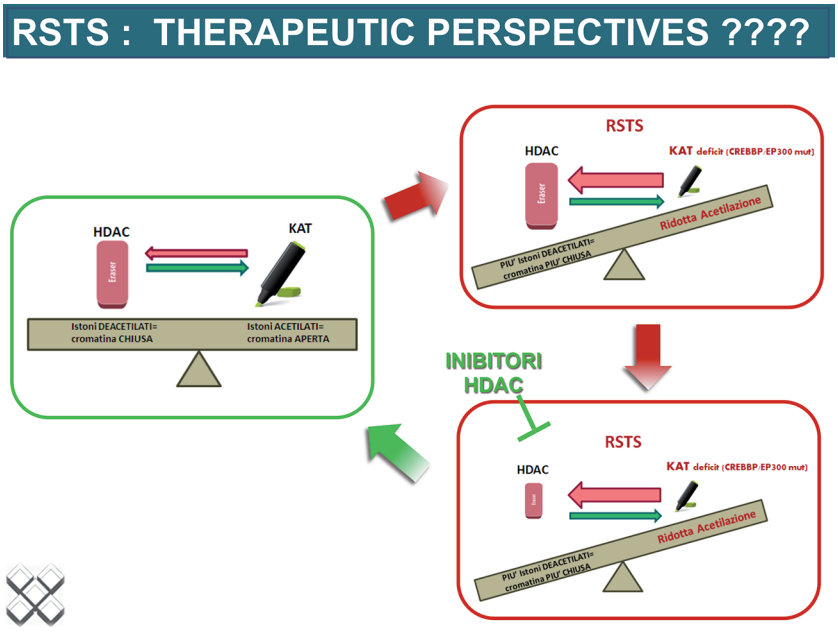


Fig. 26. Prospettive di terapia postnatale per i pazienti RSTS: il difetto di CBP o p300 determina un deficit di acetilazione di proteine istoniche e non istoniche che fa pendere la bilancia tra istonacetiltransferasi (KAT) e istondeacetilasi (HDAC) a favore di queste ultime. Inibitori delle HDAC (HDACi) posssono ripristinare l'equilibrio.

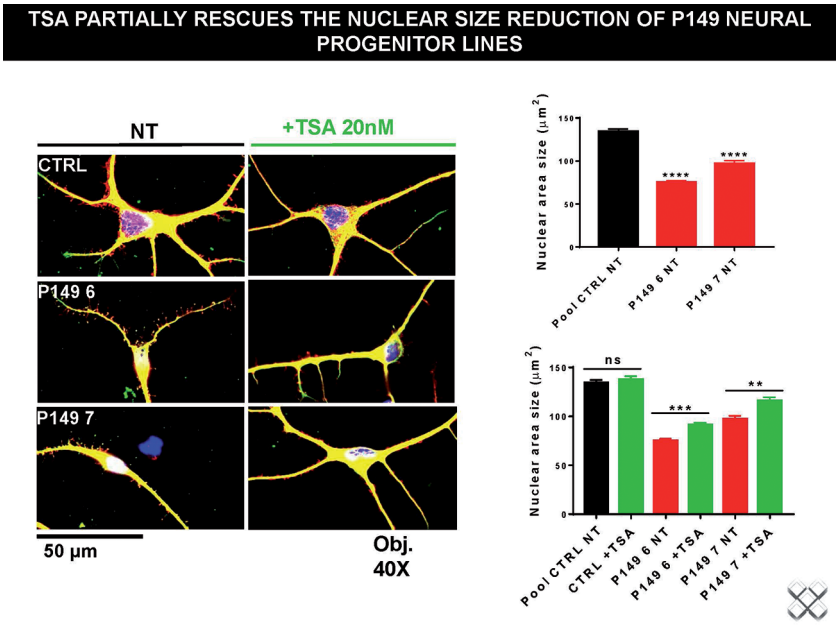


Fig. 27. Trattamento con Tricostatina A (TSA) e attenuazione di alterazioni morfologiche dei neuroni giovani del paziente 149, il più grave per il fenotipo neurocognitivo tra quelli modellati. TSA 20nM supplementata per una settimana al terreno di coltura di due linee indipendenti di progenitori neurali di P149 corregge la ridotta dimensione dell'area nucleare osservabile nelle linee non trattate.

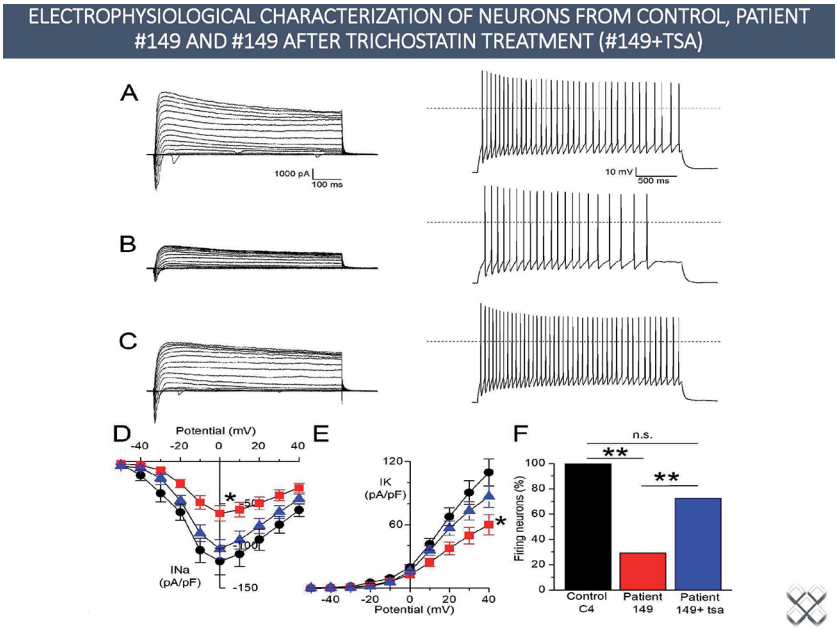


Fig. 28. Trattamento cronico con TSA e parziale reversione delle disfunzioni agli e-recordings di neuroni >70 giorni di P149. A, B, C: correnti ioniche totali dei neuroni di un controllo, di P149 e P149 dopo trattamento «long term» (35 giorni) con TSA (0.2 nM). D: plot della quantificazione della densità delle correnti di sodio a diversi potenziali per C4 (cerchietti neri, P149 (quadrati rossi) e P149 +TSA (triangoli bleu). E: plot della quantificazione delle densità delle correnti di potassio a diversi potenziali (simboli analoghi). F: quantificazione della percentuale di neuroni in grado di generare almeno un potenziale d'azione per C4, P 149 e P149 trattato con TSA. La differenza tra C4 e P149+TSA esiste, ma non è significativa, mentre lo è tra C4 e P149 e tra P149 e P149+TSA.

QUANDO L'INSIEME È MEGLIO DELLA SOMMA: NANOMATERIALI IBRIDI ARGENTO-TITANIA ED ORO-TITANIA PER SENSORI ELETTROANALITICI

LUIGI FALCIOLA (*), VALENTINA PIFFERI (*), ANNA TESTOLIN (*)

Nota presentata dal s.c. Carlo Edoardo Gatti
(Adunanza del 19 dicembre 2019)

SUNTO. – L'interesse della comunità scientifica per le nanotecnologie ed i nanomateriali nasce nel 1959, in occasione della riunione annuale della American Physical Society, quando ebbe inizio un nuovo approccio con cui considerare e manipolare la materia su scala micro- e nano-metrica. Il successivo aumento dell'utilizzo di nanomateriali nella ricerca scientifica è dovuto alle loro caratteristiche peculiari: aumentato rapporto superficie/volume, nuove proprietà ottiche e fisiche, elevata superficie attiva, trasporto di materia incrementato o modificato, selettività aumentata. Tutte queste caratteristiche rendono i nanomateriali estremamente adatti per un vasto spettro di applicazioni, dall'energetica, alla catalisi ed alla sensoristica. Negli ultimi decenni, a parte l'applicazione di nanomateriali caratterizzati da un solo componente, sta rapidamente crescendo anche l'utilizzo di nanomateriali "ibridi", composti da due o più componenti in "intimo" contatto. Ciò è connesso al fatto che, poiché in un materiale ibrido l'instaurarsi di proprietà sinergiche è confinato alla regione di contatto tra i differenti componenti, nei nanomateriali, date le piccole dimensioni, questi effetti possono diventare un fattore dominante dell'intera struttura. Il risultato è appunto l'instaurazione di proprietà nuove e inaspettate (spesso volute ed interessanti) che non sono la semplice somma della proprietà dei contributi individuali. Considerato tutto ciò, i materiali ibridi sono degli ottimi candidati per applicazioni nei campi più svariati, tra i quali l'elettroanalisi, che può trarre grande vantaggio da questi sistemi. In questo contesto, questo lavoro vuole presentare la costruzione di dispositivi basati su nanomateriali ibridi metallo-semiconduttore (a base di nanoparticelle d'argento o d'oro e biossido di titanio). In questi sistemi la presenza delle nanoparticelle metalliche viene sfruttata per la determinazione elettroanalitica di ana-

(*) Gruppo di Chimica ELettroANalitica (ELAN), Dipartimento di Chimica, Università degli Studi di Milano, Italy. E-mail: luigi.falciola@unimi.it

liti di interesse medico/diagnostico o ambientale, mentre la presenza del TiO_2 risulta essenziale per superare i problemi di sporcamento e passivazione della superficie elettrodica a seguito della determinazione della molecola in esame. La possibilità di rinnovare la superficie elettrodica semplicemente irradiando il dispositivo con luce ultravioletta, permette di utilizzare i dispositivi in remoto per lungo tempo, senza il loro continuo rispristino. Nel caso del sistema Ag-TiO_2, si è inoltre riscontrato l'instaurarsi di una interfase di "intimo contatto" tra le nanoparticelle d'argento ed il biossido di titanio, che rende l'ibrido un vero e proprio nuovo materiale, con caratteristiche differenti e con applicazioni interessanti rispetto ai suoi precursori.

***

ABSTRACT. – The interest of the scientific community for nanotechnologies and nanomaterials was born in 1959, on the occasion of the annual meeting of the American Physical Society, when a new approach began with which to consider and manipulate matter on a micro- and nano-metric scale. The subsequent increase in the use of nanomaterials in scientific research is due to their peculiar characteristics: increased surface / volume ratio, new optical and physical properties, high active surface, increased or modified material transport, increased selectivity. All these features make nanomaterials extremely suitable for a wide range of applications, from energy, to catalysis and sensors. In recent decades, apart from the application of nanomaterials characterized by a single component, the use of "hybrid" nanomaterials, composed of two or more components in "intimate" contact, is also rapidly growing. This is connected to the fact that, since in a hybrid material the establishment of synergistic properties is confined to the contact region between the different components, in nanomaterials, given the small size, these effects can become a dominant factor of the entire structure. The result is precisely the establishment of new and unexpected properties (often desired and interesting) which are not the simple sum of the ownership of individual contributions. Given all this, hybrid materials are excellent candidates for applications in the most varied fields, including electroanalysis, which can greatly benefit from these systems. In this context, this work aims to present the construction of devices based on hybrid metal-semiconductor nanomaterials (based on silver or gold nanoparticles and titanium dioxide). In these systems, the presence of metal nanoparticles is exploited for the electroanalytical determination of analytes of medical / diagnostic or environmental interest, while the presence of TiO2 is essential for overcoming the problems of fouling and passivation of the electrode surface following the determination of the molecule under examination. . The possibility of renewing the electrode surface simply by irradiating the device with ultraviolet light, allows you to use the devices remotely for a long time, without their continuous recovery. In the case of the Ag-TiO2 system, an "intimate contact" interphase between the silver nanoparticles and titanium dioxide was also found, which makes the hybrid a real new material, with different characteristics and with interesting applications compared to its precursors.

1. Introduzione

1.1 Nanomateriali ed applicazioni elettroanalitiche

L'inizio dell'era delle nanotecnologie viene fatto risalire al 29 dicembre 1959 quando, durante l'incontro annuale dell'American Physical Society, il noto fisico Richard Feynman (1918-1988), pronunciò l'ormai nota e citata frase:"*There's plenty of room at the bottom*" ("C'è un sacco di spazio laggiù in fondo") [1]. Con questa frase Feynman ha spianato la strada verso un nuovo approccio nella manipolazione e nel controllo delle proprietà della materia su scala miniaturizzata e quindi micro- e nano-metrica, a livello atomico e molecolare, accentuando l'attenzione sul nuovo mondo delle nanotecnologie.

Una definizione più completa e condivisa di "nanomateriale" è stata data il 18 ottobre 2011 dalla Commissione Europea: un nanomateriale è un "*materiale naturale, incidentale o fabbricato contenente particelle, in uno stato non legato o come aggregato o agglomerato dove, per il 50% o più delle particelle nella distribuzione dimensionale numerica, una o più dimensioni esterne sono comprese nel range dimensionale 1 nm - 100 nm*". Quindi, è possibile distinguere nanoparticelle (zero dimensionali), nanofili (monodimensionali) e nanofogli (bidimensionali). Molte entità biologiche, come proteine, acidi nucleici e virus rientrano nella scala nanodimensionale, ma anche diversi materiali sintetici di varia composizione chimica, dagli elementi chimici singoli (es. metalli, carbonio, silicio) alle molecole inorganiche (es. ossidi metallici) ed organiche (es. polimeri), con numerose applicazioni scientifiche e tecnologiche, dalla sensoristica alla catalisi, grazie alle superiori proprietà ed in particolare all'elevato rapporto superficie/volume.

Nel campo specifico dell'elettrochimica e dell'elettroanalisi, la peculiarità dei nanomateriali si esprime in nuove proprietà quantomeccaniche (effetti quantici dimensionali), chimico-fisiche ed ottiche. Inoltre, i nanomateriali permettono un trasporto di materia aumentato (che facilita lo studio di processi elettrochimici veloci), hanno un'elevata superficie attiva, una migliore selettività e costi ridotti. Sono inoltre caratterizzati da un elevato rapporto segnale/rumore e da un meccanismo di diffusione migliorato o in generale modificato, passando generalmente ad una diffusione non planare (convergente o radiale). Si

hanno inoltre diminuzione dei contributi capacitivi (la capacità del doppio strato è correlata all'area elettrodica) e riduzione della caduta ohmica (che è proporzionale alla corrente totale che passa nel circuito). La possibilità di costruire array di micro- o nano-elettrodi, permette di incrementare ulteriormente le prestazioni di questi dispositivi, che sono in grado di fornire una risposta voltammetrica di altezza (intensità di corrente) simile ai loro analoghi in dimensione macro, ma con correnti capacitive considerevolmente inferiori. [2]

Tutte queste proprietà consentono anche di fornire un maggior controllo sul microambiente locale, rendendo i nanomateriali buoni candidati per applicazioni innovative in sensori elettroanalitici. [3-4]

1.2 Materiali ibridi ed applicazioni elettroanalitiche

Secondo lo IUPAC (International Union of Pure and Applied Chemistry) Gold Book [5] un "ibrido" è definito come un "materiale composto da una miscela intima di componenti o inorganici o organici o di entrambi i tipi, in cui di solito tali componenti si compenetrano su scale inferiori a 1 μm". E sta proprio nel termine "intimo" che caratterizza la relazione tra i due componenti, la differenza con i materiali compositi (ossia, semplici materiali multicomponente [5]): negli ibridi, lo stretto legame ("intimo") tra i due componenti è in grado spesso di generare veri e propri nuovi materiali, con caratteristiche e proprietà differenti e molto spesso superiori rispetto ai componenti singoli. Mentre in un composito le proprietà possono essere considerate come la somma delle proprietà dei precursori, in un ibrido l'insieme risulta essere diverso e molto spesso migliore rispetto ai precursori.

In questo contesto, la chimica dei materiali ibridi rappresenta un campo molto interessante di ricerca e sviluppo, poiché permette una grande varietà di nuove applicazioni. [6]

Nei **nano**materiali ibridi, definiti come la classe di materiali costituiti da due o più componenti in cui almeno uno di essi deve avere una dimensione nel regime della nanoscala [5], proprio grazie alle nanodimensioni, gli effetti sinergici causati dall'intimo contatto possono diventare un fattore dominante e caratterizzante [7].

I singoli componenti di una nanostruttura ibrida possono essere scelti tra una vasta gamma di materiali: metalli, ossidi metallici, calco-

genuri metallici, polimeri, materiali carboniosi, …, mentre la loro sintesi può essere fatta in diversi modi, risultando in dimensioni, forme e composizioni differenti [8]. Ad esempio, nel caso di nanoibridi metallo-metallo è possibile trovare sistemi tipo lega o *core-shell*, tipo *Janus* (ossia caratterizzati da una doppia natura), sintetizzati tramite approcci *bottom-up* (dal basso) o *top-down* (verso il basso) [9]. Per quanto riguarda altri tipi di materiali, tra gli ibridi più diffusi si possono trovare i nanomateriali organico-inorganici, dove il controllo della natura chimico-fisico e l'estensione delle interfacce ibride risultano di primaria importanza in quanto regolano proprietà quali la trasparenza, l'omogeneità chimica e la stabilità del materiale [10-13].

Una delle caratteristiche essenziali di questi ultimi materiali sono le loro dimensioni su scala nanometrica, che generano effetti di dimensione sfruttabili in fase di progettazione delle proprietà del materiale. Gli effetti del confinamento spaziale diventano sempre più importanti quando le dimensioni di un nanocristallo diminuiscono al di sotto di un certo limite critico, portando a una struttura elettronica dipendente dalla dimensione e dalla forma [10].

1.3 Nanomateriali ibridi ed applicazioni elettroanalitiche

Considerando le sopracitate caratteristiche e prestazioni dei materiali nanostrutturati e dei materiali ibridi, negli ultimi anni la Ricerca Scientifica si sta volgendo verso l'uso di una combinazione dei due aspetti, studiando materiali ibridi nanostrutturati, in modo da esplorare gli effetti sinergici della combinazione dei due aspetti. Tra le innumerevoli applicazioni, l'elettrochimica ed in particolar modo l'elettroanalisi possono beneficiare moltissimo di queste scoperte, permettendo di creare dispositivi a basso costo e di facile costruzione sempre più sensibili e con limiti di rilevabilità sempre più bassi, caratterizzati anche da alta selettività ed alcuni casi specificità, per l'analisi di sostanze anche presenti in tracce (ppm, ppb, ppt) in ambito medico/diagnostico e/o ambientale, come nella rilevazione di inquinanti emergenti. Questi dispositivi, data la loro natura miniaturizzata e a basso costo con produzione di massa, possono essere facilmente utilizzati come "sensori d'allarme", posizionati in remoto su reti idriche o su impianti industriali, permettendo un monitoraggio continuo ed affidabile, a lungo termine.

2. Il sistema nanoibrido argento-titania

Il dispositivo con nanoparticelle di argento (*Fig. 1*) è stato preparato deponendo uno strato di silice su vetro conduttivo tramite una procedura sol-gel. La silice è stata funzionalizzata con un silossano (APTES, (3-aminopropyl)triethoxysilane) con terminazione positiva in modo da immobilizzare le nanoparticelle di argento, cariche negativamente. Il tutto è stato ricoperto con uno strato di biossido di titanio precedentemente preparato tramite sintesi sol-gel e trattato ad alta temperatura (450°C).

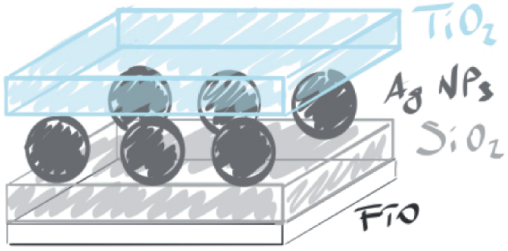


Fig. 1. Schema del dispositivo con nanoparticelle di argento.

Ogni strato del dispositivo è stato accuratamente pensato e progettato per conferire le proprietà finali desiderate. Il supporto su cui si costruisce tutto il dispositivo è un elettrodo otticamente trasparente (OTE) basato su un vetro conduttivo poiché ricoperto di un sottile strato di ossido di stagno fluorurato (FTO), resistente alla calcinazione ad alte temperature (450° C). Lo strato di silice (SiO_2) viene messo perché, in combinazione con APTES, che poi è in grado di agganciare le nanoparticelle metalliche di argento (cariche negativamente) è possibile così creare un "random assembly of nanoelectrodes", ossia un assemblaggio casuale di nanoelettrodi che però risultano statisticamente ordinati e ben distanziati. Questa caratteristica presenta grandi vantaggi in applicazioni elettroanalitiche.

Lo strato finale è costituito da TiO_2 (in fase anatasio, con tracce di brookite) che protegge le nanoparticelle metalliche dall'ossidazione e dal possibile rilascio in soluzione [14] e conferisce fotoattività e proprietà autopulenti al dispositivo. Maggiori dettagli sulla preparazione

del dispositivo e sulle sue caratterizzazioni chimico-fisiche sono reperibili in [15].

Il dispositivo è stato utilizzato per la rilevazione di dopamina, norepinefrina e serotonina con limiti di rilevabilità di 0.03, 0.04 e 0.04 μM e fattori di recupero apparenti del 100-99%, valori migliori di quelli richiesti dalla diagnostica clinica nelle analisi di routine [16]. Studi di interferenza hanno dimostrato il suo possibile utilizzo anche in presenza di acido ascorbico e acido urico, presenti solitamente in concentrazione da 100 a 1000 volte superiore rispetto ai neurotrasmettitori [16], anche sfruttando la differente carica elettrostatica della titania rispetto agli interferenti, all'opportuno valore di pH. L'utilizzo con buoni risultati (LOD = 0.1 μM) anche in matrici reali (liquor, siero e urina), ha infine dimostrato l'estrema versatilità analitica del dispositivo. [16]

Infine, la fotorinnovabilità della superficie elettrodica del sistema, che può sporcarsi o passivarsi, a causa della matrice reale o della stessa reazione elettroanalitica che può formare sottoprodotti che si adsorbono sull'elettrodo, è stata accuratamente valutata utilizzando irraggiamenti con luce UV di diversa potenza, forma e costo. Il ripristino totale della superficie viene raggiunto dopo poco tempo di irraggiamento, aprendo la strada verso un sistema di monitoraggio integrato remoto o per analisi in flusso [15, 17].

Per approfondire maggiormente le ragioni delle proprietà peculiari del dispositivo, è stato condotto un lavoro concertato elettrochimico-teorico [18] che ha dimostrato che i due materiali costituenti il dispositivo (AgNP e TiO_2) sono in intimo contatto tra loro, creando una interfase in cui le proprietà elettrochimiche dell'argento vengono drasticamente modificate rispetto a quelle attese.

I risultati del doppio approccio (caratterizzazione voltammetrica combinata con spettroscopia di impedenza e calcoli DFT) hanno portato ad evidenziare come in questo sistema avvenga una donazione di elettroni dagli atomi di Ag alla titania. Gli atomi di Ag più vicini alla titania assumono una parziale carica positiva. E' possibile quindi evidenziare una vera e propria zona (quella più vicina alla titania) in cui l'elettrodo, che si ossida durante la scansione anodica del potenziale, rimane stabile nella forma ossidata (grazie alla presenza della

titania) e non viene ridotto nella scansione inversa (*Fig. 2*), fino al raggiungimento di un equilibrio, in cui si osserva il ciclo ossidazione/riduzione solo per gli atomi di argento più lontani da TiO_2. Per ripristinare lo strato di Ag più vicino alla titania, ossia per portarlo al suo stato di ossidazione 0, si può operare con l'irraggiamento del dispositivo con luce UV, ottenendo il completo ripristino del picco voltammetrico (1 in *Fig. 2A*) posizionato a circa 0.2 V (SCE).

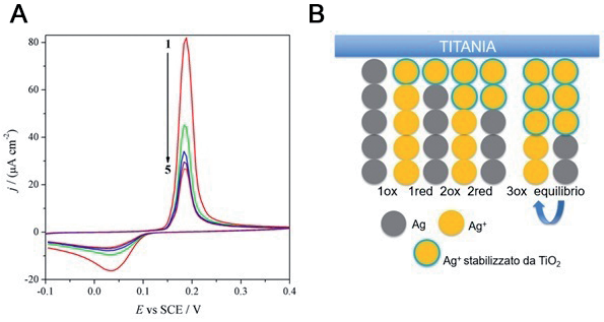


Fig. 2. A) Curve voltammetriche del dispositivo a base di Ag, registrate a 0.1 V s^{-1}*, in* $NaClO_4$ *0.1 M, a temperatura ambiente e B) schema del relativo processo ciclico di ossidazione/riduzione dell'argento (stabilizzato nella sua forma ossidata alla e dalla interfase con la titania), fino al raggiungimento di un equilibrio "dinamico".*

La presenza di questa interfase di intimo contatto tra Ag e TiO_2 è anche confermata da studi di spettroscopia di impedenza in cui è necessario aggiungere nel circuito equivalente che descrive il dispositivo e interpola i dati sperimentali, un circuito RC in parallelo nel caso dell'ibrido, proprio relativo a questa interfase [18].

Questo fenomeno di trasferimento di elettroni è stato quindi ulteriormente confermato da esperimenti di fotocorrente (*Fig. 3*), in cui risulta evidente che, a parità di potenziale di eccitazione, l'ibrido Ag/TiO_2 mostra una densità di fotocorrente inferiore rispetto al solo biossido di titanio, poiché parte degli elettroni fotoprodotti vengono utilizzati per ridurre gli atomi di argento più lontani dalla titania. Questo effetto è maggiormente apprezzabile ad alti potenziali di eccitazione (es. 1 V, SCE) dove il numero di atomi di argento ossidato è superiore, rispetto a potenziali di eccitazione più bassi (0.12 V, SCE).

In quest'ultimo caso, solo una piccola parte di fotoelettroni viene utilizzata per la riduzione degli elettroni [18].

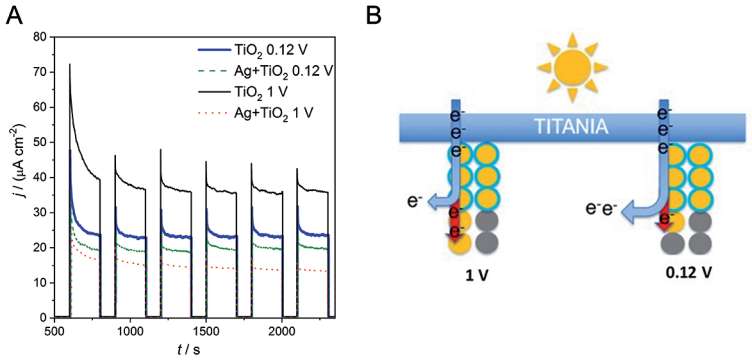


Fig. 3. A) Spettri di fotocorrente del dispositivo a base di argento, registrati a due diversi potenziali (0.12 e 1 V, SCE), sotto irraggiamento con lampada UV ad alogenuro di ferro (280-400 nm, densità di potenza effettiva di 23 mW cm^{-2}) e B) schema del relativo processo di riduzione dell'argento da parte dei fotoelettroni generati all'interfase. Figura adattata da riferimento [18].

Tutte queste evidenze sperimentali e teoriche hanno permesso di fare luce sulle peculiari proprietà elettroniche di questa eterogiunzione, caratterizzata dall'intimo contatto di Ag e TiO_2.

3. Il sistema nanoibrido oro-titania

Il dispositivo (*Fig. 4*) con le nanoparticelle di oro (10 nm) precedentemente sintetizzate con tecnica sol-gel, è stato preparato in maniera simile al precedente (vedi paragrafo 2). Le uniche differenze sono l'assenza dello strato di silice e l'aggiunta di un secondo silossano (in rapporto 36:4 con APTES) che ha permesso una funzionalizzazione diretta. Il vetro conduttivo è stato infatti direttamente funzionalizzato con l'aggiunta di un silossano con terminazione contenente zolfo (MPTMS: (3-mercaptopropyl) trimethoxysilane), che ha permesso l'immobilizzazione delle nanoparticelle d'oro, aventi grossa affinità per tale elemento.

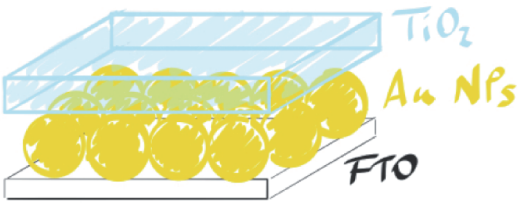


Fig. 4. Schema del dispositivo con nanoparticelle di oro.

Rispetto all'ibrido Ag/TiO_2, le voltammetrie cicliche (*Fig. 5A*) del dispositivo Au/TiO_2, mostrano una maggiore somiglianza con quelle del supporto (FTO) modificato con le sole nanoparticelle di Au, evidenziando come le proprietà intrinseche dell'oro siano preservate. Inoltre, l'area del picco di ossidazione e quella di riduzione sono simile e non sbilanciate come nel caso precedente. La diversa forma del plateau di ossidazione potrebbe essere dovuta a una diversa disposizione delle nanoparticelle d'oro, che nel caso del dispositivo sono meno aggregate e "più protette" dalla presenza dello strato di titania protettivo. Si ipotizza pertanto che in questa struttura non avvenga alcun trasferimento di carica tra i due materiali del composito, come anche confermato dagli esperimenti di fotocorrente (*Fig. 5B*).

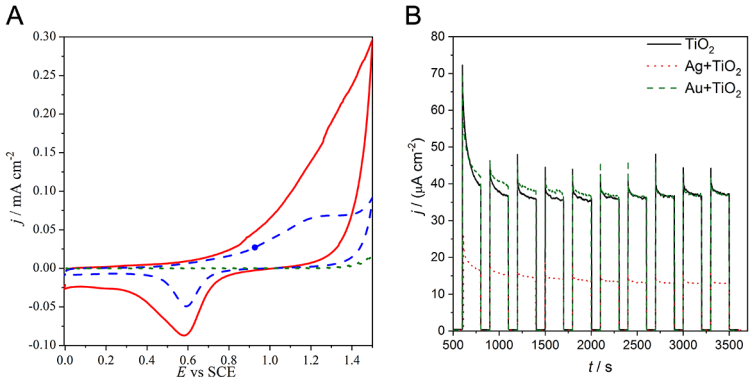


Fig. 5. A) Curve voltammetriche del dispositivo a base di Au (curva blu tratteggiata)*, di AuNPs* (curva rossa continua)*, di* TiO_2 (curva verde, tratto piccolo)*, registrate a 0.1 V* s^{-1}*, in* $NaClO_4$ *0.1 M, a temperatura ambiente e B) spettri di fotocorrente del dispositivo a base di oro (paragonato con quello a base di argento), registrati a 1 V (SCE), sotto irraggiamento con lampada UV ad alogenuro di ferro (280-400 nm, densità di potenza effettiva di 23 mW* cm^{-2}*).*

L'ibrido e il TiO_2 semplice mostrano la stessa densità di fotocorrente, il che significa che nel caso dell'ibrido tutti gli elettroni coinvolti sono utilizzati per la generazione della fotocorrente anodica e non è quindi presente alcuna donazione elettronica dall'oro alla titania. Anche gli esperimenti di spettroscopia di impedenza (*Fig. 6A*) evidenziano come l'ibrido sia più simile al sistema con le sole nanoparticelle d'oro, che non si stanno adattando alla titania come precedentemente osservato nel caso dell'argento. I circuiti equivalenti utilizzati per interpolare i dati sperimentali (*Fig. 6B*), non presentano infine, il terzo circuito RC parallelo, necessario nel dispositivo precedente, denotando l'assenza dell'interfase metallo/titania nel caso dell'oro.

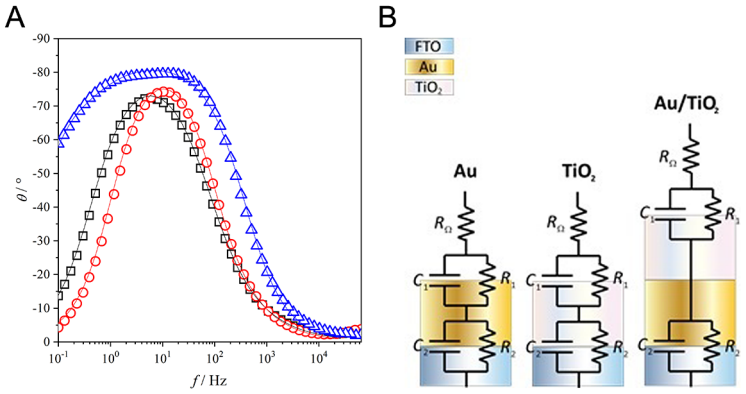


Fig. 6. A) Curve di impedenza (Bode Plot) del dispositivo a base di Au (curva rossa a pallini)*, di AuNPs* (curva nera a quadretti)*, di* TiO_2 (curva blu a triangoli)*, registrati a -0.1 V (SCE) e B) relativi circuiti equivalenti per il fitting (linee continue) delle curve sperimentali.*

3.1 Applicazioni elettroanalitiche del dispositivo Au/TiO_2

Anche questo secondo dispositivo a base di nanoparticelle di oro è stato applicato nella rilevazione elettroanalitica di Diclofenac (2- [2-(2,6-dichloroanilino) fenil] acido acetico), un farmaco antinfiammatorio non steroideo ampiamente usato per il trattamento del dolore e delle malattie infiammatorie. Decimo farmaco più venduto in Italia nel 2018, è classificato come contaminante emergente e nel 2013 è stato inserito, insieme ad altri prodotti farmaceutici, in un elenco di controllo

da parte dell'Unione Europea nella "Direttiva quadro sulle acque europee", contenente tutti quei composti emergenti che dovrebbero essere dannosi per l'ambiente. Secondo questo documento, è stato fissato il seguente standard di qualità ambientale per Diclofenac: le concentrazioni massime consentite sono di 0,1 g / L in acque dolci e 0,01 mg / L in acque marine (2013/39 / CE, Direttiva Europea, Diclofenac EQS Dossier 2011). La presenza di diclofenac nell'ambiente è principalmente dovuta alla sua escrezione, insieme ai suoi metaboliti, da parte di esseri umani e animali. Questi composti raggiungono l'impianto di trattamento delle acque reflue dove possono seguire due percorsi diversi: (i) da un lato, parte di Diclofenac viene assorbita dai fanghi che finendo in discarica, permettono a Diclofenac di entrare nelle acque superficiali per percolazione; (ii) dall'altro lato, Diclofenac rimane nelle acque reflue e viene scaricato direttamente nelle acque superficiali. Pertanto, anche i sistemi convenzionali di trattamento delle acque reflue non sono in grado di abbattere la presenza di questo inquinante con una buona percentuale di rimozione [19,20].

Sfruttando l'affinità delle nanoparticlle di Au per Diclofenac [21], abbiamo testato la capacità del dispositivo di rilevare questa molecola, sfruttando poi la presenza di TiO_2, al fine di rinnovare la superficie del dispositivo dopo l'analisi, che generalmente porta alla formazione di sottoprodotti che sporcano l'elettrodo, rendendolo inutilizzabile per le rilevazioni successive. [15, 16, 21]

In particolare, in *Fig. 7*, viene mostrato il picco voltammetrico per la rilevazione di Diclofenac ad una concentrazione di $7\ 10^{-5}$ M, in tampone fosfato a pH = 4. Dopo la determinazione dell'analita, l'elettrodo si sporca, a causa dei sottoprodotti della reazione elettrodica durante la rilevazione [21, 22]. Tale sporcamento non è risolvibile per semplice lavaggio dell'elettrodo, come dimostra la permanenza del picco voltammetrico dopo il lavaggio con acqua.

Solo dopo trattamento con luce UV (irraggiamento di un'ora con lampada UV a 365 nm posta a 20 cm dal supporto, anche in presenza della soluzione e del vetro del contenitore in cui è alloggiato l'elettrodo), la situazione elettrodica originaria viene ripristinata. Anche l'utilizzo di un'altra fonte luminosa più facilmente reperibile e maneggiabile (LED, con una lunghezza d'onda di emissione a 390 nm), ha portato al

perfetto recupero della superficie elettrodica, confermando l'applicabilità di questa procedura per il mantenimento in perfetta pulizia dell'elettrodo di lavoro ed aprendo quindi la strada per applicazioni in monitoraggio continuo e remoto.

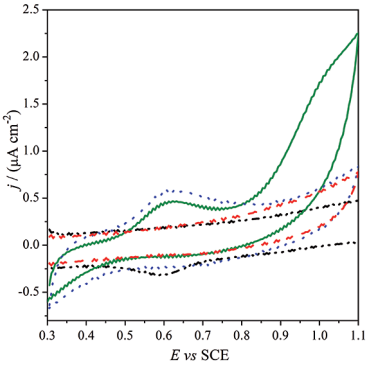


Fig. 7. Curve voltammetriche del dispositivo a base di oro, nelle varie fasi di "sporcamento" e "pulizia": prima registrazione in assenza di analita (curva nera tratto-punto); registrazione in presenza di Diclofenac 7 10^{-5} M (curva verde continua); registrazione dopo lavaggio del dispositivo per semplice risciacquo in acqua (curva blu a punti); registrazione dopo recupero fotocatalitico sotto irraggiamento per un'ora (lampada UV con (max) = 365 nm) (curva rossa tratteggiata). Curve registrate a 0.1 V s^{-1}, in tampone fosfato a pH = 4, a temperatura ambiente.

4. Conclusioni

L'uso di nanomateriali ibridi, caratterizzato da comportamenti e caratteristiche molto interessanti, ha aperto la strada verso applicazioni promettenti in molti campi, tra cui l'elettroanalisi, che si sta molto avvantaggiando di questi sistemi per la realizzazione di dispositivi sensoristici caratterizzati da prestazioni elevate.

Le nano-eterogiunzioni, che costituiscono il cuore di questi sistemi, se opportunamente progettate, permettono di migliorare le funzionalità sinergiche e consentire di ottenere "nuovi materiali " caratterizzati da proprietà chimico-fisiche macroscopiche che non sono semplicemente la somma di quella dei rispettivi precursori, ma vanno oltre il semplice materiale.

Tuttavia, la ricerca su tali materiali è spesso dominata da procedure di prova ed errore, mentre risulta essenziale una profonda comprensione atomistica dei fenomeni all'interno della regione di giunzione, guidata da solide evidenze sperimentali tratte da tecniche e procedure diverse, per giungere ad una progettazione appropriata del dispositivo finale.

Questo è l'approccio adottato nel lavoro qui presentato, che può o meglio dovrebbe essere utilizzato anche in altri casi dove il contributo delle eterogiunzioni risulta cruciale.

Ringraziamenti

In memoria del caro Professor Torquato Mussini.

BIBLIOGRAFIA

1. Feynman, R. P. Caltech Engineering and Science. *23:5*. 1960, pp 22–36
2. Compton, R. G., Banks, C. E. *Understanding Voltammetry*; Imperial College Press, London, 2001.
3. Welch, C. M., Compton, R. G. The Use of Nanoparticles in Electroanalysis: A Review. Anal. Bioanal. Chem. 2006, 384, 601–619.
4. Campbell, F. W., Compton, R. G. The Use of Nanoparticles in Electroanalysis: An Updated Review. *Anal. Bioanal. Chem.* **2010**, *396* (1), 241–259.
5. PAC, 2007, 79, 1801. (Definitions of terms relating to the structure and processing of sols, gels, networks, and inorganic-organic hybrid materials (IUPAC Recommendations 2007)) on page 1811
6. Sanchez, C., Belleville, P., Popalld, M., Lionel, N. Applications of Advanced Hybrid Organic–Inorganic Nanomaterials: From Laboratory to Market. *Chem. Soc. Rev.* **2011**, *40* (2), 588–595.
7. Costi, R., Saunders, A. E., Banin, U. Colloidal Hybrid Nanostructures: A New Type of Functional Materials. *Angew. Chemie - Int. Ed.* **2010**, *49* (29), 4878–4897
8. Huang, X., Tan, C., Yin, Z., Zhang, H. 25th Anniversary Article: Hybrid Nanostructures Based on Two-Dimensional Nanomaterials. *Adv. Mater.* **2014**, *26* (14), 2185–2203.
9. Toshima, N., Yonezawa, T. Bimetallic Nanoparticles: Novel Materials for Chemical and Physical Applications. **1998**, 1179–1201.
10. Donegá, C. D. M. Synthesis and Properties of Colloidal Heteronanocrystals. *Chem. Soc. Rev.* **2011**, *40* (3), 1512–1546.

11. Ingrosso, C., Bianco, G. V., Corricelli, M., Comparelli, R., Altamura, D., Agostiano, A., Striccoli, M., Losurdo, M., Curri, M. L., Bruno, G. Photoactive Hybrid Material Based on Pyrene Functionalized PbS Nanocrystals Decorating CVD Monolayer Graphene. *ACS Appl. Mater. Interfaces* **2015**, *7* (7), 4151–4159.
12. Corricelli, M., Altamura, D., De Caro, L., Guagliardi, A., Falqui, A., Genovese, A., Agostiano, A., Giannini, C., Striccoli, M., Curri, M. L. Self-Organization of Mono- and Bi-Modal PbS Nanocrystal Populations in Superlattices. *CrystEngComm* **2011**, *13* (12), 3988–3997.
13. Ingrosso, C., Corricelli, M., Bettazzi, F., Konstantinidou, E., Bianco, G. V., Depalo, N., Striccoli, M., Agostiano, A., Curri, M. L., Palchetti, I. Au Nanoparticle: In Situ Decorated RGO Nanocomposites for Highly Sensitive Electrochemical Genosensors. *J. Mater. Chem. B* **2019**, *7* (5), 768–777.
14. Cloake, S.J., Toh, H.S., Lee, P.T., Salter, C., Johnston, C., Compton, R.G. Anodic stripping voltammetry of silver nanoparticles: Aggregation leads to incomplete stripping. *Chemistry Open* **2015**, 4(1), 22-26.
15. Soliveri, G., Pifferi, V., Panzarasa, G., Ardizzone, S., Cappelletti, G., Meroni, D., Sparnacci, K., Falciola, L. Self-Cleaning Properties in Engineered Sensors for Dopamine Electroanalytical Detection. *Analyst* **2015**, 140 (5), 1486–1494.
16. Pifferi, V., Soliveri, G., Panzarasa, G., Cappelletti, G., Meroni, D., Falciola, L. Photo-Renewable Electroanalytical Sensor for Neurotransmitters Detection in Body Fluid Mimics. *Anal. Bioanal. Chem*. **2016**, 408 (26), 7339–7349.
17. Pifferi, V., Soliveri, G., Panzarasa,G., Ardizzone, S., Cappelletti, G., Meroni, D., Falciola, L. Electrochemical sensors cleaned by light: a proof of concept for on site applications towards integrated monitoring systems. RSC Advances 2015, 5, 71210-71214.
18. Di Liberto, G., Pifferi, V., Lo Presti, L., Ceotto, M., Falciola, L. Atomistic Explanation for Interlayer Charge Transfer in Metal-Semiconductor Nanocomposites: The Case of Silver and Anatase. *J. Phys. Chem. Lett.* **2017**, 8 (21), 5372–5377.
19. Lonappan, L., Brar, S. K., Das, R. K., Verma, M., Surampalli, R. Y. Diclofenac and Its Transformation Products: Environmental Occurrence and Toxicity - A Review. *Environ. Int.* **2016**, 96, 127–138.
20. Mohapatra, D. P., Kirpalani, D. M. Advancement in Treatment of Wastewater: Fate of Emerging Contaminants. *Can. J. Chem. Eng*. **2019**, (October), 2621–2631.
21. Afkhami, A., Bahiraei, A., Madrakian, T. Gold Nanoparticle/Multi-Walled Carbon Nanotube Modified Glassy Carbon Electrode as a Sensitive Voltammetric Sensor for the Determination of Diclofenac Sodium. *Mater. Sci. Eng. C* **2016**, 59, 168–176.
22. Yang, X., Wang, F., Hu, S. Enhanced Oxidation of Diclofenac Sodium at a Nano-Structured Electrochemical Sensing Film Constructed by Multi-Wall Carbon Nanotubes-Surfactant Composite. *Mater. Sci. Eng. C* **2008**, 28 (1), 188–194.

Scienze naturali
Ecologia

Istituto Lombardo (Rend. Scienze) 153, 207-240 (2019)

LA RICOLONIZZAZIONE DELL'ANTARTIDE: STORIE DI PINGUINI E FOCHE, SENTINELLE DEI CAMBIAMENTI CLIMATICI NEGLI ULTIMI 10 MILA ANNI

CARLO BARONI (*)

Nota presentata dal m.e. Giuseppe Orombelli
(Adunanza del 1 febbraio 2018)

SUNTO. – Nel Pleistocene superiore l'espansione del sistema glaciale antartico ha invaso la fascia costiera costringendo pinguini, foche e altri animali che nidificavano nelle aree deglaciate a migrare verso aree rifugio poste a latitudini più settentrionali. La deglaciazione che è seguita all'ultimo periodo di massima espansione glaciale, nel settore del Mare di Ross, si è completata intorno a 8000 anni fa, liberando le coste antartiche che sono state ri-colonizzate da pinguini di Adelia ed elefanti marini. Queste due specie hanno convissuto, anche alternandosi, nelle aree deglaciate della fascia costiera per diversi millenni. Nei suoli ornitogenici delle colonie di pinguini abbandonate e nei siti di riproduzione degli elefanti marini recentemente scoperti sulle spiagge emerse è conservato un importante patrimonio di resti organici che ci permette di studiare diversi aspetti della storia ambientale recente. Le fasi della colonizzazione e della coesistenza tra pinguini ed elefanti marini sono molto articolate e riflettono le variazioni climatico-ambientali che hanno caratterizzato lo scenario antartico per diverse migliaia di anni. Pinguini ed elefanti marini, sensibili sentinelle dell'ecosistema antartico stanno anche adesso recitando un ruolo di rilievo nello scenario dell'Oceano Meridionale, interpretando la parte di veri e propri campanelli d'allarme dei cambiamenti climatici.

***

ABSTRACT. – *The re-colonization of Antarctica: stories of penguins and seals, sentinels of climate change in the last ten thousand years.* During the Late Pleistocene the

(*) Istituto Lombardo Accademia di Scienze e Lettere, Milano. Dipartimento di Scienze della Terra, Università di Pisa. CNR-IGG Istituto di Geoscienze e Georisorse, Pisa, Italy. E-mail: carlo.baroni@unipi.it

Antarctic glacial system expanded and advanced on the coastal margin of the continent forcing penguins, seals and other animals nesting in the deglaciated areas to migrate toward refuge areas located at northern latitudes. The deglaciation that followed the last glacial maximum, in the Ross Sea area, was accomplished around 8000 years ago, freeing the Antarctic coasts that were re-colonized by Adélie penguins and elephant seals. These two species have coexisted, even alternating, in the deglaciated areas of the coastal strip for several millennia. In the ornithogenic soils of the abandoned penguin colonies and in the breeding sites of the elephant seals, recently discovered on the Holocene raised beaches, a precious heritage of organic remains is preserved, allowing us to study different aspects of recent environmental history. The phases of colonization and coexistence between penguins and elephant seals are very complex and reflect the climatic-environmental variations that have characterized the Antarctic scenario for several thousand of years. Penguins and elephant seals, sensitive sentinels of the Antarctic ecosystem, are now playing an important role in the Southern Ocean scenario, interpreting the dramatic role of bellwether of climate change.

1. Introduzione

Nelle regioni polari, dominate dalla criosfera, i processi glaciali e periglaciali costituiscono i principali e più efficaci agenti morfogenetici, in grado di scolpire spettacolari forme del rilievo che rendono unici i paesaggi di queste regioni. I sistemi glaciali che si sviluppano attualmente sull'Antartide (circa $13.5 * 10^6$ km^2) e sulla Groenlandia (circa $1.7 * 10^6$ km^2) comprendono il 96% della superficie glacializzata del nostro Pianeta, trattenendo circa il 99% del volume di ghiaccio totale (WGMS). L'enorme quantità di ghiaccio immagazzinata nelle calotte antartiche (circa $26.6 * 10^6$ km^3) e in Groenlandia (circa $2.85 * 10^6$ km^3) è in grado di condizionare il livello medio del mare a scala globale: l'eventuale fusione delle calotte rilascerebbe una quantità d'acqua in grado di determinare un innalzamento del livello del mare di oltre 68 m (61 m circa dalla fusione delle calotte antartiche, 7.2 m dalla fusione della calotta groenlandese; Morlighem *et al.*, 2020).

Oltre a condizionare le variazioni del livello del mare, i sistemi glaciali che si sviluppano nelle regioni polari, guidano la circolazione oceanica e il relativo trasferimento di calore attraverso l'attivazione o la disattivazione della circolazione termoalina (*conveyor belt*) e controllano i livelli di gas serra attraverso i processi di emissione/assorbimento di CO_2 e CH_4. Le regioni polari svolgono un ruolo di primo piano nel bilancio energetico planetario da almeno 34 Ma (Eocene-Oligocene),

cioè dall'epoca in cui si sviluppò la prima calotta (cenozoica) sull'Antartide Orientale (Zachos *et al.*, 2001).

La più importante conseguenza delle strette relazioni tra cambiamenti climatici e comportamento della criosfera è che le risposte dei processi glaciali e periglaciali hanno lasciato evidenti tracce sia nelle forme del rilievo conservate nelle aree deglaciate dell'Antartide, sia nei depositi superficiali del continente, sia nei sedimenti marini attorno al continente di ghiaccio, che quindi rappresentano veri e propri archivi di informazioni paleoambientali. Informazioni dirette sulla storia glaciale pleistocenica e sulla sensitività dei sistemi glaciali ai cambiamenti climatici sono peraltro fornite dalle carote di ghiaccio perforate nelle calotte glaciali (EPICA-Community-Members, 2004, 2006; Jouzel *et al.*, 2007; Lambert et al, 2008; Masson-Delmotte *et al.*, 2010). Questi preziosi archivi naturali forniscono informazioni fino a 800.000 anni fa anche sulla storia dell'atmosfera terrestre ma, potenzialmente, potrebbero estendersi fino a 1.5 milioni di anni fa (Fisher *et al.*, 2013).

La lunga storia glaciale dell'Antartide è stata caratterizzata da una complessa sequenza di fasi di espansione e di contrazione delle calotte, ben riconoscibili almeno dal Miocene (Barker and Cooper, 1997; Anderson, 1999; Bart *et al.*, 2000; Cape Roberts Science Team, 2000; McKay *et al.*, 2009; Naish *et al.*, 2009). In particolare, i ripetuti ispessimenti delle calotte antartiche sono stati accompagnati da espansioni sulla piattaforma continentale (*Fig. 1*) fino a posizioni massime simili, o addirittura più avanzate, rispetto a quelle raggiunte nell'Ultimo Massimo Glaciale -UMG- (Denton e Hughes, 1981; Denton *et al.*, 1984, 1991; Anderson, 1999; Nakada *et al.*, 2000; Huybrechts, 2002; Denton and Hughes, 2002; Pollard and DeConto, 2009; Bentley *et al.*, 2016).

Tutti gli archivi paleoclimatici disponibili e le evidenze geomorfologiche e geologiche concordano nel riconoscere, durante il Pleistocene, ripetute fasi di avanzata e di ritiro delle calotte glaciali antartiche, accompagnate da variazioni del livello del mare, con abbassamenti fino a circa 125-130 m rispetto al livello attuale, come avvenuto nell'UMG (Fairbanks, 1989; Clark e Mix, 2002).

Queste imponenti variazioni climatiche e ambientali hanno avuto un notevole impatto sull'*habitat* costiero, modificando *in primis* la conformazione delle aree deglaciate potenzialmente colonizzabili da pinguini, foche e altre specie animali.

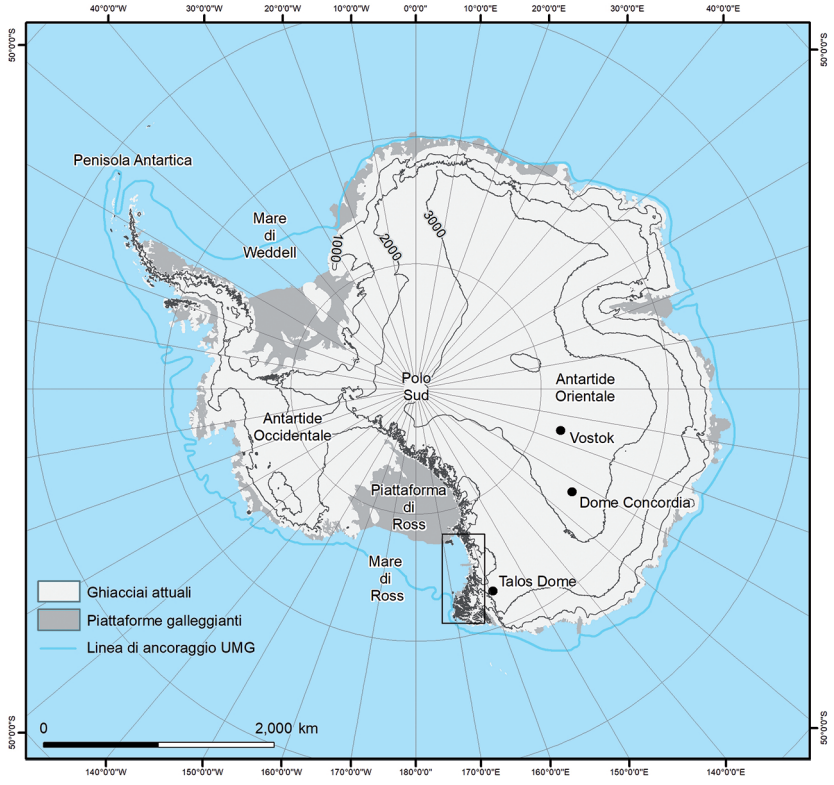


Fig. 1. L'Antartide. In grigio chiaro i ghiacciai attuali, in grigio scuro le piattaforme di ghiaccio (da Fretwell et al., 2013); in ciano la linea di galleggiamento nell'Ultimo Massimo Glaciale, intorno a 20.000 anni fa. (Da Bentley *et al.*, 2014). Il riquadro indica l'area rappresentata nella *Fig. 5*. Le curve di livello derivano dall'Antartcic Digital Database, modificate.

2. Sentinelle biologiche dei cambiamenti climatici nelle regioni polari

Un'idea fondamentale della biologia evoluzionistica è che gli esseri viventi si adattino agli ambienti in cui vivono, ma gli ambienti sono instabili e le popolazioni, sollecitate dalle nuove condizioni che si vengono a creare, rispondono a questi cambiamenti in modo diverso (Lambert *et al.*, 2010). In particolare, le popolazioni possono: 1) adat-

tarsi alle nuove condizioni, 2) migrare per allontanarsi da ambienti non più idonei alla loro sopravvivenza, 3) approfittare dei nuovi *habitat* che si stanno sviluppando altrove, 4) regolare il numero di individui che costituisce quella popolazione, 5) cercare una diversa combinazione tra le possibilità sopra elencate.

Il modo più semplice che molte specie animali adottano per reagire ai cambiamenti climatici e ambientali, piuttosto che adattarsi alle nuove condizioni, è semplicemente quello di migrare (se possibile), per rimanere nelle condizioni di vita preferite seguendo l'ambiente ideale che si sta spostando altrove (a latitudini più alte oppure a quote più elevate). Talvolta la possibilità di migrare è preclusa dalla definitiva scomparsa dell'*habitat* ideale, che può avvenire in tempi troppo rapidi per consentire la migrazione in aree rifuglio (perché troppo distanti o semplicemente ignote o isolate e quindi irraggiungibili).

Durante le fasi glaciali, le specie che vivevano negli ambienti polari hanno avuto la possibilità di espandersi andando ad occupare (come ospiti freddi) ambienti posti a latidudini più basse. In realtà, i cambiamenti climatici e ambientali condizionano la biodiversità inducendo reazioni sia ecologiche sia evolutive e le risposte possono essere differenti ai diversi livelli della catena trofica, secondo meccanismi anche complessi e non sempre definibili a priori. Questo meccanismo è ben spiegato ed è chiaramente comprensibile se pensiamo alla dinamica delle specie mutualmente esclusive che si sono alternate nel corso del Pleistocene come ospiti caldi e freddi negli ambienti marini di media latitudine (ad esempio, lo *Strombus bubonius* e l'*Arctica islandica* nel Mar Mediterraneo). Un altro esempio, molto caratteristico, è il "pinguino dell'Emisfero Settentrionale" (*Alca impennis* Linnaeus, 1758; Sin. *Pinguinus impennis*) un alcide oggi estinto (*Fig. 2*). Seguendo le fredde correnti atlantiche, nelle fasi glaciali si spinse molto a Sud fin sulle coste della Florida lungo il margine orientale del continente americano; sul versante europeo, attraverso lo stretto di Gibilterra, l'Alca si stabilì nel Mediterraneo come ospite freddo, predato dall'uomo durante il Paleolitico medio e superiore (Ascenzi e Segre, 1971; Bonfiglio *et al.*, 1986). Dopo l'UMG, seguendo la progressiva deglaciazione, l'Alca si ritirò nell'Atlantico settentrionale dove fu oggetto di progressiva predazione da parte dell'uomo per l'intero Olocene. L'ultima colonia sull'Isola di Funk (Canada) vide il progressivo declino già nel XVI secolo, fino all'estinzione avvenuta tra il XVIII e la prima metà del XIX secolo (Grieve, 1885; Fuller, 1999; Serjeantson, 2001).



Fig. 2. Alca impennis *Linn.; Sin. Pinguinus impennis un alcide estinto equivalente, nell'emisfero settentrionale, ai pinguini, ospite freddo nel Mar Mediterraneo durante le fasi glaciali.* (Collector cards from Tiere der Urwelt, Animals of the Prehistoric World, Serie Ia, datato1916). Creato da Heinrich Harder.

Da quanto sopra riportato emerge l'importanza di distinguere gli effetti dei cambiamenti climatici a seconda che si consideri il passaggio da una fase interglaciale a una fase glaciale o viceversa. La questione non è di poco conto, dato che nel primo caso il passaggio avviene tramite una più lenta transizione alle nuove condizioni, con una progressione che richiede tempi relativamente lunghi. Al contrario, è ampiamente riconosciuto che la deglaciazione o, meglio, il passaggio da fasi glaciali a fasi interglaciali, avvenga in tempi molto rapidi. Pertanto, gli ecosistemi e gli *habitat* specifici delle regioni fredde sono più vulnerabili all'impatto dei cambiamenti climatici durante le fasi di riscaldamento rispetto alla condizione inversa che si verifica durante le fasi di raffreddamento. Un esempio noto è quello dei mammuth e di altre specie che si sono estinte nella transizione tra l'UMG e l'Olocene.

Attualmente, molte specie sono a rischio di estinzione se il futuro riscaldamento globale continuerà con questi ritmi e molte prove convincenti al riguardo sono riportate per quasi tutte le componenti degli ecosistemi marini, dalle comunità planctoniche ai grandi mammiferi (IPCC, 2007, 2013; Simmonds and Isaac, 2007; Wassmann *et*

al., 2011). Il caso più drammaticamente noto è rappresentato dall'orso polare (*Ursus maritimus*) che, nell'Artico, è costretto ad adattarsi rapidamente alle nuove condizioni oppure dovrà necessariamente estinguersi. D'altro canto, molte specie subartiche troveranno *habitat* più estesi da occupare per effetto della progressiva riduzione dell'ambiente polare e la contestuale estensione di ambienti a clima più mite.

Nell'Antartico, a partire da 34 Myr fa (limite Eocene-Oligocene), dopo la formazione della calotta continentale e del Fronte Polare, gli organismi si sono adattati a un ambiente estremamente freddo e isolato, caratterizzato da una marcata stagionalità. In questo ambiente, mai toccato dall'uomo prima del XIX secolo, emblematico è il caso dei pinguini, che sono certamente da considerare i veri "nativi" di quel remoto continente. A parte i pinguini imperatore (*Aptenodytes forsteri*) e poche altre eccezioni, i pinguini si riproducono nelle aree deglaciate delle coste antartiche. Tra questi, il pinguino di Adelia (*Pygoscelis adeliae*, Hombron & Jacquinot 1841) è la specie dominante (Ainley, 2002; Santora *et al.*, 2020).

Grazie alla loro ben nota elevata sensibilità ambientale, i pinguini di Adelia rappresentano attualmente una delle specie più studiate, soprattutto per quanto riguarda la comprensione dei cambiamenti climatico-ambientali che si registrano nell'Oceano Meridionale (Taylor *et al.*, 1990; Woeheler, 1993; Croxall *et al.*, 2002; Ainley, 2002; Ainley *et al.*, 2005; Forcada and Trathan, 2009; Shepherd *et al.*, 2005; Lynch e LaRue, 2014; Borowicz *et al.*, 2018; Santora *et al.*, 2020).

Se dovessero realizzarsi gli scenari più drammatici previsti da un'accelerazione del riscaldamento globale, i pinguini di Adelia non avrebbero alcun luogo abbastanza freddo in cui rifugiarsi per sfuggire alle mutate condizioni ambientali e dovrebbero necessariamente adattarsi in tempi rapidi oppure sarebbero destinati all'estinzione (Lambert, 2010). Questo li rende formidabili indicatori del clima glaciale dell'Antartide, veri e propri campanelli d'allarme dei cambiamenti climatici (Ainley, 2002).

3. I pinguini di Adelia

I pinguini di Adelia (*Pygoscelis adeliae*) sono considerati la specie più resistente alle crude ed estreme condizioni antartiche. Durante

l'estate australe, tra fine ottobre e febbraio, i pinguini nidificano per riprodursi sulle aree deglaciate delle coste antartiche, aggregandosi in colonie di numerosi individui. Letteralmente, milioni di individui nidificano in zone libere dai ghiacci lungo la costa del continente e delle isole che la costellano. Durante l'inverno, quando l'oscurità persiste 24 ore al giorno e la temperatura scende sotto i -40 °C, i pinguini vivono al margine esterno del ghiaccio marino, ben lontani dalle aree costiere dell'Antartide. Quando la luce del sole ritorna, le temperature possono essere ancora al di sotto di -20 °C ma, tra settembre e ottobre, i pinguini viaggiano per centinaia e centinaia di chilometri, camminando o scivolando sul ghiaccio marino per raggiungere le colonie sulla terraferma. Per primi arrivano i maschi che occupano e riassettano i nidi, costituiti da concentrazioni di ciottoletti centimetrici dove la femmina depone due uova (*Fig. 3*). Le colonie sono costituite concentrazioni di questi nidi, da alcune migliaia a diverse centinaia di migliaia di nidi (*Fig. 4*).



Fig. 3. Pinguino di Adelia in un nido di ciottoli con uovo appena deposto a Cape Bird, Isola di Ross, Antartide. (Foto C. Baroni, dicembre 2011, @ PNRA).

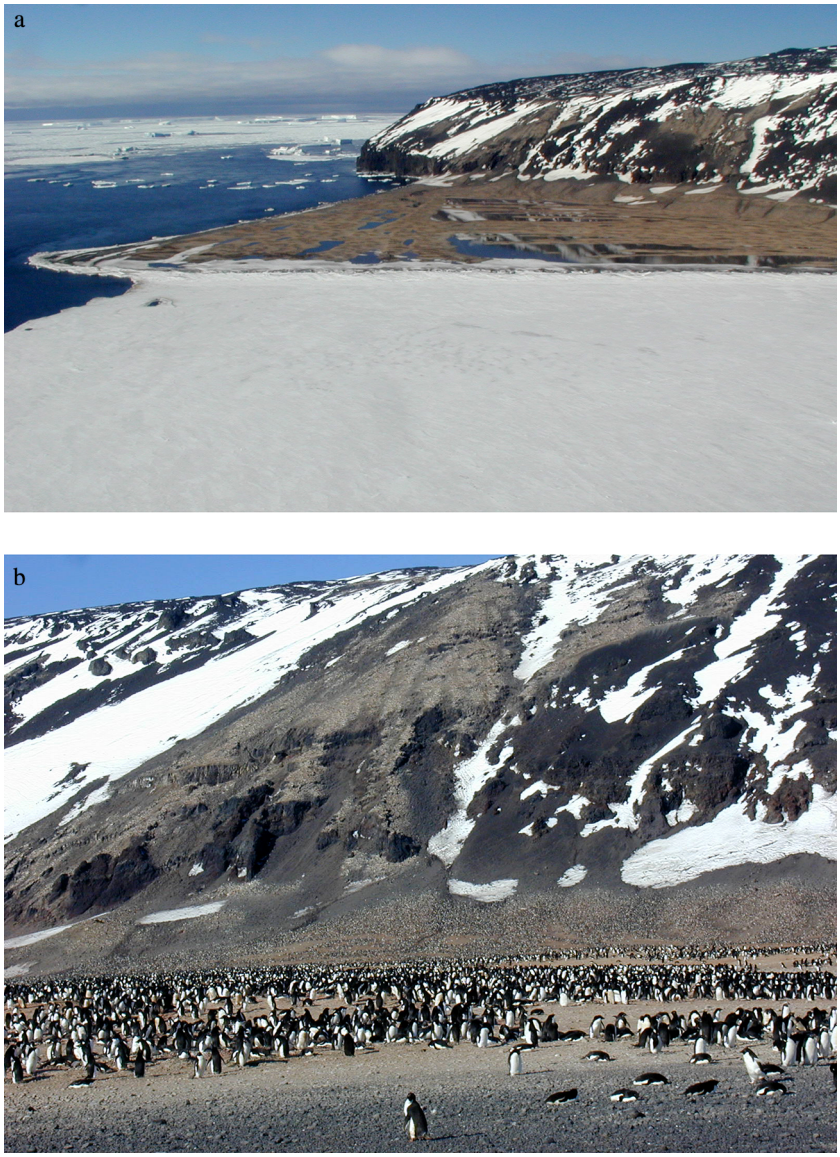


Fig. 4. a) La più estesa colonia al Mondo di pinguini di Adelia a Cape Adare, Terra Vittoria, Antartide. (Foto C. Baroni, dicembre 2002, @ PNRA). *b) Nidi di pinguini di Adelia a perdita d'occhio, colonia di Cape Adare, Terra Vittoria, Antartide.* (Foto C. Baroni, gennaio 2003, @ PNRA).

Il censimento della popolazione riproduttiva degli Adelia ha rilevato 3,79 milioni di coppie (oltre 7,5 milioni di pinguini) distribuite in oltre 250 siti di riproduzione (Lynch e Larue, 2014), che corrisponderebbero a 189 colonie *sensu* Ainley (2002; Santora *et al.*, 2020). Il censimento non considera i pinguini in età non ancora riproduttiva e, quindi, la popolazione totale di pinguini di Adelia potrebbe superare i 9-10 milioni di individui. Un terzo della popolazione adulta si trova nel Mare di Ross (33%), il 21% occupa la Penisola Antartica, il 30% l'Antartide Orientale e il resto le rimanenti aree deglaciate dell'Antartico (Lynch e Larue, 2014). Nel Mare di Ross si trovano sia la più estesa colonia del mondo, a Cape Adare, che ospita 338.000 coppie (Lynch e Larue, 2014), sia la colonia più meridionale del nostro Pianeta, a Capo Royds (77° 30' S), nell'Isola di Ross (*Fig. 5*).

I pinguini di Adelia tornano istintivamente nelle stesse colonie dove sono stati generati, mostrando quindi una elevata filopatria natale; in passato, si pensava che esistessero bassi livelli di migrazione tra le diverse colonie, il che, in realtà, è abbastanza vero in periodi di condizioni ambientali e climatiche relativamente stabili (Ainley, 2002; Shepherd *et al.*, 2005). D'altra parte, una forte filopatria natale dovrebbe promuovere la divergenza genetica tra le popolazioni. Tuttavia, stanno emergendo prove di un elevato livello di flusso genico episodico tra le colonie di pinguini, molto probabilmente causato dalla migrazione forzata da condizioni ambientali che impediscono l'accesso ad alcune colonie (Shepherd *et al.*, 2005). Emblematico è il caso dei mega-iceberg distaccati dalla Piattaforma di Ross, che si sono arenati recentemente impedendo l'accesso ad alcune colonie dell'Isola di Ross. Lo sbarramento delle normali rotte migratorie dei pinguini ha avuto drammatiche conseguenze non solo sull'occupazione dei siti di riproduzione ma anche sul foraggiamento e, in alcuni casi, ha pesantemente influito addirittura sul successo riproduttivo.

I principali fattori ecologici che regolano e limitano la distribuzione dei pinguini di Adelia intorno all'Antartide sono: 1) le aree deglaciate lungo i margini costieri, 2) l'estensione e la persistenza della banchisa, 3) la presenza di acque libere dal ghiaccio marino nei pressi della colonia per consentire l'accesso al mare (foraggiamento), 4) la disponibilità di cibo (krill, piccoli pesci e calamari), 5) la relativa assenza di predatori.

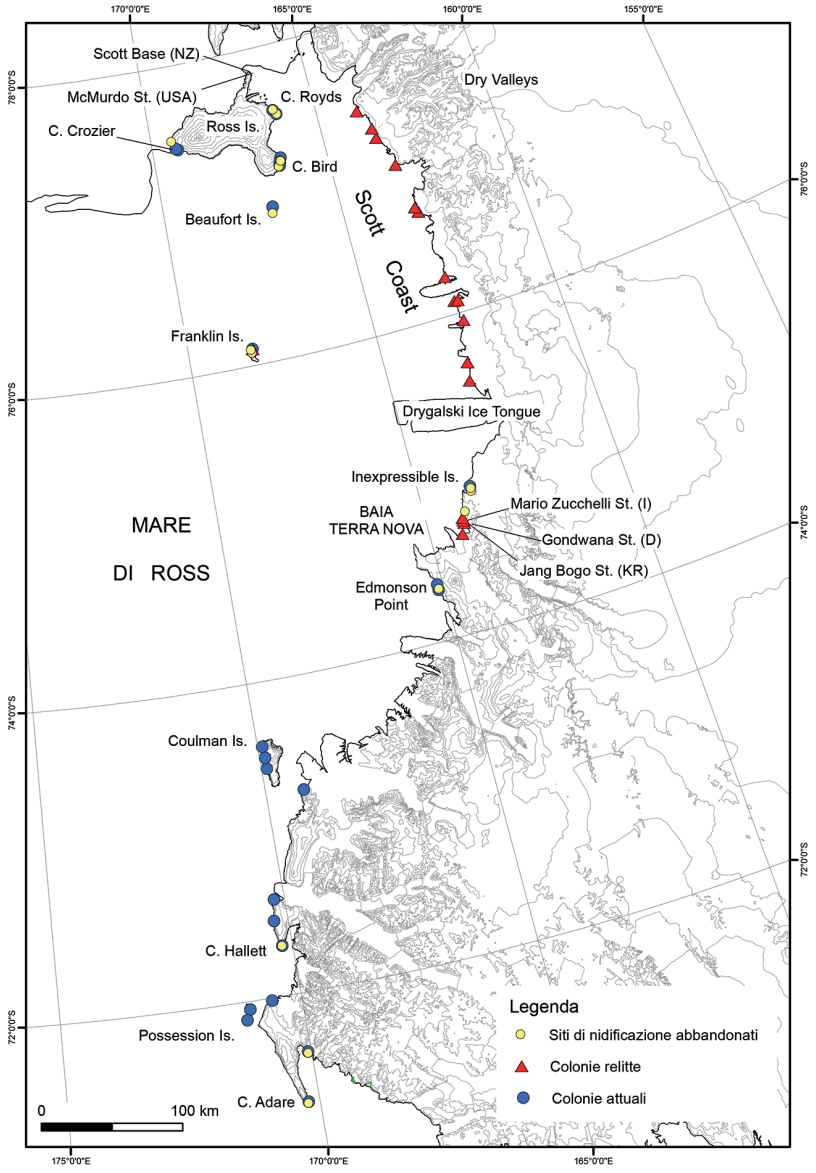


Fig. 5. La Terra Vittoria, Antartide. Distribuzione delle attuali colonie di pinguini di Adelia (pallini blu), dei siti di nidificazione abbandonati ai margini delle colonie attualmente occupate (pallini gialli) e delle colonie relitte (triangoli rossi). [Dati dall'Antarctic Digital Database - ADD (https://www.scar.org/data-products/antarctic-digital-database/)].

I pinguini stanno attualmente reagendo ai cambiamenti climatici spostando le loro colonie riproduttive a latitudini più meridionali. Uno dei più evidenti segnali del riscaldamento in atto viene dalla Penisola Antartica, dove la temperatura media invernale dell'aria è aumentata più rapidamente (6 °C dal 1950) che in qualsiasi altra parte del mondo (Stokstad, 2007). Nuovi dati sul censimento dei pinguini di Adelia mostrano che la popolazione della Penisola si sarebbe notevolmente ridotta rispetto al 1993 e almeno otto colonie sarebbero state abbandonate (Lynch e LaRue, 2014). Contestualmente, la popolazione del Mare di Ross e dell'intera Antartide Orientale è aumentata sia in termini di numero di individui (53%) sia in termini di nuove colonie (almeno 11 nuove colonie). Allo stesso tempo, altre specie di pinguini hanno occupato le nicchie ecologiche abbandonate dagli Adelia nella Penisola Antartica (Forcada e Trathan, 2009).

4. Il patrimonio di informazioni custodito nelle colonie dei pinguini di Adelia

4.1 Colonie di pinguini e suoli ornitogenici

Le attuali colonie di pinguini di Adelia si sviluppano sulle spiagge emerse oloceniche e sui terrazzi marini, ma si espandono anche sui coni e sulle falde detritiche, sui rock glaciers pedemontani, sulle morene deposte e/o a nucleo di ghiaccio delle aree costiere. In alcune colonie i pinguini possono arrampicarsi per nidificare sui versanti in roccia spingendosi fino a qualche centinaio di metri di quota (*Fig. 4*). L'estensione delle colonie varia in funzione della popolazione di pinguini e può estendersi nell'entroterra costiero da diverse centinaia di metri fino ad alcuni chilometri. Ogni anno, i genitori dei pinguini sistemano i sassolini selezionati con il becco per costruire il nido (*Fig. 6*). In una colonia si osservano aggregazioni di nidi che possono variare da alcune migliaia a diverse centinaia di migliaia di unità (*Fig. 4*). Durante la stagione riproduttiva, la predazione di uova e di pulcini da parte degli skua antartici (*Catharacta maccormicki; Fig. 7*), lascia al suolo una gran quantità di resti di pinguino (gusci d'uovo, carcasse di pulcini, ossa, piume, frammenti di pelle e altri resti organici), che si mescolano alle deiezioni (guano) e ai ciottoletti dei nidi. Con il trascorrere del tempo, il guano, i nidi e i resti della predazione da parte degli skua si sovrappongono e si saldano tra loro accumulandosi a

formare veri e propri suoli organici, riconosciuti fin dalle spedizioni antartiche del XIX secolo (*Fig. 8*). Henryk Arctowski, geologo della Spedizione Antartica Belga al comando di de Gerlache (1898-99), raccolse in depositi di guano i primi insetti mai trovati in Antartide.



Fig. 6. Pinguino di Adelia che trasporta col becco un ciottolo per assemblare il nido, colonia di Cape Bird, Isola di Ross, Antartide. (Foto C. Baroni, gennaio 2007, @ PNRA).



Fig. 7. Skua antartico, il principale predatore delle colonie di pinguini di Adelia, Cape Bird, Isola di Ross, Antartide. (Foto C. Baroni, gennaio 2007, @ PNRA).



Fig. 8. Nidi di pinguini di Adelia, Edmonson Point, Terra Vittoria, Antartide. Si noti il suolo ornitogenico visibile al centro della foto, creato dall'accumulo di deiezioni (guano) resti organici e nidi sovrapposti. (Foto C. Baroni, dicembre 2006, @ PNRA).

I terreni organici delle colonie di pinguini furono definiti suoli ornitogenici da Syroechkovsky (1959) e sono stati descritti per la prima volta nella Terra Vittoria (Harrington, 1960; Ugolini, 1972; Campbell e Claridge, 1987). Questi suoli costituiscono un eccellente archivio di informazioni sui processi pedogenetici e sulla chimica dei suoli organici che si sviluppano in condizioni ambientali veramente estreme (Ugolini, 1972; Speir e Cowling, 1984; Campbell e Claridge, 1987; Heine e Speir, 1989; Tatur e Myrcha, 1989; Zhu et al, 2005; Michel *et al.*, 2006; Simas *et al.*, 2007, 2008).

In una colonia occupata da lungo tempo il guano si accumula creando orizzonti di diversi decimetri di spessore: tanto maggiore è la durata dell'occupazione quanto maggiore sarà lo spessore dei suoli (Ugolini, 1972). Una duratura fase di occupazione delle colonie si riflette in un progessivo accrescimento dei depositi organici; inoltre, i nidi, le carcasse e gli altri resti organici possono eventualmente essere coperti da depositi di versante o di altra natura. Tipici sono i depositi colluviali trasportati dalle acque di fusione nivale che localmente possono intercalarsi ai suoli organici.

Alla periferia o nelle vicinanze delle colonie attualmente occupate si trovano siti di nidificazione abbandonati, che testimoniano sia la persistenza nel tempo delle condizioni favorevoli alla colonizzazione, sia la variabilità del numero di pinguini che costituiscolo le popolazioni che occupano quelle colonie (*Fig. 9*). Questi siti sono stati riconosciuti anche a chilometri di distanza dal centro delle colonie attuali (Baroni e Salvatore, dati inediti). Inoltre, siti di nidificazione abbandonati sono stati individuati anche in aree oggi non più occupate da colonie di pinguini (Baroni e Orombelli, 1994; *Fig 5*). Si tratta di vere e proprie colonie relitte che custodiscono preziose informazioni sulla storia dei pinguini e dell'ambiente in cui hanno vissuto.



Fig. 9. Scavo di una colonia relitta di pinguini di Adelia, Scott Coast, Terra Vittoria, Antartide. (Foto C. Baroni, gennaio 2013 @ PNRA).

Dopo l'abbandono dei siti di nidificazione, i forti venti antartici hanno spazzato la superficie di occupazione asportando le componenti più superficiali e più fini, generando veri e propri pavimenti eolici di deflazione costituiti da plaghe di ciottoletti concentrate in corrispondenza dei vecchi nidi, che raggiungono anche estensioni di decine/centinaia di metri quadrati. Una diffusa copertura di licheni ammanta i ciottoli delle più antiche colonie relitte.

Sia le colonie relitte sia i suoli ornitogenici ci forniscono un importante archivio di informazioni sulle popolazioni di pinguini di Adelia che le hanno occupate nel tempo. Per prima cosa il guano e gli altri resti di pinguino ci permettono di ottenere dati cronologici (^{14}C; *Fig. 10*) che scandiscono la storia della colonizzazione delle aree costiere, le fasi di ritiro delle calotte antartiche dopo l'Ultimo Massimo Glaciale e contribuiscono alla ricostruzione delle curve di sollevamento relativo delle aree costiere (Baroni e Orombelli, 1991; Baroni, 1984, 2013; Emslie, 1995; Emslie *et al.*, 1998, 2007; Baroni e Hall, 2004; Hall *et al.*, 2004). Se consideriamo le colonie di pinguini "sentinelle" dei recenti cambiamenti climatici (Ainley, 2002), i dati ottenuti dalla distribuzione delle colonie relitte e dai siti di nidificazione abbandonati in colonie ancora occupate ci forniscono informazioni rilevanti sulle risposte ecologiche dei pinguini agli episodi dei passati cambiamenti ambientali e climatici.

Inoltre, l'ambiente antartico, freddo e secco, conserva nei suoli ornitogenici carcasse di pinguini e altri resti (ossa, pelle, piume, guano e gusci d'uovo) che hanno fornito il DNA antico (sia nucleare che mitocondriale) meglio conservato e mai scoperto (Lambert *et al.*, 2002; Ritchie *et al.*, 2004, Millar *et al.*, 2008; Subramanian *et al.*, 2009; Parks *et al.* 2015).

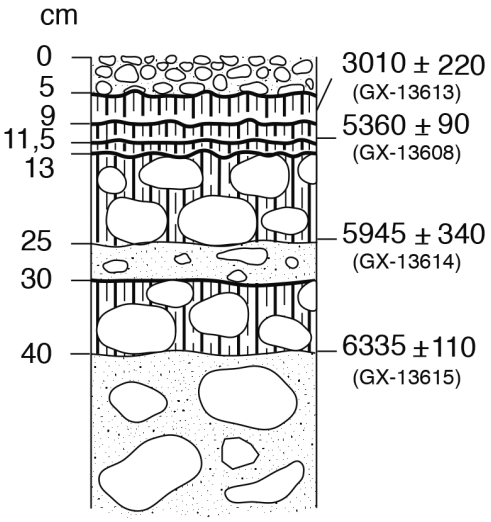


Fig. 10. Inexpressible Island (Fig. 5), sovrapposizione di diversi livelli di guano (linee verticali) generati da fasi di occupazione sovrapposte che hanno restituito diverse date ^{14}C (numeri sulla destra; tra parentesi le sigle dei campioni datati. (Ridisegnato da Baroni e Orombelli, 1991).

Le sequenze di DNA estratte dalle ossa di pinguino di Adelia hanno permesso di definire variazioni genetiche e trend evolutivi indotti dai cambiamenti climatico-ambientali che hanno accompagnato le fasi glaciali e interglaciali (tipicamente le fasi di raffreddamento climatico che hanno accompagnato le glaciazioni e le fasi di riscaldamento registrate negli interglaciali e nell'Olocene).

4.2 La deglaciazione e la ri-colonizzazione delle aree costiere

Durante il Pleistocene, le ripetute fasi di espansione e di collasso delle calotte antartiche hanno impattato sull'*habitat* dei pinguini a scala continentale, allontanandoli dalle numerose ma discontinue aree deglaciate distribuite lungo le coste (*Fig. 1*). Durante l'ultimo periodo di massima espansione glaciale (UMG, tra 25.000 e 18.000 anni fa), le aree costiere antartiche erano in gran parte coperte da ghiaccio, in seguito all'ispessimento dei margini delle calotte e alla loro avanzata su ampi settori della piattaforma continentale. In conseguenza dell'abbassamento del livello del mare a scala globale (circa -120/130 m rispetto all'attuale livello) la linea di ancoraggio delle piattaforme è avanzata di centinaia di chilometri spingendo molto più a Nord, rispetto alla posizione attuale, sia il margine delle piattaforme galleggiani, sia il limite del ghiaccio marino (Dawson, 1992; Bentley *et al.*, 2016). Lo spessore dei ghiacciai raggiungeva, sul margine costiero, diverse centinaia metri, innalzandosi localmente fino a 400-500 m al di sopra del livello del mare attuale (come documentato dai depositi glaciali e dalla distribuzione dei massi erratici; Denton e Hughes, 1981, 2002; Orombelli *et al.*, 1990; Denton *et al.*, 1991).

L'ultima transizione dal periodo glaciale all'attuale interglaciale (Olocene) ha indotto un innalzamento della temperatura sull'Antartide di circa 12 °C che si è verificato molto repentinamente. In pochi millenni le calotte antartiche si sono ritirate nella loro posizione attuale, insieme alle piattaforme galleggianti e alla loro linea di ancoraggio. Di conseguenza, la deglaciazione è stata accompagnata dall'ingressione marina e dal sollevamento glacio-isostatico delle zone costiere. Questi eventi hanno portato alla formazione di spiagge emerse oloceniche.

Molto significativo è il caso di studio del Mare di Ross e, in particolare, della colonizzazione della Terra Vittoria, dove i pinguini di Adelia occupano un ecosistema marino molto particolare, tra i più meridionali del Pianeta. Attualmente, circa 1/3 della popolazione totale

di pinguini di Adelia occcupa il margine occidentale del Mare di Ross, distribuita in 20 colonie lungo il margine costiero della Terra Vittoria (*Fig. 5*).

Tamite l'analisi di fotografie aeree e di immagini da satellite, rilevamenti geomorfologici regionali e studi geomorfologici di dettaglio condotti in siti chiave sono state studiate le aree costiere dell'intera Terra Vittoria. L'accurato rilevamento geomorfologico regionale per l'analisi del paesaggio delle aree deglaciate (*Landscape analysis*), affiancato da un approccio tipicamente "archeologico" per il ritrovamento dei siti di colonizzazione ("*survey*" areale) ci ha permesso di scoprire numerose colonie di pinguini relitte tra Cape Adare e l'Isola di Ross (Baroni e Orombelli, 1991, 1994; Baroni, 1994; Lambert *et al.*, 2002; Baroni e Hall, 2004; Hall *et al.*, 2004; Ritchie *et al.*, 2004; Baroni e Salvatore, dati inediti; *Fig. 5*).

Applicando le tecniche proprie dell'archeologia, attraverso scavi stratigrafici dei siti di nidificazione abbandonati, è stato possibile identificare le fasi occupazionali delle colonie (eventualmente separate da strati minerali come sabbia e ghiaia di origine colluviale o periglaciale, depositi eolici, ecc.) e ricostruire la storia della colonizzazione costiera. Sono stati eseguiti saggi di scavo (generalmente 1-2 m^2) o veri e propri scavi estesi alcuni m^2, pulendo la superficie dall'alto e scavando livello dopo livello i suoli organici, fino ai sedimenti sterili sottostanti oppure fino al substrato roccioso. Inoltre, con un accurato approccio stratigrafico, è stato possibile discriminare, già durante il lavoro di scavo, tra i resti appartenenti a singoli individui e quelli attribuibili a organismi distinti, consentendo anche di riconoscere e caratterizzare i livelli organici eventualmente rimaneggiati da processi periglaciali o di altra natura. La elevata concentrazione di resti organici ha fornito numerose date ^{14}C (circa 200) che ci consentono di scandire con buona approssimazione le fasi della deglaciazione, l'arretramento della linea di ancoraggio delle piattaforme galleggianti e i tempi dell'ingressione marina nell'Olocene.

Nel Mare di Ross, la regressione glaciale si è completata intorno a circa 8000 anni fa, che all'incirca corrisponde all'età della più antica spiaggia emersa olocenica che si trova circa 30 m al di sopra del livello del mare attuale e all'età delle più antiche colonie di pinguini occupate dopo la deglaciazione (Baroni e Orombelli, 1991, 1994; Baroni e Hall, 2004). In realtà, il ritiro della linea di ancoraggio delle piattaforme galleggianti è stata un poco più articolata ed è avvenuta in fasi distinte tra circa 8000 e circa 3000 anni fa (Convay *et al.*, 1999; Hall *et al.*, 2004).

Il sollevamento delle aree costiere è ben descritto da curve di innalzamento relativo del livello del mare basate su numerose date ^{14}C ottenute da conchiglie raccolte nei sedimenti marini (che hanno fornito età contemporanee e/o età massime per il livello del mare nel corso dell'Olocene) e da resti organici provenienti da suoli ornitogenici di colonie relitte e siti abbandonati di nidificazione individuati sulle spiagge oloceniche (che hanno fornito età minime; Baroni e Orombelli, 1991; Baroni e Hall, 2004; Hall *et al.*, 2004). Nella zona di Baia Terra Nova, il tasso di sollevamento iniziale nei primi millenni dopo la deglaciazione è stato di 8-10 mm/anno, ridotto a meno di 2 mm/a negli ultimi tre millenni (<0.5 mm/a nell'ultimo millennio).

4.3 Le paleodiete dei pinguini

L'eccellente conservazione della sostanza organica nell'ambiente secco e freddo dell'Antartide permette anche la conservazione dei resti delle specie predate dai pinguini, concentrate nelle deiezioni trattenute nei suoli ornitogenici. Ciò consente, pertanto, anche di studiare i resti delle paleodiete dei pinguini e di ricostruire le strategie di foraggiamento. Ossa, otoliti e altri resti di pesci, oltre a becchi di calamaro (chitinosi), sono perfettamente conservati nei suoli e hanno fornito rilevanti informazioni sulle paleodiete dei pinguini di Adelia. L'identificazione tassonomica del "*silverfish*" antartico (*Pleuragramma antarcticum*) e dei calamari glaciali (*Psychroteuthis glacialis*) ha rivelato queste due specie come importanti componenti nella dieta del pinguino di Adelia fin dall'inizio dell'Olocene (Emslie *et al.*, 1998; Emslie, 2001; McDaniel and Emslie, 2002; Emslie and McDaniel, 2002; Polito *et al.*, 2002; Emslie and Woehler, 2005; Lorenzini *et al.*, 2009).

La proporzione di questi resti nella paleodieta dei pinguini è variata in accordo con le modificazioni climatico ambientali che hanno interessato le coste antartiche e, in particolare, il settore del Mare di Ross. E' noto, infatti, che le strategie di foraggiamento degli Adelia sono legate alle condizioni ambientali locali e regionali e i dati recentemente acquisiti nella Terra Vittoria sembrano confermare che questo sia avvenuto anche nel passato. Pertanto, la variabilità delle specie predate dai pinguini rappresenta un valido strumento per studiare, da un altro punto di vista, gli effetti indotti dai cambiamenti globali sugli *habitat* dei pinguini (Ainley *et al.*, 1998; Ainley, 2002).

I campioni di suolo raccolti nelle colonie della Terra Vittoria

hanno fornito migliaia di resti di pesce (vertebre, ossa, otoliti, …) e di altri organismi predati dai pinguini. Gli otoliti raccolti sono stati identificati sulla base di tavole anatomiche comparative dei diversi taxa antartici (Williams e McEldowney, 1990). L'identificazione tassonomica degli otoliti ha permesso di dimostrare che la specie principalmente predata dai pinguini è il *Pleuragramma antarcticum* Boulenger 1902, noto come "*silverfish*" (88% del numero totale di otoliti) e, in ordine decrescente d'importanza, il *Trematomus bernacchii* Boulenger 1902 (1,1%) e il *Trematomus scottii* Boulenger 1907 (0,13%); un certo numero di otoliti è stato identificato anche come *Trematomus* sp. (0.3%) e *Pagothenia* sp. (1.4%).

Inoltre, analisi quantitative condotte su circa 400 otoliti hanno consentito di definire anche le dimensioni medie della taglia dei pesci preferibilmente predati, compresa tra 40 e 80 mm nell'83% dei casi (67±23 mm; Lorenzini *et al.*, 2009).

Numerosi campioni di guano non contengono resti di specie predate, riflettendo presumibilmente diete a base di krill, i gamberetti antartici che non restano conservati nei suoli, ma che sono eventualmente individuabili attraverso l'analisi degli isotopi stabili estratti dal guano di pinguino e dai gusci d'uovo. Queste analisi hanno infatti consentito di individuare anche il contributo dietetico di altre prede, tra le quali principalmente il krill (*Euphausia superba*), non osservato direttamente nei suoli nella forma di resti macroscopici (Lorenzini *et al.*, 2010, 2012).

5. Gli elefanti marini e la colonizzazione delle coste antartiche

Avvistamenti di elefanti marini meridionali (*Mirounga leonina*) sulle coste antartiche sono stati segnalati sin dall'epoca eroica dell'esplorazione antartica e avvengono tuttora anche alle latitudini più meridionali, lungo la costa della Terra Vittoria, ma questi mammiferi marini non sono annoverati tra gli elementi più comuni dell'ecosistema antartico (Le Boef e Laws, 1994; Biew *et al.*, 2007). Gli elefanti marini meridionali sono, infatti, mammiferi molto mobili, ma principalmente subantartici, che prediligono acque stagionalmente libere dai ghiacci marini. Nel 1990 la popolazione totale era di 664.000 individui; oggi si stima che siano diminuiti ma non ci sono dati su censi-

menti recenti dell'intera popolazione. Colonie riproduttive con un consistente numero di individui si trovano nella Penisola di Valdes, in Sud America, nella Georgia del Sud (circa il 60% della popolazione totale), sulle isole subantartiche (Kerguelen, Heard e altre minori, circa il 28% della popolazione totale) sulle isole Macquarie, Campbell, Auckland e sulle isole Antipodi (circa il 12%). La maggior parte delle colonie si trovano a cavallo (oppure all'esterno) della convergenza Antartica e solo poche piccole colonie sono segnalate entro questo limite (De Bruyn et al,, 2009). Tra queste, le uniche colonie presenti sulle coste deglaciate dell'Antartide si trovano alla Statione Palmer sulla Penisola Antartica e sulle Isole Windmill nella Terra di Wilkes (Antartide orientale).

Nella famiglia dei focidi, gli elefanti marini, pur essendo caratterizzati da un forte dimorfismo sessuale, sono la specie che raggiunge le più grandi dimensioni. I maschi adulti possono raggiungere 4.5 m di lunghezza e fino a 3.7 t di peso, mentre le femmine restano al di sotto di 2.8 m di lunghezza e 800 kg di peso (Le Boeuf and Laws, 1994).

Gli elefanti marini si alimentano in acque aperte, in assenza di banchisa, dove si muovono ad ampio raggio (migliaia di km) coprendo aree di foraggiamento molto estese.

Sono grandi predatori di calamari e pesci, in grado di raggiungere la sorprendente profondità di 1430 m. La loro impressionante capacità di immersione può durare per un tempo incredibilmente lungo, fino a 120 minuti (Le Boeuf and Laws, 1994).

Gli elefanti marini si riposano sulla terraferma, rotolandosi sulle spiagge sabbiose o ghiaiose, libere dai ghiacci, per la muta e per riprodursi. Il goffo e pesante movimento degli elefanti marini a terra ha l'effetto di bulldozer che arano le spiagge, travolgono e schiacciano qualsiasi cosa si trovi sul loro cammino. Colonie di pinguini eventualmente occupate da questi enormi mammiferi subiscono gravi danni e raramente si contano superstiti.

Durante i secoli XIX e XX questi mammiferi sono stati cacciati per la produzione di olio e la popolazione era stata pesantemente predata, giungendo quasi sull'orlo dell'estinzione (Le Boeuf eLaw, 1994). Attualmente, la popolazione sembra sottoposta a una forte riduzione, con estinzione delle colonie più settentrionali e conseguente migrazione degli elefanti marini a latitudini più meridionali, sempre più in direzione dei mari antartici (Hindel *et al.*, 2017; Jones *et al.*, 2020).

Hall *et al.* (2006) hanno scoperto che gli elefanti marini hanno in passato occupato la Terra Vittoria, colonizzandola sin da 7000 anni fa (*Fig. 11*). Sulle spiagge sabbioso-ghiaiose e ciottolose presenti lungo la costa deglaciata della Terra Vittoria sono stati raccolti numerosi campioni di pelle, ossa e resti di scheletri mummificati, inclusi neonati, a dimostrazione che molti dei siti occupati lungo la Terra Vittoria non erano semplici siti occupati per la muta o come semplici aree di riposo, ma erano vere e proprie colonie riproduttive (Hall *et al.*, 2006: Koch *et al.*, 2019). Attualmente, a parte la polynia di Inexpressible Island nella zona di Baia Terra Nova, la Scott Coast è sbarrata da una banchisa costiera, localmente caratterizzata da ghiaccio pluriennale, e non è pertanto raggiungibile dagli elefanti marini. La colonia più vicina attualmente si trova sull'Isola Macquarie (54° 30' S), 2500 km a Nord, circa a mezza strata tra l'Antartide e l'Australia.



Fig. 11. Un elefante marino mummificato sulle spiagge di Inespressible Island, Baia Terra Nova, Antartide. (Foto C. Baroni, gennaio 2002, @ PNRA).

Oltre 200 campioni di resti di elefanti marini sono stati datati con il ^{14}C e utilizzati per estrarre il DNA antico, applicando stretti protocolli per sequenziare e amplificare il DNA mitocondriale (mtDNA), al fine di ricostruire la storia evolutiva della popolazione ancestrale di elefanti

marini nel Mare di Ross (De Bruyn *et al.*, 2009, 2014). Questa popolazione aveva stretti legami genetici con la popolazione che attualmente occupa la colonia dell'Isola Macquarie e si è insediata a partire da circa 7500 anni fa, subito dopo la deglaciazione che ha seguito l'UMG, più o meno contemporaneamente (o poco dopo) l'arrivo dei primi pinguini di Adelia. Il collasso di queste colonie è iniziato intorno a 1000 anni fa e si è completato circa 400 anni fa, in un periodo che, nell'emisfero settentrionale, corrisponde alla "piccola età glaciale".

6. La risposta dei pinguini di Adelia e degli elefanti marini ai cambiamenti climatici

Le condizioni ambientali dell'UMG hanno obbligato i pinguini e tutti gli animali che nidificano nelle aree costiere a migrare verso Nord, nel settore circum-antartico, in cerca di luoghi adatti per la nidificazione e la riproduzione. A riprova di questo comportamento dei pinguini sono emerse recentemente molte evidenze di occupazioni di colonie precedenti all'UMG, che vanno dalla scoperta di suoli ornitogenici (sepolti da depositi glaciali dell'UMG), a resti di scheletri subfossili di pinguini, ad altri resti organici associati alle colonie (Ritchie et al 2004; Gardner *et al.*, 2006; Subramanian *et al.*, 2009; Emslie *et al.*, 2007; Parks *et al.*, 2015). Dopo la deglaciazione, le zone costiere libere dal ghiaccio erano perfettamente disponibili per essere ri-colonizzate. I pinguini di Adelia hanno riconquistato i siti ancestrali colonizzati dai loro progenitori del Pleistocene.

Il tasso evolutivo calcolato dal DNA antico prelevato dai resti fossili di pinguini, rispetto agli eventi di divergenza osservati nel Pleistocene, ha permesso a Ritchie *et al.* (2004) di riconoscere due popolazioni con un distinto patrimonio genetico (*lineage*) nel DNA mitocondriale. Un primo "*lineage*" è stato osservato principalmente nelle colonie del Mare di Ross, mentre un secondo è stato evidenziato nella popolazione circumantartica. La diversa distribuzione dei due patrimoni genetici è coerente con l'esistenza di due rifugi ancestrali (pleistocenici) che erano soggetti a strozzature nelle dimensioni della popolazione (*bottleneck*). Il tasso di evoluzione molecolare della regione HVRl, determinato estraendo il DNA mitocondriale da 162 ossa subfossili che abbracciano un arco temporale di circa 37.000 anni, era di 0,96 sostituzioni per sito ogni miIione di anni (Lambert *et al.*, 2002).

Applicando questo tasso evolutivo, il più recente progenitore di questi pinguini avrebbe un'età compresa tra 75.000 e 122.000 anni, certamente più antico dell'UMG, possibilmente più antico dell'ultimo ciclo glaciale (> MIS 4) e, molto probabilmente attribuibile all'ultimo interglaciale (MIS 5e).

La limitata distribuzione del patrimonio genetico del "*lineage*" del Mare di Ross suggerisce che le aree rifugio di questa popolazione nel Pleistocene superiore dovevano essere molto vicine a questo settore antartico. Inoltre, la relativamente alta frequenza del patrimonio genetico di questa popolazione al margine settentrionale della Terra Vittoria e la sua progressiva diminuzione man mano ci si sposti verso la piattaforma di Ross, al margine meridionale della colonizzazione attuale, a vantaggio della popolazione circumantartica, indica che entrambe le popolazioni avrebbero progressivamente colonizzato le aree costiere del Mare di Ross ma con un comportamento relativamente diversificato, relativamente meno dinamico nel caso della popolazione ancestralmente legata alle aree rifugio al margine esterno del Mare di Ross.

La colonizzazione della Terra Vittoria da parte dei pinguini di Adelia è intimamente legata alla presenza di un'altra importante specie di predatori, gli elefanti marini meridionali, che si sono recentemente rivelati importanti attori nella storia della colonizzazione delle coste antartiche durante l'Olocene (Hall *et al.*, 2006; De Bruyn *et al.*, 2009, 2014).

Una popolazione originaria della lontanissima Isola Macquarie ha colonizzato le coste della Terra Vittoria non appena questo *habitat* si è reso disponibile, dopo il ritiro dei ghiacciai avvenuto dopo l'UMG (De Bruyn *et al.*, 2009, 2014). Nell'arco di poche generazioni si è sviluppata una popolazione geneticamente molto diversa da quella originale. I fattori che hanno guidato questo evento sono rappresentati dall'assenza stagionale della banchisa, dalla disponibilità di spiagge sabbioso-ghiaiose e ciottolose facilmente raggiungibili e dalla disponibilità di cibo. La variabilità delle condizioni ambientali che rispondono a questi fattori ha peraltro guidato la fluttuazione del numero di siti occupati nel corso dell'Olocene, la variazione delle dimensioni della popolazione e l'abbandono di questo settore antartico nell'ultimo millennio.

Sulla base delle numerose date ^{14}C associate all'occupazione delle colonie di pinguini e di elefanti marini, è evidente che sono esistite nell'Olocene condizioni più o meno favorevoli per i fattori che limitano o favoriscono la presenza di queste due specie animali (Baroni e

Orombelli, 1994, Hall *et al.*, 2006, Mezgec *et al.*, 2017; *Fig. 12*). In particolare, la presenza o l'assenza di elefanti marini è condizionata dalla estensione e dalla persistenza del ghiaccio marino e, soprattutto, della banchisa costiera costituita da ghiaccio pluriennale, come evidenziato dall'associazione di diatomee e altri microfossili che si trovano nei sedimenti marini (Mesgec *et al.*, 2017).

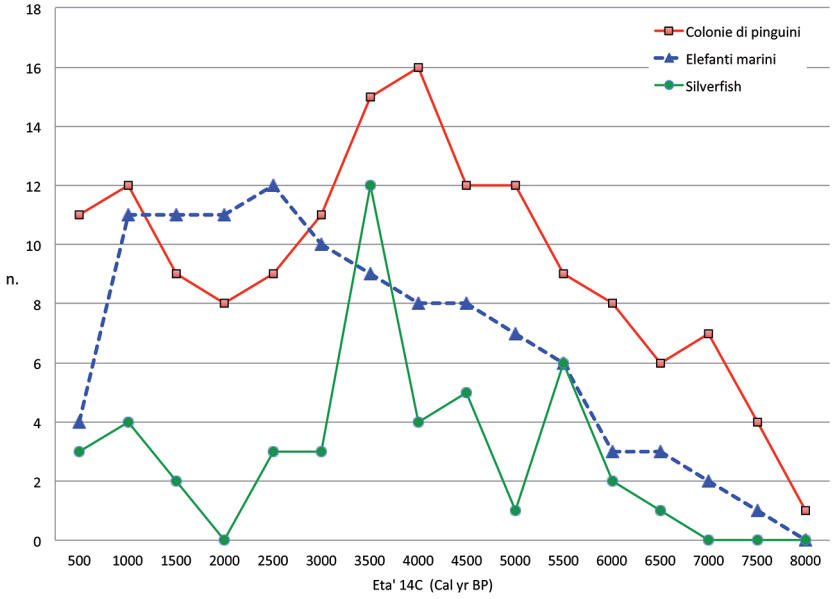


Fig. 12. Distribuzione e frequenza delle colonie di pinguini di Adelia (attuali e relitte), delle colonie di elefanti marini e di resti di Pleurogramma antarcticum (silverfish) da livelli di guano campionati in siti abbandonati di pinguini nella Terra Vittoria (i dati sono raggruppati per intervalli di 500 anni. (Modificato da Mezgec *et al.*, 2017, con dati inediti di Baroni e Salvatore).

A seguito della deglaciazione completata intorno a 8000 anni fa, una polynia si stabilì al largo di Inexpressible Island nella zona di baia Terra Nova. Tra 8000 e circa 5000 anni elefanti marini e pinguini di Adelia occuparano estesamente le aree deglaciate della Terra Vittoria. Tra 5000 e circa 3000 anni fa si è verificato un periodo particolarmente favorevole allo sviluppo di colonie di pinguini di Adelia, che hanno colonizzato estesi settori costieri, oggi non più occupati ("penguin optimum", Baroni e Orombelli, 1994; Hall *et al.*, 2004).

In questo periodo la popolazione di elefanti marini ha subito una forte riduzione, persistendo principalmente nei siti colonizzati più vicini alla polynia di Inexpressible Island.

Tra 2500 e 1100 anni fa, condizioni subantartiche hanno caratterizzato le coste della Terra Vittoria, favorendo una significativa espansione della presenza di elefanti marini, che in quel periodo hanno occupato il maggior numero di colonie. In questo periodo l'estensione della banchisa era ai minimi termini, almeno stagionalmente. Le colonie di pinguini, in questa fase, avrebbero subito una drastica riduzione, sia in termini di colonie occupate sia in termini di dimensioni della popolazione.

La scomparsa delle colonie è avvenuta attraverso un collo di bottiglia che si osserva, su base genetica, intorno a 1000 anni fa (De Bruyn *et al.*, 2014) e che vede l'espulsione degli elefanti marini dal Mare di Ross intorno a 500 anni fa.

In passato, la variabilità delle condizioni ambientali era direttamente associata a condizioni climatiche semplicemente intese in relazione a conditioni di temperatura più o meno elevata. Mezgec *et al.* (2017) hanno recentemente dimostrato che le condizioni che favoriscono o limitano la presenza di pinguini, elefanti marini e pesci preferenzialmente predati da queste specie (*Pleurogramma antarticum*) sono in realtà controllati da un più complesso insieme di condizioni ambientali, tra le quali emerge il ruolo delle polynia nel condizionare la produzione di ghiaccio marino e la persistenza di ghiaccio stagionale e pluriennale nelle aree costiere. Complessivamente, i nuovi dati indicano che i venti catabatici sono i principali attori che regolano la formazione del ghiaccio marino (controllando l'efficienza della polynia) e la persistenza della banchisa costiera, mentre i venti geostrofici guidano l'allontanamento e l'estensione del ghiaccio marino nel Mare di Ross, chiarendo che le variazioni ambientali registrate dall'ecosistema marino in Antartide non sono meramente riconducibili a semplici variazioni di temperatura ma vanno considerate in un più ampio contesto.

BIBLIOGRAFIA

Ainley D.G., 2002. The Adélie penguin: bellwether of climate change, Columbia University Press, New York, 310 pp.

Ainley D.G., Clarke E.D., Arrigo K., Fraser W.R., Kato A., Barton K.J., Wilson P.R., 2005. Decadal-scale changes in the climate and biota of the Pacific sector of the Southern Ocean, 1950's to the 1990s. Antarctic Science, 17, 171-182.

Ainley D.G., Wilson P.R., Barton K.J., Ballard G., Nur N., Karl B., 1998. Diet and foraging effort of Adélie penguins in relation to pack-ice conditions in the southern Ross Sea. Polar Biology, 20, 311-319.

Anderson J.B., 1999. Antarctic Marine Geology. Cambridge University Press, Cambridge, 289 pp.

Ascenzi A., Segre A., 1971. A new Neanderthal child mandible from an Upper Pleistocene site in Southern Italy. Nature, 233, 280–282.

Barker P.F., Cooper A.K. (Eds.), 1997. Geology, Seismic Stratigraphy of the Antarctic Margin. Part 2. Antarctic Research Series. American Geophysical Union 71, Washington, USA, 187 pp.

Baroni C., 1994. Notes on late-glacial retreat of the Antarctic ice sheet and Holocene environmental changes along the Victoria Land Coast. Memoirs National Institute of Polar Research (Japan). Series C. Earth Sciences, Special Issue, 50, 85-107.

Baroni C., 2013. Climate Change Impacts on Cold Climates. In: Shroder J. (Editor in Chief), Giardino M., Harbo J. (Eds.), Treatise on Geomorphology. Academic Press, San Diego, CA, vol. 8, Glacial and Periglacial Geomorphology, 430-459. doi: 10.1016/B978-0-12-374739-6.00222-0

Baroni C., Hall B.L., 2004. A new Holocene relative sea-level curve for Terra Nova Bay, Victoria Land. Journal of Quaternary Science, 19, 377-396.

Baroni C., Orombelli G., 1991. Holocene raised beaches at Terra Nova Bay, Victoria Land, Antartica. Quaternary Research, 36, 157-177.

Baroni C., Orombelli G., 1994. Abandoned penguin rookeries as Holocene paleoclimatic indicators in Antarctica. Geology, 22, 23-26.

Bart P.J., Anderson J.B., Trincardi F., Shipp S.S., 2000. Seismic data from the Northern basin, Ross Sea, record extreme expansions of the East Antarctic Ice Sheet during the late Neogene. Marine Geology 166, 31–50.

Bentley M.J., Ocofaigh C., Anderson J.B., Conway H., Davies B., Graham A.G.C., Hillenbrand C.D., Hodgson D.A., Jamieson S.R., Larter R.D., Mackintosh A., Smith J.A., Verleyen E., Ackert R.P., Bart P.J., Berg S., Brunstein D., Canals M., Colhoun E.A., Crosta X., Dickens W.A., Domack E., Dowdeswell J.A., Dunbar R., Ehrmann W., Evans J., Favier V., Fink D., Fogwill C.J., Glasser N.F., Gohl K., Golledge N.R., Goodwin I., Gore D.B., Greenwood S.L., Hall B.L., Hall K., Hedding D.W., Hein A.S., Hocking E.P., Jakobsson M., Johnson J.S., Jomelli V., Jones R.S., Klages J.P., Kristoffersen Y., Kuhn G., Leventer A., Licht K., Lilly K., Lindow J., Livingstone S.J., Mass G., McGlone M.S., McKay R.M., Melles M., Miura H., Mulvaney R., Nel W., Nitsche F.O., O'Brien P.E., Post A.L., Roberts S.J., Saunders K.M., Selkirk P.M., Simms A.R., Spiegel C., Stolldorf, T.D., Sugden D.E., van der Putten N., van Ommen T., Verfaillie D., Vyverman W., Wagner B., White D.A., Witus A.E., Zwartz D., 2014. A Community-based Geological Reconstruction of Antarctic Ice Sheet Deglaciation since the Last Glacial Maximum Quaternary Science Reviews, 100, 1-9.

Biuw M., Boehme L., Guinet C., Costa D., Charrassin J.-B., Roquet F., Bailleul F., Meredith M., Thorpe S., Tremblay Y., McDonald B., Park Y.-H., Rintoul S.R., Bindoff N., Goebel M., Crocker D., Lovell P., Nicholson J., Monks F., Fedak M. A., 2007. Variations in behavior, condition of a Southern Ocean top predator in rela-

tion to in situ oceanographic conditions. Proceedings of the National Academy of Sciences of the United States of America 104 (34), 13705–13710.
Bonfiglio L., Cassoli P.F., Mallegni F., 1986. Neanderthal parietal, vertebrate fauna, and stone artifacts from the Upper Pleistocene deposits of Contrada Ianni di SanCalogero (Catanzaro, Calabria, Italy). American Journal of Physical Anthropology, 70 (2), 241–250.
Borowicz A., McDowall P., Youngflesh C. Sayre-McCord T., Clucas G., Herman R., Forrest S., Rider M., Schwaller M., Hart T., Jenouvrier S., Polito M.J., Singh H., Lynch H.J., 2018. Multi-modal survey of Adélie penguin mega-colonies reveals the Danger Islands as a seabird hotspot. Scientific Reports, 8, 3926. doi.org/10.1038/s41598-018-22313-w
Campbell I.B., Claridge G.C., 1987. Antarctica: Soils, Weathering Processes, Environment. Elsevier, Amsterdam, 368 pp.
Cape Roberts Science Team, 2000. Studies from the Cape Roberts Project, Ross Sea, Antarctica. Initial Report on CRP-3, Terra Antartica 7, 209 pp.
Clark P.U., Mix A.C., 2002. Ice sheets and sea level of the Last Glacial Maximum, Quaternary Science Reviews, 21 (1-3), 1-7. doi: 10.1016/S0277-3791(01)00118-4
Conway H., Hall B.L., Denton G.H., Gades A.M., Waddington E.D., 1999. Past and future grounding-line retreat of the west antarctic ice sheet. Science, 286 (5438), 280-283. doi:10.1126/science.286.5438.280
Croxall J.P., Trathan P.N., Murphy E.J., 2002. Environmental change and Antarctic Seabird populations. Science, 297, 1510-1514.
Dawson, A.G., 1992. Ice Age Earth: Late Quaternary Geology and Climate. Routledge, London, 299 pp.
De Bruyn M., Hall B.L., Chauke L.F., Baroni C., Koch P.L., Hoelzel A.R., 2009. Rapid Response of a Marine Mammal Species to Holocene Climate and Habitat Change. PLoS Genetics, 5 (7), e1000554. doi:10.1371/journal.pgen.1000554
De Bruyn M., Pinsky M.L., Hall B., Koch P., Baroni C., Hoelzel A.R., 2014. Rapid increase in southern elephant seal genetic diversity after a founder event. Proceedings - Royal Society. Biological Sciences, 281, 20133078-20133085. doi: 10.1098/ rspb.2013.3078
Denton G.H., Hughes T.J., 1981. The Last Great Ice Sheets, Wiley, New York, 484 pp.
Denton G.H., Hughes T.J., 2002. Reconstructing the Antarctic ice sheet at the lastglacial maximum. Quaternary Science Reviews 21, 193–202.
Denton G.H., Prentice M.L., Burckle L.H., 1991. Cainozoic history of the Antarcticice sheet. In: Tingey R.J. (Ed.), The geology of Antarctica. Oxford University Press, Oxford, 365–433.
Emslie S.D., 1995. Age and taphonomy of abandoned penguin rookeries in the antarctic peninsula region. Polar Record, 31 (179), 409-418. doi:10.1017/ S0032247400027388
Emslie S.D., 2001. Radiocarbon dates from abandoned penguin colonies in the Antarctic Peninsula region. Antarctic Science 13 (3), 289–295.
Emslie S.D., Coats L., Licht K., 2007. A 45,000 yr record of Adélie penguins and climate change in the Ross Sea, Antarctica. Geology, 35, 61-64.
Emslie S.D., Fraser W., Smith R.C., Walker W., 1998. Abandoned penguin colonies and

environmental change in the Palmer Station area, Anvers Island, Antarctic Peninsula. Antarctic Science, 10 (3), 257-268. doi:10.1017/s0954102098000352

Emslie S.D., McDaniel J.D., 2002. Adélie penguin diet, climate change during the middle to late Holocene in northern Marguerite Bay, Antarctic Peninsula. Polar Biology 25(3), 222–229.

Emslie S.D., Woehler E.J., 2005. A 9000-year record of Adélie penguin occupation, diet in the Windmill Islands, East Antarctica. Antarctic Science 17 (1), 57–66.

EPICA-Community-Members, 2004. Eight glacial cycles from an Antarctic ice core. Nature 429, 623–628.

EPICA-Community-Members, 2006. One-to-one coupling of glacial climate variability in Greenland and Antarctica. Nature 444, 195–198.

Fairbanks R.A., 1989. 17,000-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. Nature 342, 637–642. doi: 10.1038/342637a0

Fischer H., Severinghaus J., Brook E., Wolff E., Albert M., Alemany O., Arthern R., Bentley C., Blankenship D., Chappellaz J., Creyts T., Dahl-Jensen D., Dinn M., Frezzotti M., Fujita S., Gallee H., Hindmarsh R., Hudspeth D., Jugie G., Kawamura K., Lipenkov V., Miller H., Mulvaney R., Parrenin F., Pattyn F., Ritz C., Schwander J., Steinhage D., van Ommen T., Wilhelms F., 2013. Where to find 1.5 million yr old ice for the IPICS "Oldest-Ice" ice core. Climate of the Past, 9, 2489–2505. doi: 10.5194/cp-9-2489-2013

Forcada J., Trathan P.N., 2009. Penguin responses to climate change in the Southern Ocean. Global Change Biology 15, 1618–1630.

Forcada J., Trathan P.N., Reid K., Murphy E.J., Croxall J.P., 2006. Contrasting population changes in sympatric penguin species in association with climate warming. Global Change Biology, 12, 411-423.

Fretwell P., and 59 altri, 2013. Bedmap2: improved ice bed, surface and thickness datasets for Antarctica. Cryosphere 7, 375e393. doi: 10.5194/tc-7-375-2013.

Fuller E., 1999. The Great Auk. Harry N. Abrams, Inc. Publishers, New York, 448 pp.

Gardner N., Hall B., Wehmiller J., 2006. Pre-holocene raised beaches at Cape Ross, southern Victoria Land, Antarctica. Marine Geology, 229 (3-4), 273-284. doi:10.1016/j.margeo.2006.01.006

Grieve S., 1885. The Great Auk, or Garefowl: Its History, Archaeology, Remains. Thomas C. Jack, London, 141 pp.

Hall B.L., Baroni C., Denton G.H., 2004. Holocene relative sea-level history of the Southern VL Coast, Antarctica, Global and Planetary Change, 42, 241-263.

Hall B.L., Hoelzel A.R., Baroni C., Denton G.H., Le Boeuf B.J., Overturf B., Töpf A.L. 2006. Holocene elephant seal distribution implies warmer-than-present climate in the Ross Sea. Proc. National Academy of Sciences of the USA (PNAS), 103, 10213-10217.

Harrington H.J., 1960. Adélie penguin rookeries in the Ross Sea region. Notornis 9 (2), 33–39.

Heine J.C., Speir T.W., 1989. Ornithogenic soils of the Cape Bird Adélie penguin rookeries, Antarctica. Polar Biology 10, 89–99.

Hindell M.A., Sumner M., Bestley S., Wotherspoon S., Harcourt R.G., Lea M.,

Alderman R., McMahon C.R. 2017. Decadal changes in habitat characteristics influence population trajectories of southern elephant seals. Global Change Biology, 23 (12), 5136-5150. doi: 10.1111/gcb.13776

Huybrechts P., 2002. Sea-level changes at the LGM from ice-dynamic reconstructions of the Greenland and Antarctic ice sheets during the glacial cycles. Quaternary Science Reviews 21 (1-3), 203-231.

IPCC, 2007. Climate Change 2007: The Physical Science Basis. Cambridge University Press, Cambridge, 996 pp.

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Cambridge University Press, Cambridge, UK and New York, NY, USA, 1535 pp.

Jones C.W., Risi M.M., Bester M.N., 2020. Local extinction imminent for southern elephant seals Mirounga leonina at their northernmost breeding site, Gough Island—South Atlantic Ocean. Polar Biol 43, 893–897. doi: 10.1007/s00300-020-02679-2

Jouzel J., Masson-Delmotte V., Cattani O., Dreyfus G., Falourd S., Hoffmann G., Minster B., Nouet J., Barnola J.M., Chappellaz J., Fischer H., Gallet J.C., Johnsen S., Leuenberger M., Loulergue L., Luethi D., Oerter H., Parrenin F., Raisbeck G., Raynaud D., Schilt A., Schwander J., Selmo E., Souchez, R. Spahni R., Stauffer B., Steffensen J.P., Stenni B., Stocker T.F., Tison J.L., Werner M., Wolff E., 2007. Orbital and millenial Antarctic climate variability over the past 800,000 years. Science 317, 793–796.

Koch P.L., Hall B.L., de Bruyn M., Hoelzel A.R., Baroni C., Salvatore M.C., 2019. Mummified and skeletal southern elephant seals (*Mirounga leonina*) from the Victoria Land coast, Ross Sea, Antarctica. Marine Mammal Science, 35 (3), 934-956. doi:10.1111/mms.12581

Lambert D.M., Millar C.D., Swaminathan S., Baroni C., 2010. Evolution on a frozen continent. American Scientist, 98, 386-393. doi: 10.1511/2010.86.386

Lambert D.M., Ritchie P.A., Millar C.D., Holland B., Drummond A.J., Baroni C., 2002. Rates of Evolution in Ancient DNA from Adélie Penguins. Science, 295, 2270-2273.

Lambert F., Delmonte B., Petit J.R., Bigler M., Kaufman P.R., Hutterli M., Stocker T.F., Ruth U., Steffensen J.P., Maggi V., 2008. Dust–climate couplings over the past 800,000 years from the EPICA Dome C ice core. Nature 452, 616–619.

Le Boeuf, B.J., Laws R.M., (Eds.), 1994. Elephant Seals: Population Ecology, Behavior, and Physiology. Berkeley, University of California Press, 414 pp.

Lorenzini S., Baneschi I., Fallick A.E., Salvatore M.C., Zanchetta G., Dallai L., Baroni C., 2012. Insights into the Holocene environmental setting of Terra Nova Bay region (Ross Sea, Antarctica) from oxygen isotope geochemistry of Adélie penguin eggshells. The Holocene, 22, 63-69. doi: 10.1177/ 0959683611409773

Lorenzini S., Olmastroni S., Pezzo F., Salvatore M.C., Baroni C, 2009. Holocene Adélie Penguin diet in Victoria Land, Antarctica. Polar Biology, 32 (7), 1077-1086. doi: 10.1007/s00300-009-0607-4

Lorenzini S., Baroni C., Fallick A.E., Baneschi I., Salvatore M.C., Zanchetta G., Dallai L., 2010. Stable isotopes reveal Holocene changes in the diet of Adélie penguins in Northern Victoria Land (Ross Sea, Antarctica). Oecologia, 164 (4), 911-919. doi: 10.1007/s00442-010-1790-2

Lynch H.J., LaRue M.A., 2014. First global census of the Adélie penguin. The Auk, 131, 457-466. doi: 10.1642/AUK-14-31.1

Masson-Delmotte V., Stenni B., Pol K., Braconnot P., Cattani O., Falourd S., Kageyama M., Jouzel J., Landais A., Minster B., Barnola J.M., Chappellaz J., Krinner G., Johnsen S., Röthlisberger R., Hansen J., Mikolajewicz U., Otto-Bliesner B., 2010. EPICA Dome C record of glacial and interglacial intensities. Quaternary Science Reviews, 29 (1–2), 113-128. doi: 10.1016/j.quascirev.2009.09.030.

McDaniel J.D., Emslie S.D., 2002. Fluctuations in Ade lie penguin prey size in the mid to late Holocene, northern Marguerite Bay, Antarctic Peninsula. Polar Biology 25 (8), 618–623.

McKay R., Browne G., Carter L., Cowan E., Dunbar G., Krissek L., Naish T., Powell R., Reed J., Talarico F., Wilch T., 2009. The stratigraphic signature of the late Cenozoic Antarctic Ice Sheets in the Ross Embayment. Bulletin of the Geological Society of America 121 (11-12), 1537–1561.

Mezgec K., Stenni B., Crosta X., Masson Delmotte V., Baroni C., Braida M., Ciardini V., Colizza E., Melis R., Salvatore M.C., Severi M., Scarchilli C., Traversi R., Udisti R., Frezzotti M. (2017) - Holocene sea ice variability driven by wind and polynya efficiency in the Ross Sea. Nature Communications, 8, 1-12. doi: 10.1038/s41467-017-01455-x

Michel, R., Ferriera, M.C.E., Shaefer, G.R., *et al.*, 2006. Ornithogenic Gelisols (Cryosols) from Maritime Antarctica. Soil Science Society of America Journal 70, 1370–1376

Millar C.D., Dodd A., Anderson J., Gibb G.C., Ritchie P.A., Baroni C., Woodhams M.D., Hendy M.D., Lambert D.M. (2008) - Mutation and Evolutionary Rates in Adélie Penguins from the Antarctic. PLoS Genetics 4(10): e1000209. doi: 10.1371/journal.pgen.1000209

Morlighem M. *et al.* 2017, *IceBridge BedMachine Greenland, Version 3*. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center (updated 2018) doi: 10.5067/2CIX82HUV88Y.

Nakada M., Kimura R., Okuno J., Moriwaki K., Miura H., Maemoku H., 2000. Late Pleistocene and Holocene melting history of the Antarctic ice sheet derived from sea-level variations. Marine Geology, 167 (1-2), 85-103.

Naish T., Powell R., Levy R., Wilson G., Scherer R., Talarico F., Krissek L., Niessen F., Pompilio M., Wilson T., Carter L., DeConto R., Huybers P., McKay R., Pollard D., Ross J., Winter D., Barrett P., Browne G., Cody R., Cowan E., Crampton J., Dunbar G., Dunbar, N., Florindo F., Gebhardt C., Graham I., Hannah M., Hansaraj D., Harwood D., Helling D., Henrys S., Hinnov L., Kuhn G., Kyle P., Läufer A., Maffioli P., Magens D., Mandernack K., McIntosh W., Millan C., Morin R., Ohneiser C., Paulsen T., Persico D., Raine I., Reed J., Riesselman C., Sagnotti L., Schmitt D., Sjunneskog C., Strong P., Taviani M., Vogel S., Wilch T., Williams T. 2009. Obliquity-paced pliocene west antarctic ice sheet oscillations. Nature, 458 (7236), 322-328. doi:10.1038/nature07867

Orombelli G., Baroni C., Denton G.H. (1990) - Late Cenozoic glacial history of the Terra Nova Bay Region, northern Victoria Land, Antarctica. Geografia Fisica e Dinamica Quaternaria, 13 (2), 139-163.

Parks M., Subramanian S., Baroni C., Salvatore M.C., Zhang G., Millar C.D., Lambert D.M., 2015. Ancient population genomics and the study of evolution. Philosophical Transactions Royal Society London Series B: Biological Sciences, 370, 1-10. doi: 10.1098/rstb.2013.0381

Polito M., Emslie S.D., Walker W., 2002. A 1000-year record of Adélie penguin diets in the southern Ross Sea. Antarctic Science, 14, 327-332.

Pollard D., DeConto R.M., 2009. Modelling West Antarctic ice sheet growth, collapse through the past five million years. Nature, 458 (7236), 329-332.

Ritchie P.A., Millar C.D., Gibb G.C., Baroni C., Lambert D.M., 2004. Ancient DNA Enables Timing of the Pleistocene Origin and Holocene Expansion of Two Adélie Penguin Lineages in Antarctica. Molecular Biology and Evolution, 21 (2), 240-248. doi: 10.1093/molbev/msh012

Santora J.A., LaRue M.A., Ainley D.G., 2020. Geographic structuring of Antarctic penguin populations. Global Ecology and Biogeography, 1-13. doi: 10.1111/geb.13144

Serjeantson D., 2001. The Great Auk, the Gannett: a Prehistoric Perspective on the Extinction of the Great Auk. International Journal of Osteoarchaeology 11, 43–55.

Shepherd L.D., Millar C.D., Ballard G., Ainley D.G., Wilson P.R., Haynes G.D., Baroni C., Lambert D.M., 2005. Microevolution and mega-icebergs in Antarctica. Proceedings of the National Academy of Sciences of the United States of America (PNAS), 102 (46), 16717-16722. doi:10.1073/pnas.0502281102

Simas F.N.B., Schaefer C.E., Albuquerque Filho, M.R., *et al.*, 2008. Genesis, properties and classification of cryosols from Admiralty Bay, maritime Antarctica. Geoderma 144, 116–122.

Simas, F.N.B., Schaefer, C.E., Melo, V.F., *et al.*, 2007. Ornithogenic cryosols from maritime Antarctica: phosphatization as a soil forming process. Geoderma138, 191–203.

Simmonds M.P., Isaac S.J., 2007. The impacts of climate change on marine mammals: early signs of significant problems. Oryx 41, 19–26. doi: 10.1017/S00306053 07001524

Speir T.W., Cowling J.C., 1984. Orthogenetic soils of the Cape Bird Adelie penguin rookeries, Antarctica. Polar Biology 2, 207-212.

Stokstad, E., 2007. Boom and bust in a Polar hot zone. Science 315 (5818), 1522–1523. doi: 10.1126/science.315.5818.1522.

Subramanian S., Denver D.R., Millar C.D., Heupink T., Aschrafi A., Emslie D.S., Baroni C., Lambert D.M., 2009. High mitogenomic evolutionary rates and time dependency. Trends in Genetics, 25 (11), 482-486. doi: 10.1016/j.tig.2009.09.005

Syroechkovsky E.E., 1959. The role of animals in the formation of primary soils under the conditions of circumpolar regions of the earth (Antarctica). Zoologicheskiy Zhurnal 38, 1770-1775.

Tatur A., Myrcha A., 1989. Soils and vegetation in abandoned penguin rookeries. Proceeding NIPR Symposium Polar Biology 2, 181-189.

Taylor R.H., Wilson P.R., Thomas B.W., 1990. Status and trends of Adélie Penguin populations in the Ross Sea region. Polar Record 26, 293-304.

Ugolini F.C., 1972. Ornithogenic soils of Antarctica, Antarctic Research Series, 20, 181-193.

Wassmann P., Duarte C.M., Agusti S., Sejr M.K., 2011. Footprints of climate change in the Arctic marine ecosystem. Global Change Biology 17 (2), 1235-1249.

Williams R., McEldowney A., 1990. A guide to the fish otoliths from waters off the Australian Antarctic Territory, ANARE Research Notes 75, Antarctic Division Australia.

Woehler E.J., 1993. The Distribution and Abundance of Antarctic and Subantarctic Penguins. Scientific Committee on Antarctic Research, Cambridge, UK.

Zachos J.C., Pagani M., Sloan L., Thomas E., Billups K., 2001. Trends, rhythms, aberrations in global climate 65 Ma to present. Science 292, 686–693.

Zhu R., Sun L., Yin X., Xie Z., Liu X., 2005. Geochemical evidence for rapid enlargement of a gentoo penguin colony on Barton Peninsula in the maritime Antarctic. Antarctic Science 17, 11-16.

COMMEMORAZIONI

CARLO SOAVE

s.c. MARTIN KATER (*)

(Adunanza del 24 ottobre 2019)

Il professor Carlo Soave nasce a S. Benedetto Po (MN) il 31 luglio 1941 e muore improvvisamente a Milano il 10 marzo 2019.

Terminati gli studi classici presso il Liceo Berchet di Milano, nel novembre 1966 consegue la Laurea in Medicina (cum laude) presso l'Università degli Studi di Milano.

All'esercizio della professione medica preferisce la ricerca e si dedica da subito ad approfondire le sue competenze nel campo della Genetica lavorando presso il laboratorio del prof. Giorgio Bernardi "Centre de Récherches sur les Macromolécules", Strasburgo (Francia) dove "scopre" la biochimica e comincia la sua attività di ricerca studiando i DNA satelliti e gli enzimi idrolitici del DNA.

Dal 1970 al 1983 lavora come Ricercatore C.N.R. presso l'Istituto Biosintesi Vegetali di Milano dove inizia una collaborazione con il prof. Francesco Salamini sulla genetica e biochimica delle proteine di riserva del mais, culminata con la scoperta del sistema poligenico preposto alla sintesi di queste proteine ed alla identificazione e clonaggio dei geni regolatori. Successivamente diventa direttore pro tempore della sezione di Bergamo dell'Istituto per la Cerealicoltura del Ministero dell'Agricoltura, dove però resta solo pochi anni perché nel 1986 diventa Professore ordinario di "Genetica Agraria" presso l'Università della Basilicata, dove rimane fino al 1991. Nel periodo in cui era in Basilicata, avendo la famiglia a Milano, si spostava frequentemente e cercava di

(*) Dipartimento di Bioscienze, Università degli Studi di Milano, Italy.
E-mail: martin.kater@unimi.it

mantenere una collaborazione con i botanici dell'Università di Milano (proff. Marziani, Longo, Galli).

Dal 1992 è stato professore ordinario nel settore scientifico disciplinare BIO/04 - Fisiologia Vegetale presso l'Università degli Studi di Milano. L'attività scientifica e i campi di ricerca più recenti di Carlo Soave sono stati focalizzati sui seguenti temi: studi sul 5S-RNA nelle piante; studi genetici, biochimici e fisiologici sulle proteine di riserva nel mais e sui geni che controllano la loro espressione; studi su determinazione del sesso e differenziazione in *Asparagus officinalis*; sviluppo di procedure selettive e isolamento di mutanti di trasporto in *Arabidopsis thaliana* nel ambito della resistenza alla tossina fusicoccina; studi su metabolismo ed omeostasi del ferro in mais ed *Arabidopsis*; studi su mutanti fotosintetici.

La fotosintesi è stata pure una delle grandi passioni di Carlo Soave nell'ambito della fisiologia vegetale: ha infatti studiato la possibilità di mutare le caratteristiche di resistenza allo stress, alla luce e alla bassa temperatura dei vegetali e ha dimostrato che tale resistenza è correlata al rallentamento del trasporto elettronico nel fotosistema. Negli ultimi 10 anni della sua carriera si è molto interessato ai temi di carattere ambientale. Un argomento su cui ho personalmente collaborato con lui è stata l'identificazione di specie vegetali potenzialmente adatte alla produzione di biocarburanti. Era molto interessato alla ricerca utilizzando piante, come la *Arundo donax*, che cresce su terreni poco adatti per l'agricoltura ma in grado di produrre molta biomassa; ne studiava la propagazione ma anche la modificazione genetica, per arrivare a piante transgeniche ottimalizzate per il loro eventuale uso nella produzione di biocarburante.

È stato docente di corsi di genetica, citogenetica, biologia generale, fisiologia vegetale, bioenergetica e biotecnologie alimentari.

È stato relatore di numerose tesi di laurea e dottorato di ricerca. Ha svolto anche diversi incarichi di tipo scientifico-gestionale: è stato direttore del Dipartimento di Biologia all'Università degli Studi di Milano (1998-2004), Presidente della Società Italiana di Fisiologia Vegetale (2001-2004), membro del Consiglio del Dipartimento del CRA (Consiglio per la Ricerca in Agricoltura, adesso CREA), coordinatore del collegio dei docenti di dottorato in Biologia Vegetale (2005-2009), delegato del Rettore per gli Orti Botanici, per la loro valorizzazione e conservazione. Gli orti botanici gestiti dall'Università degli Studi di Milano sono tre, di cui uno, il più noto, è quello di Brera, ma

un altro molto bello si trova a Città Studi. E' proprio l'Orto botanico di Città Studi il luogo dove Soave conduceva le sue ricerche di fisiologia vegetale, e raccoglieva nel corso degli anni, tante specie vegetali significative per la Lombardia. E' stato Presidente della Fondazione Aqualab (dal 2010) e Socio Corrispondente dell'Istituto Lombardo (2006) dove diviene Membro Effettivo nel 2015.

Carlo Soave è autore di 151 pubblicazioni, per la gran parte su riviste internazionali.

La divulgazione scientifica è sempre stato uno dei temi di particolare interesse ed importanza cui Soave ha dedicato gran parte del suo tempo, anche dopo il suo pensionamento. Ha organizzato corsi, convegni, mostre in centri culturali e scuole. Tre di questi convegni/mostre sono stati realizzati proprio nell'ambito delle attività dell'Istituto Lombardo: *Una storia da scoprire: l'agricoltura all'Istituto Lombardo nell' 800*, *Arte nella Scienza*, con modelli di frutti in gesso e cera, e *La bellezza dei fiori*, realizzata, oltre che presso l'Università degli Studi di Milano, anche presso l'Istituto Lombardo. Il tema di quest' ultima mostra sottolineava come i fiori facessero parte della cultura italiana ed europea e come, in definitiva, essi vengano usati come simboli in tutte le culture del mondo.

Personalmente sono stato per anni il coordinatore nazionale per l'Italia per l'evento *Fascination of Plants Day*, giornata dedicata alla ricerca sulle piante, e Carlo Soave era sempre molto attivo nell'organizzazione di questo evento, credendo profondamente nella importanza della diffusione di ogni tipo di conoscenza. Per questo aveva ideato una mostra sulla domesticazione delle piante selvatiche rivolta a studenti, insegnanti e pubblico generico e concretizzatasi in numerosi scritti (es. Naturale, artificiale e coltivato: l'antico dialogo dell'uomo con la natura). Questo mostra evidenza che tutte le piante usate nella agricoltura sono un prodotto della selezione umana (domesticazione). Queste piante non sono "naturali" e sono per la loro sopravvivenza completamente dipendenti dall' uomo. Viceversa, l'uomo basa la sua alimentazione sull'utilizzo di queste piante domesticate. Carlo Soave era convinto che questo messaggio era di fondamentale importanza per far comprendere al pubblico l'importanza delle biotecnologie vegetali moderne. Far comprendere che le piante che consumiamo sono già state geneticamente modificate dall'uomo durante gli ultimi 10.000 anni era uno dei suoi obiettivi comunicativi. Nel futuro porteremo avanti ogni anno questa interessante mostra che Carlo Soave ha creato.

Il rapporto tra la biologia evoluzionistica e la genetica è stato tra i suoi interessi più recenti : su questo tema aveva pubblicato alcuni saggi ed era impegnato nella scrittura di un libro, la cui edizione viene curata dai colleghi De Bernardi e Fascio: "La Biologia evoluzionistica tra scienza e storia".

Carlo è sempre stato un grande fautore della buona scuola, appassionato come era dell'educazione dei giovani sempre coltivata come docente universitario. Personalità equilibrata ed autorevole era stimato e apprezzato da tutti i colleghi.

Un commento finale: il percorso di studi e di ricerca di Carlo Soave si delinea chiaramente ed emerge come esempio di un percorso altamente interdisciplinare: medico, genetista, botanico, fisiologo vegetale, ambientalista, anche filosofo: il tutto sotteso da un filo conduttore che appare essere la genetica vegetale come disciplina da lui preferita e nella quale può essere considerato un gran comunicatore.

Istituto Lombardo (Rend. Scienze) 153, 247-258 (2019)

LUCA CAVALLI SFORZA E I GENI DEGLI EUROPEI

s.c. GUIDO BARBUJANI (*)

(Adunanza del 6 giugno 2019)

È davvero un motivo di orgoglio essere stato invitato a parlare di Luca Cavalli Sforza: un grande maestro ma anche un italiano di straordinaria intelligenza e di singolari capacità imprenditoriali, grande scienziato ma anche persona che è riuscita a cambiare il volto della ricerca in genetica promuovendo progetti di grande portata, e conducendoli al successo. So che alcuni dei presenti non hanno familiarità con il suo lavoro, e quindi parlerò rapidamente di alcuni dei campi di cui Cavalli Sforza si è occupato, per poi concentrarmi su un tema in cui ha dato un contributo fondamentale e su cui si continua a lavorare: i geni degli europei. Vi dirò anche che gli ultimi studi demoliscono il mito secondo il quale gli europei hanno, e hanno sempre avuto, la pelle bianca. Oggi sappiamo che non è così: lo studio del DNA antico ci dice che, fino a poche migliaia di anni fa, *gli europei non erano quelli con la pelle bianca*, come siamo abituati a pensare.

Cavalli Sforza è nato il 25 gennaio del 1922 ed è morto l'anno scorso a 96 anni, lasciando la propria impronta sulla ricerca genetica a livello mondiale. La prima cosa importante che vorrei ricordare risale agli anni '50, quando aveva trent'anni. Fu a Cambridge, dove si occupava di batteri: era laureato in Medicina e i batteri gli interessavano dal punto di vista clinico, ma era persona di singolare curiosità e dopo un po' cominciò a interessarsi a questioni relative ai batteri che hanno poco a che vedere con le malattie che questi provocano e molto invece

(*) Dipartimento di Scienze della Vita e Biotecnologie, Università degli Studi di Ferrara, Italy. E-mail: bjg@unife.it

con l'evoluzione, cioè il tema di cui successivamente si è occupato in maniera prevalente. Ha trattato in particolare la sessualità dei batteri, un tema delicato dell'evoluzione. Negli anni Cinquanta si scopre che anche nei batteri ci sono delle forme di sessualità, cioè di scambio di pezzi di DNA. Una delle forme batteriche più abili in questo scambio fa parte della specie *Escherichia coli*, un ceppo di batteri che Cavalli Sforza individua e chiama *Hfr*.

Ma i suoi contributi sono soprattutto legati allo studio della genetica umana. Negli anni '60 e '70 Cavalli Sforza comincia a rendersi conto che il paradigma darwiniano, in cui gli organismi sono variabili e si adattano all'ambiente perché i portatori di certe caratteristiche vantaggiose lasciano più discendenti, in realtà non spiega tutto. C'è infatti una componente fondamentale, che possiamo chiamare "caso", non mirata all'adattamento, che deriva da singoli eventi, unici e in un certo senso irripetibili, nella storia delle popolazioni. Può capitare ad esempio che ci sia una carestia e una popolazione si estingua o si restringa, mentre (altrove, o in un altro momento) può accadere che siano disponibili molte risorse e che quindi la popolazione si espanda. Cavalli Sforza è tra i primi a esplorare le conseguenze di questi fenomeni, che noi chiamiamo di "deriva genetica", sia lavorando in Africa con i Pigmei, su cui ha pubblicato molti lavori, sia lavorando in Val di Parma. Questo suo interesse lo porta, nel momento in cui ha inizio il grande progetto del Genoma Umano, a rendersi conto di un limite molto forte. Il genoma umano è il complesso del nostro DNA; negli anni '90, quando viene lanciato questo progetto di lettura e di studio complessivo del genoma umano, sembra già molto ambizioso proporsi di descrivere completamente un genoma, i sei miliardi di basi o nucleotidi che lo compongono. Ma il genoma di chi? Ognuno di noi è infatti diverso dagli altri. Chi dobbiamo quindi scegliere? Da un lato è chiaro che avere a disposizione un genoma letto completamente è meglio che non averne nessuno; dall'altro, è altrettanto chiaro che se vogliamo capire perché certe persone hanno la pressione alta e certe l'hanno bassa; perché a certe persone l'aspirina fa bene, ad altri non fa niente e ad altri addirittura fa male; perché c'è chi digerisce il lattosio e chi no; per rispondere a tante domande come questa dobbiamo fare dei confronti. Un solo genoma non basta, e Luca Cavalli Sforza interviene vigorosamente per promuovere lo sviluppo in parallelo di un altro progetto, il "Progetto della Diversità Genomica Umana". Ottiene di farlo finanziare, si costituisce a Parigi un centro nel quale vengono raccolte

linee cellulari di 52 popolazioni di tutto il mondo. Linee cellulari significa cellule che continuano a riprodursi, per cui chiunque nell'ambito della genetica potrà nei decenni successivi accedere a questi campioni e studiarne un piccolo pezzo, e così contribuire alla conoscenza del grande mosaico della nostra diversità genomica. È interessante vedere la mappa di queste 52 popolazioni perché costituiscono anche una descrizione dei viaggi di Cavalli Sforza. Quasi tutti questi campioni, se non li ha portati a casa lui personalmente, se li è infatti procurati attraverso contatti di lavoro e, a volte, anche amicizie. Ci sono nove campioni del Pakistan; ce ne sono molti dell'Africa, il continente a cui si è dedicato di più; e ci sono tre importanti campioni italiani: i sardi, i toscani e i bergamaschi. Disponendo di un tale archivio, è possibile farsi tante domande: come mai siamo fatti così, come mai ci sono al nostro interno certe differenze? E, quando è evidente che non è stata solo la selezione a farci così, quando ci sono ottimi motivi per pensare che la selezione abbia sì giocato un ruolo, importante ma circoscritto; allora quello che resta è la traccia di eventi storici. Le nostre somiglianze e differenze sono il frutto delle migrazioni che hanno portato un piccolo gruppo di africani, che era stato per milioni di anni sul ciglio dell'estinzione (si stima che fossero qualche migliaio) ad espandersi a un certo punto con stupefacente rapidità, colonizzando tutto il pianeta e portando all'estinzione altre specie umane, fino a diffondersi dappertutto assieme ai nostri compagni di viaggio, cioè i topi, che con noi hanno colonizzato tutto il mondo.

Se si parla di migrazioni (oggi è ancora più evidente di quanto non lo fosse 20 anni fa) si finisce per parlare delle interazioni fra popolazioni che si muovono, e quindi di come guardiamo il nostro prossimo e di come lo classifichiamo. Quando l'ho conosciuto, Cavalli Sforza non pensava che valesse la pena mettere in discussione il concetto di razza umana. Abbiamo discusso parecchio su questo tema, ma quando si è convinto è partito con la sua formidabile capacità organizzativa e da qui è nata una serie di iniziative culminate non tanto in progetti scientifici specifici (la parte scientifica l'abbiamo fatta negli anni '90 e poi si è sostanzialmente conclusa) ma in una serie di libri molto importanti. I testi di Luca Cavalli-Sforza, spesso firmati insieme a Francesco Cavalli-Sforza, hanno contribuito come pochi altri a diffondere una cultura biologica di base e a mettere in discussione alcuni paradigmi dell'antropologia classica, per esempio l'idea che l'umanità possa essere divisa in gruppi razziali distinti. Oggi tra gli scienziati che si occupano

seriamente di questi argomenti non c'è più dubbio che *nella nostra specie non esistano gruppi razziali distinti*.

Vorrei adesso spendere alcune parole sullo studio degli europei, perché qui il lavoro di Luca Cavalli-Sforza porta a una svolta fondamentale. La genetica di popolazioni, fino agli anni '70, era un campo di studio interessante, con una base matematica molto forte, ma forse non molto divertente. Si trattava, sostanzialmente, di calcolare il rischio che due genitori sani, ma portatori di una malattia genetica, potessero trasmetterla ai figli. In quegli anni Cavalli Sforza si rende conto che nei nostri geni c'è ben di più: c'è un messaggio dal passato, la traccia di quello che è successo prima, il segno delle nostre vicende, delle nostre migrazioni e anche di fenomeni culturali che hanno determinato dei cambiamenti demografici. Nel 1978, sulla rivista *Science*, insieme ad Alberto Piazza e Paolo Menozzi, Luca Cavalli Sforza sviluppa il primo studio sistematico delle popolazioni europee. Il gruppo di Cavalli Sforza disponeva di alcuni software sofisticati, molto potenti per quell'epoca, fornitigli dalla Nato. Mi dicono che probabilmente gli americani, in quegli anni impegnati nella guerra del Vietnam, studiavano i geni dei vietnamiti per scoprire se sarebbe stato possibile costruire un'arma biologica che colpisse solo loro e non gli altri. Questo obiettivo si è poi rivelato implausibile, e oggi sappiamo che i DNA di tutti noi, compresi i vietnamiti, sono identici per oltre il 99%. Quando i dati genetici raccolti e il software sono stati messi a loro disposizione, Menozzi, Piazza e Luca Cavalli-Sforza hanno sviluppato metodi matematici per ottenere un'immagine tridimensionale della variabilità genetica degli europei. Nelle mappe si usano il verde e il marrone per indicare, rispettivamente, montagne e pianure. Allo stesso modo, Cavalli-Sforza e i suoi collaboratori ottengono un'immagine sintetica di come variano 38 geni negli europei (Per inciso: all'epoca sembravano tantissimi, ma oggi si lavora su milioni di varianti del DNA). Nella mappa si vede bene che queste varianti genetiche non sono distribuite a caso: c'è un gradiente che parte da sud-est con le tinte più scure e diventa sempre più chiaro, a mano a mano che ci si sposta verso nord-ovest. Dunque, qualcosa ha messo in fila le caratteristiche genetiche degli europei; ma cosa? Cavalli Sforza è persona di grandissima cultura, abituato a frequentare ambienti anche lontani dal suo. Presto si rende conto che esistono dati archeologici che corrispondono fedelmente a questo andamento. Dalla nostra origine fino a 10.000 anni fa, quindi da 180.000 anni, *Homo sapiens* si è alimentato delle cose che trovava spon-

taneamente in natura, dei frutti della caccia e della raccolta dei frutti delle piante. Cominciamo a produrre cibo solo da 10.000 anni, a concentrare semi, a produrre utensili per l'agricoltura, e contenitori di ceramica per conservare il cibo. Al complesso di reperti archeologici che rimandano all'attività agricola diamo il nome di neolitico, la nuova Età della Pietra. Se guardiamo le mappe che descrivono i più antichi ritrovamenti neolitici in Europa, troviamo che sono straordinariamente simili a quella che descrive le differenze dei nostri geni. Si tratta di una correlazione, il che vuol dire che due fenomeni procedono insieme, non necessariamente che uno causa l'altro. Ma fra gli archeologi si discute se la diffusione delle tecniche neolitiche, dal Vicino Oriente verso nord e ovest, sia frutto di scambi culturali o di una migrazione degli agricoltori. L'idea è che i manufatti neolitici possano essersi diffusi o perché qualcuno migrava e li portava con sé o perché venivano imitati o trasmessi attraverso il commercio; ma se il meccanismo fosse stato solo commerciale non ci sarebbero conseguenze a livello genetico. Se scambiate il telefonino con un cinese, il cinese continua ad avere i geni dei cinesi e voi i vostri. Dal momento che la diversità genetica europea forma un gradiente che abbraccia tutto il continente, la spiegazione più logica è che ci sia stata una grande migrazione che ha prodotto il fenomeno. Solo l'ipotesi della migrazione, non quella che siano stati contatti e scambi culturali a diffondere l'agricoltura, spiega perché la distribuzione dei geni si sovrapponga a quella della diffusione dell'agricoltura.

Si chiama modello di *Diffusione demica neolitica*: i primi agricoltori del Vicino Oriente si sono gradualmente diffusi, portando con sé le proprie tecnologie agricole, i propri geni (e forse qualcos'altro, le loro lingue Indo-Europee, a cui accennerò più avanti). Però in realtà ho semplificato troppo, e adesso cerco di rimediare. La migrazione, in sé e per sé, non basta a creare gradienti. Negli ultimi trent'anni, molti romeni sono immigrati in Italia, ma non per questo abbiamo gradienti genetici che dalla Romania, attraverso Serbia, Bosnia e Croazia, arrivino in Italia. Perché si formino gradienti, e Cavalli-Sforza elabora il complesso modello matematico che lo dimostra insieme ad Albert Ammerman, bisogna che la popolazione che si diffonde abbia un potenziale di crescita superiore a quella che incontra nel corso della migrazione. Se è così; se le popolazioni che praticavano l'agricoltura disponevano di più cibo, e quindi potevano crescere più di quelle che continuavano a cacciare, allora i loro geni hanno potuto viaggiare e

diluirsi poco alla volta, formando i gradienti che, 10mila anni dopo, Menozzi, Piazza e Cavalli-Sforza hanno identificato.

Dunque, ci sono vari motivi per credere che le caratteristiche genetiche degli europei odierni derivino da questo fenomeno di diffusione demica: un movimento migratorio proveniente dal Vicino Oriente. Di fatto, ci sono voluti anni per mettere d'accordo i tanti che, come archeologi o come genetisti, si sono occupati dell'evoluzione degli europei. Secondo Cavalli Sforza, c'è stata una prima espansione di popolazioni di cacciatori e raccoglitori, nel paleolitico. Venivano dall'Africa, erano nomadi, e hanno colonizzato (a basse densità, perché la caccia non forniva risorse sufficienti a crescere più di tanto) varie regioni dell'Europa. Poi, nel neolitico, lo sviluppo delle tecnologie per la produzione di cibo dà alle popolazioni del Vicino oriente, quella che a lungo si è chiamata la Mezzaluna fertile, un vantaggio decisivo; si espandono, si sovrappongono alle popolazioni preesistenti, in certi casi le sostituiscono. Oggi sappiamo che il modello di diffusione demica è quello che funziona meglio, ma ci siamo anche resi conto del fatto che non tutto era così semplice, che in diverse regioni d'Europa altre migrazioni, più localizzate, hanno avuto importanza.

La discussione interdisciplinare provocata dal modello di diffusione demica neolitica ha però messo in moto collaborazioni ancora più vaste. Nel 1988, insieme a Alberto Piazza, Paolo Menozzi e Joanna Mountain, Luca Cavalli-Sforza pubblica un articolo in cui compaiono due alberi che rappresentano graficamente le relazioni fra le popolazioni di tutto il mondo. Uno è un classico albero genetico, con le popolazioni dell'Eurasia, e poi delle Americhe e dell'Oceania, che si staccano dal tronco delle popolazioni africane; l'altro è un albero delle lingue parlate nel mondo. I due alberi a grandi linee si corrispondono, pur con una serie di eccezioni che andranno comprese: a maggiori differenze genetiche corrispondono maggiori differenze linguistiche. È nato tutto da un'intuizione di Charles Darwin, che nel suo libro del 1871, *L'origine dell'uomo e la selezione sessuale,* scrive appunto che la ricostruzione migliore delle relazioni linguistiche tra popolazioni potrebbe essere il loro albero evolutivo. In termini moderni, potremmo dire che la storia demografica dà forma alle differenze fra popolazioni, perché i processi che portano le popolazioni a diversificarsi (isolamento) o ad assomigliarsi (contatti) sono sostanzialmente gli stessi, sia a livello biologico, sia a livello culturale. Charles Darwin, come di consueto, ci aveva pensato prima degli altri. Solo negli anni Ottanta, però, prima

nel laboratorio di Robert Sokal in cui all'epoca lavoravo, a Stony Brook, e poco dopo nel laboratorio di Stanford di Luca Cavalli-Sforza, si raccolgono i dati, linguistici e genetici, necessari per verificare se e quanto Darwin avesse avuto ragione.

L'ingresso di genetisti e antropologi, Robert Sokal e Luca Cavalli-Sforza, nel dibattito sull'origine delle lingue non poteva non avere significative conseguenze. All'epoca, l'ipotesi tradizionale, dovuta soprattutto al lavoro di una linguista dell'epoca sovietica (lituana in realtà), Marija Gimbutas, era che le lingue IndoEuropee fossero arrivate in Europa dalle steppe dell'attuale Ucraina, intorno a 5.000 anni fa, nell'Età del Bronzo. Marija Gimbutas tracciava sulla mappa linee di diffusione delle lingue IndoEuropee, corrispondenti a possibili vie di migrazione, senza però disporre di alcun dato biologico. Si immaginava quindi un numero molto elevato di fenomeni migratori, ognuno dei quali aveva determinato la diffusione di una lingua. Un brillante archeologo inglese, Colin Renfrew, ha preso sul serio la teoria della diffusione demica. Si è chiesto quindi perché immaginarsi tante migrazioni dalle steppe, per nessuna delle quali era disponibile alcun dato biologico, quando ce n'era una, dimostrata da Cavalli-Sforza con convincenti dati genetici, che potrebbe spiegare come sono arrivati in Europa gli indoeuropei: non dall'Ucraina ma dal Vicino Oriente, e non nell'età del bronzo, ma nel neolitico. Renfrew propone che i primi contadini che hanno portato i propri geni e le proprie tecnologie in Europa, hanno portato anche le lingue. Si sono spostati a ovest dal Vicino Oriente, in Anatolia e poi nei Balcani e nel resto d'Europa, semplicemente perché a est c'erano altre popolazioni di contadini, che come loro stavano crescendo di numero e si sarebbero diffusi nel Nord Africa, in Arabia e in India.

Quello di Renfrew è un magnifico modello, in cui le informazioni provenienti da molti campi diversi – l'archeologia, la genetica, la linguistica – si combinano in maniera intellettualmente soddisfacente. Non ho tempo per entrare nei dettagli; posso dire che studi successivi l'hanno in buona parte confermato, ma anche dimostrato (e non è una sorpresa) che quando si studiano le popolazioni in particolare, o qualche aspetto del lessico e della sintassi, non sempre i conti tornano. Non c'è niente da fare: fenomeni complessi sono difficili da spiegare con cause semplici. Stiamo però vivendo in una fase in cui le conoscenze genetiche sono aumentate così tanto che possiamo porci domande fino a pochi anni fa fantascientifiche. Prima di entrare nell'ultima parte

di questa chiacchierata, vorrei dire due parole sul nostro genoma. L'abbiamo già detto: chiamiamo genoma il complesso del nostro DNA. Il genoma è stato spesso confrontato metaforicamente con un testo: è il libretto delle istruzioni delle nostre cellule; come un testo, è costituito da lettere che noi sappiamo riconoscere, sono 4 molecole che indichiamo con le lettere A, C, T e G, che formano lunghissime catene, i cromosomi. Dunque, conosciamo l'alfabeto del DNA, e ne conosciamo anche il lessico, perché all'interno del genoma dei mammiferi ci sono circa 20.000 geni. Ognuno di loro produce una o più proteine e noi li conosciamo quasi tutti. Perché allora non riusciamo ancora a prevedere il cancro o il diabete o la pressione alta? Perché tante di queste caratteristiche – patologiche e no; vorremmo per esempio capire qualcosa delle basi biologiche delle nostre capacità fisiche e intellettuali — non dipendono da un solo gene ma da tantissimi; e quindi il contributo di ciascun gene in percentuale diventa molto piccolo, e difficile da riconoscere. Quindi, proseguendo con la metafora del testo, del genoma non conosciamo ancora la sintassi, cioè come i vari geni interagiscano tra loro e con vari fattori ambientali che li influenzano. Però è un testo che riusciamo a leggere ed è lunghissimo. Sono 6 miliardi e mezzo di caratteri all'interno di una cellula; secondo una mia stima, *I promessi sposi* sono lunghi un milione e mezzo di caratteri! Quindi il libretto di istruzioni delle nostre cellule è pari a 4000 volumi, ciascuno grande come *I promessi sposi*. Come faccia la cellula nel giro di secondi o minuti a trovare la riga giusta in questa biblioteca e fare la cosa giusta è un'altra cosa che ci sfugge e che in futuro ci piacerebbe imparare. Ma leggendo per intero questo genoma abbiamo capito, per esempio, che due persone prese a caso, una ragazza che viene dall'Asia e un signore nord-europeo, con il loro aspetto diverso, hanno in comune, *in media*, il 99,9% del loro DNA. Quindi tutti i ragionamenti che facciamo noi evoluzionisti sulle differenze, fra individui e popolazioni, si concentrano su una frazione molto piccola del genoma, quella in cui troviamo delle differenze. Attenzione: in media significa che, in certi casi, queste differenze possono essere più piccole, e in altri più grandi. Il posto dove abbiamo trovato le differenze genetiche più grandi è l'Africa, il che conferma quello che ci dicono i fossili: l'Africa è un continente speciale, noi veniamo da lì.

Passerei ora a parlarvi del colore della pelle. Siamo abbastanza sicuri che, prima dell'agricoltura e della diffusione demica, l'Europa fosse abitata da cacciatori e raccoglitori che venivano anch'essi

dall'Africa: le prime tracce fossili di uomini come noi, in Europa, risalgono a 45.000 anni fa. Quando sono arrivati, l'Europa non era spopolata: c'era l'Uomo di Neandertal, il vero europeo, da 300.000 anni. La nostra convivenza con loro non è durata a lungo: qualche millennio, poi siamo rimasti solo noi. Come abbiamo visto, intorno a 8.000 anni fa, dal Vicino Oriente, dalla mezzaluna fertile, arrivano i primi contadini, che nel giro di pochi millenni colonizzano tutta l'Europa. Inoltre, la Gimbutas qui aveva visto giusto, studiando le differenze nei genomi delle popolazioni europee, e anche qualche resto umano preistorico in cui il DNA si è conservato, si trovano tracce di migrazioni nell'Età del Bronzo provenienti dalle steppe pontiche. Dunque c'è un continuo sovrapporsi di strati: prima i Neandertal, poi i cacciatori-raccoglitori paleolitici, sopra di loro i contadini neolitici, e poi, soprattutto nel nord Europa, le popolazioni dell'Età del Bronzo.

Lo studio del DNA nelle ossa fossili non è banale, ma si stanno facendo grandi progressi. Una nuova prospettiva è quella di ricostruire caratteristiche fisiche di persone di cui sono rimaste solo le ossa fossili: il gruppo sanguigno, la capacità di digerire il latte. Il colore della pelle è un altro dei caratteri che dipendono da tanti geni contemporaneamente, e quindi non facile da studiare. I geni della pigmentazione sono attivi all'interno di cellule della pelle che si chiamano melanociti. I melanociti producono due pigmenti, le melanine: una più scura, l'eumelanina e un'altra più sul giallo, la feomelanina. Le melanine vengono concentrate in granuli che poi dai melanociti passano alle cellule più superficiali dell'epidermide: il particolare colore della pelle di ciascuno di noi dipende dal rapporto fra eumelanina e feomelanina nei granuli, da quanti granuli abbiamo, da quanto sono grossi. C'entrano parecchi geni e anche parecchi fattori ambientali perché, come sapete, quando ci si mette al sole ci si abbronza. Si tratta di un'interazione complessa che non abbiamo ancora del tutto decifrato, ma in cui cominciamo a orientarci. Ma come si è evoluto il colore della pelle? La pelle non lascia fossili, ma la maggiore esperta di queste cose, un'antropologa americana, Nina Jablonski, ci propone un modello, cioè un'ipotesi. Se aprite il pelo di un gorilla o di uno scimpanzé, la pelle sotto è bianca e quindi c'è motivo di ritenere che i nostri primi antenati africani avessero la pelle bianca. Per qualche motivo, quando ci separiamo dagli scimpanzè, forse sei milioni di anni fa, cominciamo a perdere pelo. A stare senza protezione sotto il cielo africano, con un'alta emissione di raggi ultravioletti è più facile sviluppare un tumore e quindi c'è una

pressione darwiniana a sviluppare pelli sempre più scure. Quando compaiono le mutazioni che rendono la pelle più scura, queste mutazioni producono un vantaggio, un adattamento, e quindi si diffondono. A un certo punto, forse 70mila anni fa, qualcuno migra dall'Africa verso l'Asia e l'Europa. Nel nuovo ambiente, il meccanismo selettivo si inverte, perché alle basse esposizioni solari del nord Europa e del nord Asia con le pelli scure non si sviluppa abbastanza vitamina D, fondamentale nella gravidanza e nell'allattamento. Quindi nel nuovo ambiente diventa vantaggioso avere le pelli chiare, e le mutazioni che riducono la pigmentazione vengono promosse dalla selezione naturale e, attraverso le generazioni, si diffondono. Quando questo sia successo non lo sappiamo con esattezza; sembrerebbe, e a lungo è sembrato, logico, pensare che, siccome siamo in Europa da 70.000 anni o giù di lì, abbiamo avuto tutto il tempo perché le pelli si schiarissero. Negli articoli di Lina Jablonski non c'è una data precisa, ma le illustrazioni indicano che pensava che intorno a 50mila anni fa europei e asiatici avessero già le pelli chiare.

Non è così, e ce l'ha dimostrato lo studio dei genomi. L'anno scorso si sono studiate le ossa, datate a 10mila anni fa, trovate in Inghilterra, in località Cheddar, in una palude. Appartenevano a un individuo che è stato, con ogni probabilità, ucciso, come dimostra il foro ritrovato nel suo cranio. È stato possibile studiare il suo DNA, e, con metodi di intelligenza artificiale usati dalle polizie scientifiche, ricostruire approssimativamente il suo aspetto. L'uomo di Cheddar aveva pelle scura e occhi celesti: una combinazione oggi insolita che però, evidentemente, era presente in Inghilterra 10.000 anni fa. Quando la notizia viene pubblicata dai giornali, ci si rende conto che in realtà non era la prima volta che ci si trovava di fronte a una osservazione del genere. Qualche anno prima, nel sito spagnolo di La Braña, era stato scoperto un individuo vissuto 7.000 anni fa, anch'egli con la pelle scura e gli occhi celesti. Finora sono stati studiati i genomi di quattro di questo periodo, il Mesolitico; tutti quanti hanno la pelle scura, e tre di loro occhi celesti. Una mia collaboratrice, Gloria Gonzales Fortes, sta utilizzando gli algoritmi di intelligenza artificiale con cui si sono studiate le pigmentazioni di questi quattro individui. Prima di tutto, abbiamo verificato il margine d'errore, analizzando campioni di cui conoscevamo il colore della pelle, e verificando se corrispondesse o meno a quella predetta dall'algoritmo. Possiamo dire che il margine di errore sperimentale è del 4%. Studiando i genomi antichi disponibili, stiamo rendendoci conto che

fino al Mesolitico, intorno a 10.000 anni fa, indipendentemente dalla latitudine, le pelli erano tutte quante scure o molto scure. In seguito, nel nord cominciavano a comparire occasionalmente pelli molto più chiare. All'inizio del Neolitico sono ancora comuni le pelli scure, ma in Spagna compaiono pelli più chiare che nel medio Neolitico diventano ancora più chiare. Infine, nell'Età del Bronzo, ormai, gli europei hanno tutti la pelle chiara. Ancora una volta, il Neolitico è il momento di snodo fondamentale. Le prime presenze di pelli bianche sono documentate nel Caucaso intorno a 12.000 anni fa; quindi, probabilmente sono stati i discendenti di queste popolazioni, nel Vicino Oriente e in Anatolia, a portare le pelli bianche in Europa. È interessante che proprio negli stessi giorni in cui usciva il lavoro sul Cheddar Man l'allora aspirante Presidente della Giunta lombarda, oggi Presidente, Attilio Fontana, avesse invocato misure contro l'immigrazione per proteggere la razza bianca. A quanto pare, invece, le pelli bianche sono arrivate in Europa proprio grazie a immigranti.

Tutto questo, e altro, lo possiamo dire perché 40 anni fa Luca Cavalli Sforza ha capito che dentro il DNA delle nostre cellule, che allora non si poteva quasi nemmeno studiare, c'erano tante storie interessanti, e ci ha dato i primi strumenti per decifrarle. Oggi la tecnologia ha fatto enormi progressi, ma è a Luca Cavalli-Sforza che dobbiamo essere grati, perché è lui che ci ha permesso di muovere i primi passi in una direzione che promette di rivelarci aspetti del nostro passato che altrimenti sarebbero rimasti sconosciuti.

INDICE DEL VOLUME

FERRETTI G., Nota sulla non linearità del sistema respiratorio: come l'emoglobina lega Gaston Tissandier a Reinhold Messner 3

ALBINI A., Quattro chiacchiere a proposito della comunicazione delle scienze. La chimica 21

MARINI L.D., Matematica e calcolo scientifico 41

VARLAMOV A., RIGAMONTI A., Gastronomy universe by eyes of physicists 57

RAOS G., Separazione di carica in miscele di materiali organici: vale la legge di Coulomb? 83

BORSA F., The physics of Leonardo da Vinci 89

TARONI P., Spettroscopia e imaging in ottica diffusa: focus sul tumore della mammella 115

CAVALLERI O., *et al.*, Self-assembly di DNA su superficie: uno studio combinato nanolitografia/spettroscopia ottica 131

LARIZZA L., Comprendere le basi molecolari della disabilità intellettiva attraverso il modello dei neuroni di cellule staminali pluripotenti indotte 147

FALCIOLA L., PIFFERI V., TESTOLIN A., Quando l'insieme è meglio della somma: nanomateriali ibridi argento-titania ed oro-titania per sensori elettroanalitici 191

BARONI C., La ricolonizzazione dell'Antartide: storie di pinguini e foche, sentinelle dei cambiamenti climatici negli ultimi 10 mila anni 207

Commemorazioni

KATER M.
Carlo Soave 243

BARBUJANI G.
Luca Cavalli Sforza e i geni degli europei 247

INDICE PER MATERIA

CHIMICA

ALBINI A., Quattro chiacchiere a proposito della comunicazione delle scienze. La chimica 21

FALCIOLA L., PIFFERI V., TESTOLIN A., Quando l'insieme è meglio della somma: nanomateriali ibridi argento-titania ed oro-titania per sensori elettroanalitici 191

CHIMICA/FISICA

RAOS G., Separazione di carica in miscele di materiali organici: vale la legge di Coulomb? 83

FISICA

VARLAMOV A., RIGAMONTI A., Gastronomy universe by eyes of physicists 57

BORSA F., The physics of Leonardo da Vinci 89

TARONI P., Spettroscopia e imaging in ottica diffusa: focus sul tumore della mammella 115

CAVALLERI O., *et al.*, Self-assembly di DNA su superficie: uno studio combinato nanolitografia/spettroscopia ottica 131

MATEMATICA

MARINI L.D., Matematica e calcolo scientifico 41

MEDICINA

FERRETTI G., Nota sulla non linearità del sistema respiratorio: come l'emoglobina lega Gaston Tissandier a Reinhold Messner 3

LARIZZA L., Comprendere le basi molecolari della disabilità intellettiva attraverso il modello dei neuroni di cellule staminali pluripotenti indotte 147

Scienze naturali

Ecologia

Baroni C., La ricolonizzazione dell'Antartide: storie di pinguini e foche, sentinelle dei cambiamenti climatici negli ultimi 10 mila anni 207

Commemorazioni

Kater M.
Carlo Soave 243

Barbujani G.
Luca Cavalli Sforza e i geni degli europei 247

Finito di stampare nel mese di novembre 2020

Decreto Pres. Tribunale 10 gennaio 1961, n. 5504 registro cancelleria
Direttore responsabile: Adele Robbiati Bianchi