

Autori

Damiano Galimberti

Professore a contratto di Nutrigenomica e Medicina Anti-Aging; Specialista in Scienza dell’Alimentazione; Presidente AMIA (Associazione Medici Italiani Antiaging); Coordinatore SINE (Società Italiana di Nutrigenomica ed Epigenetica)

Giovanni Battista Gidaro

Biologo; MSc. Statistica Medica e Genomica, Consulente Scientifico; Segretario Generale SINE (Società Italiana di Nutrigenomica ed Epigenetica)

Vittorio Calabrese

Professore Ordinario di Biochimica Clinica e Biologia Molecolare Clinica e di Neurobiologia Clinica, Università degli Studi di Catania

Alessandro Gelli

Docente e Coordinatore Scientifico dei Corsi universitari Metodologie Antiaging e Antistress, “Sapienza” Università di Roma

Stefano Govoni

Professore Ordinario, Dipartimento di Scienze del Farmaco, Università degli Studi di Pavia

Federico Licastro

Professore Alma Mater presso Università di Bologna Alma Mater Studiorum, Dipartimento di Medicina Specialistica, Diagnostica e Sperimentale (DIMES)

Luisa Bernardinelli

Professore Ordinario di Statistica Medica, Dipartimento di Scienze del Sistema Nervoso e del Comportamento, Università degli Studi di Pavia

Davide Gentilini

Specialista in Genetica Medica, Istituto Auxologico Italiano IRCCS, Cusano Milanino (Milano)

Giovanni Malerba

Biologo; Professore Associato di Genetica Medica, Università degli Studi di Verona; Specialista in Genetica Medica

Collaboratori

Rosaria Ornella Bua

Dipartimento di Scienze Biomediche e Biotechnologiche, Facoltà di Medicina, Università di Catania

Antonio Mori

Specialista in Genetica Medica, Sezione di Biologia e Genetica, Dipartimento di Neuroscienze, Biomedicina e Movimento, Università degli Studi di Verona

Elisa Porcellini

Dipartimento di Medicina Specialistica, Diagnostica e Sperimentale (DIMES), Scuola di Medicina, Università degli Studi di Bologna

Prefazione

I rapidi cambiamenti che minacciano la salute mondiale richiedono una risposta altrettanto rapida e soprattutto rivolta al futuro. Grazie ai progressi compiuti nel controllo delle malattie infettive, le grandi epidemie dei prossimi anni difficilmente saranno simili a quelle che hanno flagellato il mondo in passato. Certo, il rischio di nuove pandemie, come ad esempio quelle influenzali, richiede una vigilanza costante, ma saranno le epidemie “invisibili” di cardiopatie, ictus, diabete, tumori e altre malattie croniche a far pagare il prezzo più alto nel prossimo futuro.

Agendo globalmente sui principali fattori di rischio, si può già ridurre gran parte del carico di morti premature, malattie e disabilità. Investendo nella prevenzione e migliorando il controllo delle malattie croniche, si potrebbe migliorare la qualità della vita e il benessere, a livello sia individuale sia globale.

Normalmente l’atteggiamento classico del medico tradizionale è di *wait and see*: agisce solo in caso di disturbi e sintomi conclamati. Occorre invece intervenire per tempo, preventivamente, sulle condizioni borderline: lo scopo è quello di prevenire, rallentare o invertire un invecchiamento precoce o troppo veloce o manifestazioni negative correlate allo stato di salute e/o di benessere fisico e mentale.

L’obiettivo della medicina deve consistere nel migliorare sempre più la “pagella” della vita, nel combattere contro il tempo per conservare il più a lungo possibile la libertà dalla malattia. In fondo, il compito del medico dovrebbe essere quello di curare i propri pazienti soprattutto quando stanno bene, di occuparsi di loro proprio quando sono in buona salute. Il buon medico, infatti, dovrebbe essere considerato come il custode della salute. Un obiettivo su tutti è mantenere il più a lungo possibile la giovinezza delle nostre cellule (o almeno di alcune) per evitare e prevenire le malattie degenerative legate all’invecchiamento, con il risultato di aumentare inoltre la durata della vita anche attraverso screening personalizzati, potenzialmente in grado di prevenire decessi e disabilità, oltre che di migliorare la qualità della vita.

Il rischio personale di sviluppare una malattia, infatti, dipende dall’interazione tra l’individuo, l’ambiente circostante e la suscettibilità individuale. Mentre la riduzione e il controllo dei fattori di rischio più modificabili e dei determinanti principali restano le pietre angolari dell’azione di prevenzione e controllo delle malattie croniche, in questi ultimi anni si sono sviluppati sempre più test per malattie monogeniche e poligeniche, finalizzati a individuare la predisposizione genetica verso le patologie più comuni. Si tratta di un modo per identificare i soggetti ad alto rischio, consentendo così interventi più precoci e mirati, da effettuare prima della comparsa dei sintomi o per sottoporre a test gli individui sintomatici per confermare la diagnosi e offrire nuove terapie. Pertanto è importante sensibilizzare le persone al bisogno di salute, al fine di mirare a un miglioramento della qualità della vita. In questo contesto l’approccio più innovativo è dato

Prefazione

dall'incrocio di due tipologie di informazioni: quelle che riguardano l'“oggi”, legate ai risultati di esami ematochimici e strumentali, e quelle che riguardano il nostro “ieri-oggi-domani”, cioè la nostra primaria costituzione (il DNA).

Tutti gli esseri umani sono dotati di occhi, tuttavia non esiste un paio di occhi identico all'altro. Queste differenze genetiche sono alla base di conseguenze cruciali per la nostra vita: non solo l'apparenza esteriore, ma anche la nostra salute dipendono da esse. Se il rischio è noto, si può fare qualcosa per prevenirlo. Per ciascuna malattia, il momento in cui si interviene è cruciale. La prevenzione è la migliore arma a disposizione per ottimizzare il proprio stato di salute ed è la strategia più sostenibile a lungo termine, che consente di affrontare contemporaneamente diversi potenziali disturbi e i rispettivi fattori di rischio.

Nel DNA non c'è il fato o il destino; esiste unicamente un'indicazione della suscettibilità personale verso problematiche correlate alla salute e alle performance psico-fisiche. Se i geni ci sono stati dati alla nascita, lo sviluppo di determinate malattie può comunque verificarsi in età avanzata. Se la costellazione genetica di una persona è nota con sufficiente anticipo, le contromisure interventistiche hanno una maggiore possibilità di successo.

Conoscere le variazioni genetiche e il loro effetto sull'organismo può quindi essere molto utile per individuare queste predisposizioni individuali, al fine di mantenere e migliorare il nostro stato di benessere. I test del DNA non sono in alcun modo diagnostici, cioè non sono in grado di stabilire se una persona svilupperà oppure no una certa patologia, ma forniscono indicazioni che consentono di adottare misure preventive mirate e personalizzate, in modo da aumentare la probabilità di mantenersi in salute o di recuperare una buona salute fisica ed estetica.

Lo sviluppo della genetica e dell'epigenetica sottolinea come gli stili di vita di ciascuno di noi, interagendo con i geni, possano modificare le risposte metaboliche e l'attivazione delle proteine funzionali.

Non solo. Ecco alcuni “punti chiave”:

- il ruolo sempre più importante che riveste la nutrigenomica, cioè la possibilità di condizionare l'espressione dei nostri geni;
- l'azione dei nutraceutici, non più paragonabili ai farmaci tradizionali, ma da vedersi sempre più come endomodulatori dell'espressione dei nostri geni, in grado di interagire con il DNA, direttamente o indirettamente, così da condizionarne l'espressione;
- la possibilità di personalizzare la scelta del farmaco, grazie ai sempre più rapidi sviluppi della farmacogenomica;
- l'importanza dell'epigenetica, fin dal comportamento e dall'alimentazione della madre nel corso della sua gravidanza.

È questo il filo conduttore del presente trattato di nutrigenomica ed epigenetica per il mondo accademico, curato e descritto con grande accuratezza scientifica, profondità e ricchezza di spunti di riflessione da Damiano Galimberti e Giovanni Battista Gidaro, con il prezioso supporto di Stefano Govoni, Vittorio Calabrese e Giovanni Malerba, insieme ad altri eminenti accademici e clinici, ognuno “testimonial” dei capitoli di personale conoscenza: dalla biologia alla clinica.

Personalmente ho incontrato il dottor Galimberti 8 anni fa e con lui ho condiviso l'importanza del fatto che il DNA possa essere influenzato da sostanze nutritive, stimoli esterni e stile di vita. Nessun destino nel DNA, ma solo condizioni che possono avvantaggiarci o svantaggiarci e che, una volta apprese, possono essere condizionate positivamente dalle nostre scelte. Una medicina

a tutti gli effetti personalizzata e personalizzabile. Pertanto ho introdotto questi concetti e questi esami nella mia pratica clinica: strumenti da affiancare a quelli tradizionali per una visione più olistica del paziente così come del soggetto sano, affinché quest'ultimo possa essere condotto a preservare questo stato il più a lungo possibile nel tempo.

Spero che questo libro, curato dal mio collega e amico insieme ad altri validi professionisti della salute, possa essere di ispirazione anche per altri medici in vista non solo di una medicina curativa, ma anche e in particolare preventiva.

Prof. Dr. Michael Papacharalambous, MD PhD
Chief Internal Medicine Department,
Metropolitan Hospital, Athens
General Secretary of ESAAM – European Society of Preventive,
Regenerative and Anti-Aging Medicine

Riferimenti bibliografici

Third Joint European Societies' task Force on cardiovascular Disease Prevention in Clinical Practice. European Guidelines for Cvd Prevention. *Eur J Cardiovasc Prev Rehabil.* 2003; 10(Suppl 1).

WHO. *Comprehensive cervical cancer control. A guide to essential practice.* Geneva, 2006.

WHO. *Gaining health. The European Strategy for the Prevention and Control of Noncommunicable Diseases.* Geneva, 2006.

WHO. *National cancer control guidelines: policies and managerial guidelines*, 2nd ed., Geneva, 2002.

WHO. *Preparing a health care workforce for the 21st century. The challenge of chronic conditions.* Geneva, 2005.

WHO (OMS). *Prevenire le malattie croniche: un investimento vitale.* Ginevra, 2005.

WHO Regional Office for Europe. *A strategy to prevent chronic disease in Europe. A focus on public health action. The CINDI vision.* Copenhagen, 2004.

WHO Regional Office for Europe. *The European health report 2005. Public health action for healthier children and populations.* Copenhagen, 2005.

Wilkinson R, Marmot M (eds.). *Social determinants of health: the solid facts*, 2nd ed., WHO Regional Office for Europe, Copenhagen, 2005.

Indice

Capitolo 1. Basi biologiche della nutrigenomica e dell'epigenetica

<i>Giovanni Battista Gidaro</i>	1
Il genoma della specie umana	2
DNA e geni	2
Cromatina e cromosomi	6
Conservazione e trasmissione del patrimonio genetico	8
Il dogma centrale della biologia	9
La regolazione epigenetica dell'espressione genica.....	10
La traduzione del codice genetico e la sintesi proteica.....	16
Gli aminoacidi	17
Aminoacidi e proteine: relazione struttura-funzione.....	17
Architettura e organizzazione delle proteine.....	21
Le sequenze segnale aminoacidiche determinano il destino delle proteine	26
Ruolo degli aminoacidi nell'interazione delle proteine con le altre molecole	31
Le varianti del DNA.....	33
Variazione e riparazione del DNA.....	33
Classificazione delle varianti genomiche... ..	36
Variazioni cromosomiche.....	37
Variazioni di sequenza nucleotidica.....	38
Effetti delle variazioni e malattia.....	40
Bibliografia.....	42

Capitolo 2. Nutrigenetica e nutrigenomica

<i>Giovanni Battista Gidaro</i>	47
Relazione tra polimorfismi e nutrizione	49
Trasportatori di membrana e canali ionici	50
Recettori di membrana e intracellulari.....	52
Sistemi di trasduzione.....	55
Enzimi e proteine effettrici.....	56
Proteine strutturali	60
Bibliografia.....	60

Capitolo 3. La dieta mediterranea e la caloric restriction

<i>Damiano Galimberti</i>	63
Dieta mediterranea.....	63
Caloric restriction.....	70
Apoptosi e autofagia.....	71
Bibliografia	73

Capitolo 4. Nutraceutici ed espressione genica

<i>Damiano Galimberti</i>	75
Nutraceutici ed epigenetica.....	76
Curcuma longa.....	76
Polygonum cuspidatum e resveratrolo	78
Licopene	81
Uncaria tomentosa	82
Acido ellagico	83
Acido α -lipoico.....	84
Indolo-3-carbinolo (I3C), diindolimetano (DIM) e sulforafano	86
Zingiber officinale	90
Silybum marianum	92
Capsaicina	93
Bacopa monnieri.....	95
Quercetina	96
Fisetina.....	97
Epigallocatechina-3-gallato	98
Rhodiola rosea	100
Panax ginseng.....	100
Withania somnifera	101
Ziziphus jujuba	101
Antociani	101
Dieta, nutraceutici e preparazioni galeniche	104
Bibliografia.....	107

Capitolo 5. I test del DNA in medicina predittiva e preventiva

<i>Damiano Galimberti</i>	115
Bibliografia.....	122

Capitolo 6. Genetica clinica in medicina preventiva

Damiano Galimberti, Giovanni Battista Gidaro..... **125**

Introduzione *Damiano Galimberti*..... 125
Bibliografia..... 126

Medicina potenziativa

Damiano Galimberti 127
Inflamming 127
Stress ossidativo..... 128
Metilazione 129
Glicazione..... 130
Meccanismi di DNA-repair..... 130
Bibliografia..... 132

Medicina del wellness

Damiano Galimberti 134
Bibliografia..... 137

Medicina estetica *Damiano Galimberti*..... 138
Bibliografia..... 143

Medicina sportiva

Giovanni Battista Gidaro..... 144
Endurance 146
Potenza 148
Recupero postesercizio..... 149
Suscettibilità agli infortuni 149
Bibliografia..... 150

Capitolo 7. Genetica clinica nelle intolleranze alimentari, nel peso corporeo, nel metabolismo e nei processi di detossificazione

Damiano Galimberti, Giovanni Battista Gidaro..... **153**

Introduzione *Damiano Galimberti*..... 153
Bibliografia..... 154

Peso corporeo e aspetti metabolici

Damiano Galimberti 155
Bibliografia..... 163

Intolleranze alimentari

Giovanni Battista Gidaro..... 165
Intolleranza al lattosio 166
Intolleranza e sensibilità al glutine..... 169
 Malattia celiaca 169
 Allergia al grano 172
 Sensibilità al glutine non celiaca (NCGS)..... 172
Intolleranza all'istamina (HIT)..... 173
Conclusioni 179
Bibliografia..... 179

Detossificazione *Giovanni Battista Gidaro*... 182
Sistema di detossificazione..... 183
 Reazioni di fase I..... 183
 Reazioni di fase II..... 185
Detossificazione e chemoprevenzione della carcinogenesi..... 188
Bibliografia..... 189

Capitolo 8. Genetica clinica in medicina funzionale

Damiano Galimberti, Federico Licastro, Elisa Porcellini, Giovanni Battista Gidaro..... **193**

Introduzione *Damiano Galimberti*..... 193

Oncologia *Damiano Galimberti*..... 194
Bibliografia..... 202

Neurologia

Federico Licastro, Elisa Porcellini..... 204
Fattori di rischio della demenza 204
Epigenomica della malattia di Alzheimer... 206
Nutrigenomica nella malattia di Alzheimer... 207
Prevenzione del decadimento cognitivo e della demenza 208
Bibliografia..... 209

Cardiologia

Damiano Galimberti..... 210
Malattia tromboembolica 211
Fibrillazione atriale..... 212
Ipertensione..... 212
Morte improvvisa –
Sindrome del QT lungo 212
Profilo lipidemico..... 214
Iperomocisteinemia 214
Ictus 215
Miscellanea 216
Conclusioni 216
Bibliografia..... 216

Ginecologia

Damiano Galimberti 218
Menopausa, terapia ormonale sostitutiva (TOS) e polimorfismi genetici 218
Poliabortività 220
Endometriosi 221
Infertilità 221
 Alterazioni ovariche 222
 Resistenza all'ormone luteinizzante..... 222
 Endometriosi e infertilità..... 223
Bibliografia..... 223

Apparato respiratorio

Damiano Galimberti 225
Genetica e broncopneumopatia cronica ostruttiva (BPCO)..... 225

Genetica e asma.....	226
Bibliografia.....	229

Apparato osteoarticolare

<i>Damiano Galimberti</i>	230
Osteoporosi.....	230
Artrosi	234
Conclusioni.....	234
Bibliografia.....	235

Dermatologia *Damiano Galimberti*

Alopecia androgenetica.....	236
Farmacogenomica	
e alopecia androgenetica	237
Area celsi e alopecia areata	237
Cheloidi e cicatrizzazione	238
Skin aging, eczema, dermatite atopica	
e idratazione cutanea	239
Bibliografia.....	239

Gusto e selezione degli alimenti

<i>Giovanni Battista Gidaro</i>	240
Il gusto	240
Il gusto dolce.....	240
Il gusto amaro	242
Bibliografia.....	242

Capitolo 9. Micobiota, microbiota e DNA

<i>Alessandro Gelli, Damiano Galimberti</i>	245
---	-----

Cosa sono il gut microbiota

e il gut microbioma? <i>Alessandro Gelli</i>	245
Bibliografia.....	248

Gut microbiota, microbioma, micobiota, micobioma e DNA *Alessandro Gelli*

Gut microbiota e gut microbioma:	
correlazioni con aging/stress,	
alimentazione e salute globale umana.....	252
Azioni del gut microbiota	254
Disbiosi ed eubiosi.....	255
Cervello, gut microbiota, microbioma,	
aging e stress	260
Cervello cranico-cervello addominale-	
gut microbiota (CCCAGM).....	261
Cervello cranico, sistema nervoso	
autonomo, gut microbiota, aging/stress,	
leaky gut/tight junction e gluten sensitivity:	
ancora sulla grande connessione	263
Tight junction e GALT.....	265
Il morbo di Crohn, una patologia	
autoimmune: correlazioni tra gut	
microbiota e micobiota	265

Sostanze naturali per favorire	
lo stato di eubiosi	267
Principali categorie	267
Il gut micobiota (funghi o miceti del gut)	270
Uso/abuso di antibiotici: resistenza	
e influenze sulle popolazioni del gut	
microbiota e sul relativo microbioma	274
Bibliografia.....	274

Test del DNA del microbiota

<i>Damiano Galimberti</i>	280
---------------------------------	-----

Capitolo 10. Medicina della longevità.

Frailty, la fragilità dell'anziano

<i>Davide Gentilini, Luisa Bernardinelli,</i>	
<i>Giovanni Battista Gidaro,</i>	
<i>Damiano Galimberti</i>	283
L'invecchiamento	283
Cause dell'invecchiamento.....	284
Aspetti epigenetici dell'invecchiamento	
e della longevità.....	286
Aspettativa di vita, invecchiamento	
in salute e longevità	290
La fragilità dell'anziano	291
Dieta e invecchiamento cerebrale	292
Bibliografia.....	293

Capitolo 11. Alimentazione durante la gravidanza e implicazioni epigenetiche

<i>Davide Gentilini, Luisa Bernardinelli,</i>	
<i>Giovanni Battista Gidaro</i>	295
Alimentazione e gravidanza	295
Epigenetica, ambiente e stress	297
Epigenetica e nutrizione	299
Epigenetica e sviluppo embrionale	300
Nutrizione in gravidanza	
e aspetti epigenetici.....	301
Effetti transgenerazionali.....	303
Bibliografia.....	303

Capitolo 12. Cronobiologia

e crononutrizione *Giovanni Battista Gidaro*

Bioritmi	305
Oscillatore centrale e oscillatori periferici	307
Geni orologio	309
Modulazione dell'orologio biologico	
e cronodistruzione.....	310
Dieta e ritmo circadiano.....	313
Frequenza dei pasti	313
Composizione dei pasti	314
Bibliografia.....	315

Capitolo 13. Proteomica e metabolomica

<i>Rosaria Ornella Bua, Vittorio Calabrese</i>	319
Proteomica	319
Modificazioni post-traduzionali.....	320
Metodi di analisi	323
Metabolomica	329
Metodi di analisi	331
Bibliografia	336

Capitolo 14. Esami di laboratorio

e test genetici <i>Damiano Galimberti</i>	341
Bibliografia	345

Capitolo 15 La prospettiva del professionista e del consumatore tra etica e norme

<i>Giovanni Battista Gidaro</i>	347
La prospettiva del professionista	347
La prospettiva del consumatore	349
Aspetti etici	350
Aspetti regolatori e legali	351
Bibliografia	352

Capitolo 16. Farmacogenomica

<i>Stefano Govoni</i>	353
Introduzione e definizioni	353
La variabilità di risposta ai farmaci: ruolo e importanza della farmacogenetica ...	355
Importanza della farmacogenomica nello sviluppo dei farmaci e negli studi clinici	356
Basi genetiche della risposta ai farmaci	357
Polimorfismi di enzimi responsabili del metabolismo dei farmaci	358
Polimorfismi degli enzimi metabolizzanti di fase I	359
Polimorfismi degli enzimi metabolizzanti di fase II	360
Polimorfismi dei trasportatori dei farmaci	361
Polimorfismi in bersagli primari di farmaci...	362
Polimorfismi dei microRNA	363
Coesistenza di più genomi.....	364
Reazioni di ipersensibilità e farmacogenetica	365
Polimorfismi combinati: un esempio.....	365
Mezzi informatici a disposizione	366
Conclusioni	366
Bibliografia	367

Capitolo 17. Tecniche genomiche, bioinformatica e statistica abbinata alla genetica

<i>Giovanni Malerba, Antonio Mori</i>	369
L'informazione genetica.....	369
Ereditarietà e fattori genetici associati a malattia	369
Un gene per più caratteri.....	370
Alla ricerca del gene coinvolto	371
Analisi di linkage.....	371
Analisi di associazione	372
Studi GWAS in dettaglio.....	374
Variabilità del genoma, sequenziamento e genotyping.....	378
Lo studio caso-controllo più in dettaglio.....	378
Next Generation Sequencing: dal biologico all'in silico	383
Sequenziamento di nuova generazione o massivo parallelo.....	383
Parametri fondamentali per la qualità di un esperimento NGS	384
Applicazioni del sequenziamento di nuova generazione	384
Studi di espressione genica: RNA-seq e microarray a confronto	386
Sequenziamento di nuova generazione e bioinformatica	386
Flusso di lavoro di un esperimento NGS di II generazione.....	387
NGS per usi clinici	396
Alcuni articoli consigliati.....	397
Bibliografia.....	397

Capitolo 18. Dalla nutrigenomica alla tavola

<i>Damiano Galimberti, Giorgio Damini</i>	401
Quinoa con cavolo nero, aglio, olio, peperoncino e maitake.....	402
Riso nasigoreng con curcuma, carote e maitake	403
Tartare di salmone selvaggio affumicato, robiola di latte crudo di mucca e capra, cialda di patate e aneto e profumo di limone	403
Pappa al pomodoro	404
Broccolo fiolaro croccante aglio, olio e peperoncino	404
Crostatine di farina integrale, lemon curd e lamponi	405
Coniglio al finocchietto	405

Carne cruda, passatina di topinambur e tartufo nero.....	406	di peperoni arrostiti	408
Faraona arrostita in padella con crema di aglio dolce, cous cous di broccoli romani, frutta secca (noci, mandorle) e polvere di lamponi disidratati	407	Sarde in saor con melagrana.....	409
Tortellini di carne in brodo di carote e zenzero	407	Conclusioni <i>Damiano Galimberti</i>	411
Puccia Damini di manzo con formaggio		Nutrigenomica ed epigenetica.....	411
Dobbiaco, speck, rucola e passatina		Apoptosi, autofagia e necrosi.....	413
		Il mitocondrio.....	414
		Indice analitico	419

CAPITOLO 1

Basi biologiche della nutrigenomica e dell'epigenetica

Giovanni Battista Gidaro

La ricerca scientifica ha permesso di sviluppare numerose conoscenze in campo nutrizionale e farmacologico, tali da consentire un approccio personalizzato ai bisogni del paziente. Il concetto di personalizzazione è abbastanza vasto e relativo a numerosi aspetti della nutrizione e della terapia medica. Relativamente alla nutrizione, la personalizzazione si riferisce classicamente alla valutazione dei bisogni energetici in funzione dell'attività lavorativa, fisica e sociale; alla disponibilità di tempo per preparare i pasti e al contesto lavorativo; alla composizione e alla costituzione corporea; ai gusti, alle preferenze e ai comportamenti alimentari; e, cosa più importante, alla valutazione dei bisogni nutrizionali in funzione dello stato fisiologico, del sesso, dei quantitativi di vitamine e minerali previsti dai LARN nelle varie fasi dello sviluppo umano e infine dello stato patologico accertato, nonché dei relativi esami ematochimici o urinari. Tutto questo permette un buon grado di personalizzazione e la selezione di alimenti ottimali per garantire il fabbisogno nutrizionale, soprattutto nell'ambito della dieta mediterranea, che elenca numerosi alimenti con effetto salutistico e preventivo. La moderna era della genomica e della postgenomica ha però rivelato che gli individui differiscono tra loro per milioni di varianti genomiche e ciò li rende diversi tanto nell'aspetto esteriore quanto nel metabolismo e nei processi fisiologici, così come nella risposta ai farmaci e alle sostanze tossiche. L'individuo in quanto tale è unico e diverso dalla media della popolazione. Gli enzimi del metabolismo possono essere più lenti o più veloci e la necessità di cofattori vitaminici o minerali può dunque cambiare da individuo a individuo. Alcuni enzimi non sono addirittura espressi a causa di delezioni, così come alcuni recettori, trasportatori oppure canali di membrana possono avere una funzionalità alterata; tutto ciò dovrebbe essere tenuto in considerazione, se noto, nella personalizzazione della nutrizione e della terapia medica. Sebbene siamo ancora all'inizio della costruzione e della comprensione del puzzle genomico alla base della complessità umana, le conoscenze già disponibili in nutrigenetica e nutrigenomica permettono di migliorare il grado di personalizzazione della dieta, rispettivamente sulla base dei polimorfismi e dell'espressione genica, così come quelle disponibili in farmacogenomica consentono di prevedere in anticipo la farmacocinetica, di individuare i soggetti metabolizzatori lenti o veloci e anche i pazienti responder e non-responder sulla base delle variazioni presenti in recettori, sistemi di trasduzione ed effettori.

Tutto questo sta incoraggiando il passaggio dal vecchio approccio *trial and error*, in cui il medico provava una terapia, sbagliava e correggeva, a un nuovo paradigma di *precision medicine*, basato appunto sulla variabilità genetica individuale e sulla capacità predittiva del dato

CAPITOLO 1 - Basi biologiche della nutrigenomica e dell'epigenetica

genomico. Questo nuovo approccio, che si fonda sulla genomica e, a breve, si baserà anche sull'epigenetica, permetterà sempre più di fare la differenza in termini nutrizionali e terapeutici a livello individuale.

In questo capitolo verranno presi in esame alcuni argomenti di base utili alla comprensione della genomica applicata alla nutrizione e alla clinica, quali il genoma e i meccanismi di controllo epigenetico della trascrizione, la struttura delle proteine e il ruolo degli aminoacidi, le variazioni genetiche e il loro effetto.

Il genoma della specie umana

Intorno agli anni 2000, il **progetto Genoma** rivelò per la prima volta la sequenza del DNA umano, grazie alla collaborazione di numerosi laboratori di ricerca, con un costo che si aggirò intorno ai 3 miliardi di dollari. Seguirono poi altri progetti di ricerca, come l'**HapMap** e il **progetto 1000 Genomi**, che hanno generato informazioni su varianti del DNA e sulla sequenze di migliaia di individui [1,2].

Da un punto di vista puramente strutturale, il genoma umano è costituito da **acido desossiribonucleico (DNA)** e da **proteine** di diversa natura necessarie al funzionamento del DNA. In termini funzionali, il genoma può invece essere considerato come l'**insieme delle informazioni** ereditarie necessarie alla costruzione, al mantenimento e al funzionamento dell'organismo.

Le informazioni genetiche sono racchiuse nelle molecole di **DNA nucleare** e solo in minima parte in quello **mitocondriale**, e sono costituite dalla sequenza delle quattro basi azotate che compongono il DNA: adenina, timina, citosina e guanina. Il DNA nucleare è complessato con proteine e organizzato in strutture chiamate **cromosomi**. Ogni cromosoma è costituito da una molecola di DNA. In totale ci sono **22 tipi di cromosomi autosomici**, presenti in duplice copia e detti **omologhi** perché contengono, salvo mutazioni, gli stessi tipi di geni; a questi si aggiunge una coppia di **cromosomi sessuali (XY)**, detti **eterologhi** in quanto possiedono geni diversi. In ogni coppia di cromosomi, una copia è di origine paterna e una di origine materna; per tale ragione il **corredo cromosomico** umano è di tipo **diploide (2n)**.

In base all'ultima versione 38.p7 del Genome Reference Consortium Human Build (GRCh38.p7, aggiornato al giugno 2016), il DNA aploide ha una lunghezza di circa **3 miliardi e mezzo di basi** ($3,5 \times 10^9$) [3]. Nel DNA sono presenti sequenze codificanti, chiamate **geni**, che saranno analizzate più avanti, spaziate da regioni intergeniche di dimensioni variabili. Solo l'1-3% del DNA ha funzione codificante e contiene l'informazione genetica per la sintesi delle proteine. Secondo la versione Gencode25, ci sono circa 19.950 geni codificanti per proteine. La maggior parte del DNA, invece, ha ancora funzioni sconosciute, ma è molto probabile che una buona parte svolga funzioni regolatorie [4].

DNA e geni

Il DNA nucleare è una macromolecola a **doppia elica destrorsa** formata da due filamenti appaiati costituiti da una sequenza di 2'-desossiribonucleotidi (Figura 1.1).

In ogni nucleotide del DNA è presente un'unità di 2'-desossiribosio (ribosio senza il gruppo 2'-OH) unito tramite il carbonio 1 a una base azotata e tramite il carbonio 5 al gruppo fosfato. In ciascun filamento i **desossiribonucleotidi** sono uniti mediante **legame fosfodiesterico 5'-3'** tra il fosfato 5' di un nucleotide e l'ossidrile 3' del desossiribosio appartenente al nucleotide adiacente. Ciascun filamento ha dunque un'estremità 5' con un fosfato libero e un'estremità 3' con un ossidrile libero. Nei nucleotidi del DNA sono presenti quattro basi azotate: **adenina**

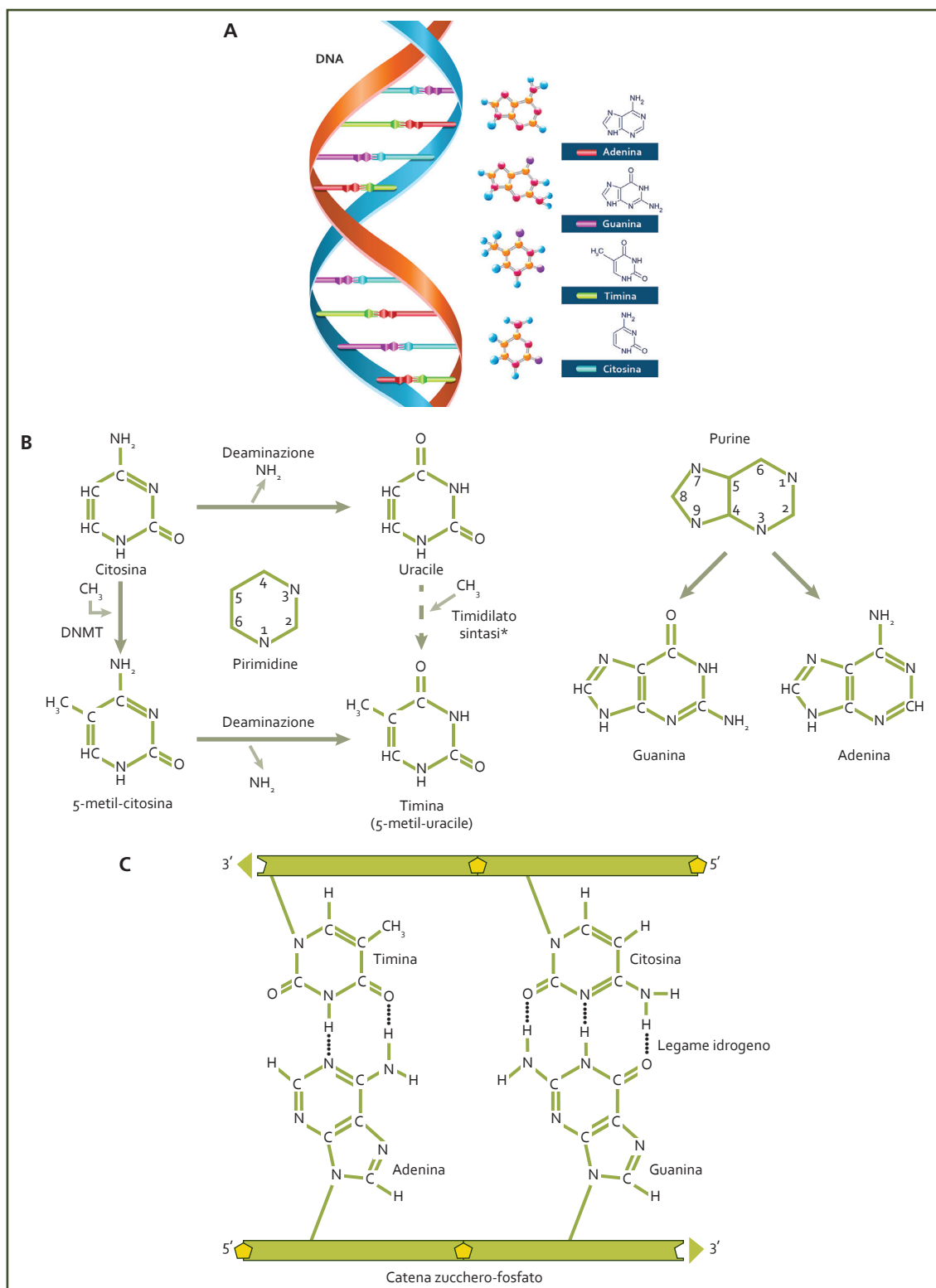


Figura 1.1 (A) La doppia elica del DNA; (B) le specie chimiche delle basi azotate; (C) l'appaiamento delle basi A=T e C=G. DNMT: DNA-metiltransferasi.

* La timidilato sintasi agisce sul deossitimidilato (dUMP) convertendolo in deossitimidilato (dTMP).

(A), **guanina (G)**, **citosina (C)** e **timina (T)**. Di queste basi, la citosina e la timina sono dette pirimidine per la presenza dell'anello pirimidinico, mentre l'adenina e la guanina, avendo un anello purinico (anello pirimidinico fuso con un anello imidazolico), sono chiamate purine. La sequenza delle basi azotate viene chiamata sequenza nucleotidica e la lunghezza di una specifica sequenza viene misurata in numero di "paia di basi" (bp), oppure in chilobasi (kb) per indicare le migliaia o in megabasi (Mb) per indicare i milioni di basi. Diversamente dal DNA, l'**RNA** (acido ribonucleico) è fatto di ribonucleotidi che presentano il ribosio al posto del deossiribosio e le medesime basi azotate del DNA, fatta eccezione per la timina (5-metil-uracile), che è sostituita da un'altra pirimidina chiamata **uracile (U)**.

Secondo le ricerche pionieristiche di Watson e Crick, i due filamenti del DNA presentano le basi azotate rivolte verso il centro dell'elica e si appaiano in modo tale da formare coppie complementari AT stabilizzate da due legami idrogeno e coppie CG legate invece da tre legami idrogeno. La quantità delle coppie CG (mediamente del 41% nel genoma umano) in una data regione incrementa la stabilità della doppia elica e la **temperatura di fusione** (temperatura di melting, T_m) necessaria per separare i due filamenti in quella regione. La doppia elica è strutturata in modo da contenere un nucleotide ogni 3,4 Å lungo l'asse verticale; i nucleotidi presentano tra loro un'angolazione di 36°, quindi la struttura si ripete ogni 10 nucleotidi, cioè a intervalli di 34 Å. I due filamenti complementari dell'elica hanno **polarità opposta**, cioè uno corre in direzione 5'→3' (forward strand) mentre l'altro in direzione 3'→5' (reverse strand); per questa ragione sono detti **filamenti antiparalleli**. Nell'elica si osservano un **solco maggiore** e un **solco minore**, che rappresentano siti di ingresso e di legame per fattori importanti per la replicazione, la trascrizione genica e la sua regolazione, e per la riparazione del DNA [5,6].

Lungo la sequenza del DNA si incontrano sequenze codificanti (geni) e sequenze non codificanti. Le sequenze non codificanti sono costituite da sequenze di **DNA ripetitivo** e si trovano negli spazi intergenici e, in parte, negli introni. A questo gruppo di DNA ripetitivo appartengono i **trasposomi**, elementi mobili del DNA molto abbondanti (circa il 45% del genoma) che possono migrare in altre regioni del genoma e che comprendono i retrotrasposomi e i trasposomi a DNA. I **retrotrasposomi** possono essere virali e non virali; questi ultimi sono abbondanti nell'uomo e sono rappresentati dai **LINE** (lunghi elementi nucleari interdispersi) e dai **SINE** (breve elementi nucleari interdispersi). I retrotrasposomi usano la trascrittasi inversa per convertire la copia di RNA a cDNA, che viene quindi inserito nel DNA; invece i **trasposomi a DNA** si spostano senza alcuna copia di sequenza. Questi elementi mobili sono da considerare come parassiti molecolari che si sono duplicati nel corso dell'evoluzione. La loro migrazione nel DNA può alterare geni ed essere causa di patologie. Solo una piccola percentuale del genoma (circa l'1,5-3%) è costituita dalle porzioni codificanti dei geni. Per **gene** si intende una sequenza di nucleotidi capace di codificare una proteina oppure un RNA funzionale. Secondo questa definizione sono geni anche quelle sequenze che trascrivono gli RNA coinvolti nella sintesi proteica, cioè gli RNA ribosomiali (rRNA) e gli RNA transfer (tRNA); inoltre, sono classificati come geni altri RNA non codificanti come i miRNA (**microRNA**) e i long ncRNA (**long non-coding RNA**), che secondo recenti evidenze svolgono funzione regolatoria. I geni per gli RNA e per le proteine istoniche sono normalmente ripetuti in tandem. I geni che codificano le proteine possono essere presenti in copia singola (**geni solitari**) oppure in diverse copie (**geni duplicati**), fino a formare anche **famiglie di geni**. Ci sono poi gli **pseudogeni**, cioè geni che, avendo accumulato molte mutazioni nel corso dell'evoluzione, non sono più funzionali, anche se si ipotizza un loro ruolo regolatorio dell'espressione genica. I geni si comportano

come unità fisiche trasmissibili alla progenie e sono quindi alla base dell'ereditarietà. Qualche decennio fa si pensava che ci fossero addirittura 100.000 geni capaci di trascrivere proteine e ciò avrebbe potuto spiegare, almeno in parte, la complessità dell'essere umano, mentre oggi si sa che ci sono solo **19.950 geni codificanti per proteine** (Tabella 1.1), un valore molto simile a quello del topo [7,8]. È noto, inoltre che i geni sembrano essere distribuiti casualmente sui due filamenti e che, a livello del gene, solo uno dei due filamenti codifica per la proteina, mentre il filamento complementare potrebbe esprimere un RNA con funzione regolatoria (ad esempio miRNA). Oggi si ritiene che il numero delle sequenze regolatorie sia in correlazione con la complessità dell'organismo.

Da un punto di vista strutturale, nel gene eucariotico si distinguono diverse regioni, tra cui la sequenza codificante, quella non codificante e quella regolatoria. La **sequenza codificante** può essere costituita da sequenze di diversa lunghezza chiamate **esoni**, separate da sequenze non codificanti dette **introni**. Gli esoni contengono molto spesso l'informazione genetica per fabbricare un dominio della proteina e possono essere combinati e uniti in modo alternativo per formare diverse isoforme proteiche mediante il meccanismo di **splicing alternativo** (dall'inglese *splice*, unire), un processo che implica la rimozione degli introni. In prossimità della giunzione introne-esone sono presenti sequenze segnale che indicano i **siti di splicing**. In particolare, all'estremità intronica 5' si trova molto spesso una sequenza **GU**, mentre all'estremità 3' una sequenza **AG** preceduta da una sequenza ricca di pirimidine. A monte del segnale 3' si trova, inoltre, un'adenina (**A**) chiamata punto di ramificazione e necessaria per il processo di splicing [9]. I siti di splicing sono riconosciuti dallo **spliceosoma**, un complesso macromolecolare costituito da cinque piccole ribonucleoproteine (snRNP) [10]. La stragrande maggioranza (>95%) dei geni multiesonici produce isoforme proteiche anche in modo tessuto-specifico [11]. Analizzando il rapporto tra trascritti codificanti proteine e il numero di geni (vedi Tabella 1.1), si deduce che ci sono in media quattro trascritti per gene. In pratica, il meccanismo dello splicing alternativo permette a un gene di esprimere più isoforme proteiche. A monte del primo esone e a valle dell'ultimo esone ci sono rispettivamente le regioni **5'-UTR** e **3'-UTR**, che vengono trascritte, ma che non vengono tradotte (UTR, Untranslated Region). La **regione regolatoria** del gene si trova principalmente a monte del primo esone e prende il nome di **promotore**; tuttavia, ci possono essere altre sequenze regolatorie, quali gli **elementi prossimali al promotore**, che si trovano a -200 bp, oppure gli **intensificatori** (enhancer), costituiti da 50-1.500 nucleotidi, che possono essere localizzati a monte (-1Mb) o nell'introne o a valle dell'esone 3'-terminale [12,13]. L'attività degli enhancer può essere bloccata da altri elementi di DNA interposti tra enhancer e promotore, chiamati **isolanti** (**insulator**), i quali sono sequenze di legame per re-

Tabella 1.1 Statistiche generali sui geni umani in base alla versione 25 del progetto GENCODE*

Numero totale di geni	58.037
Geni codificanti per proteine	19.950
Geni che trascrivono RNA lunghi non codificanti (long non-coding RNA genes)	15.767
Geni che trascrivono RNA piccoli non codificanti (small non-coding RNA genes)	7.258
Pseudogeni	14.650
Trascritti codificanti per proteine	80.087

*www.gencodegenes.org/stats/current.html.

pressori della trascrizione come il CTFC. Il promotore contiene specifiche **sequenze segnale (elementi)** che permettono la formazione del complesso di inizio della trascrizione e il reclutamento dell'RNA polimerasi II necessaria per la sintesi dell'RNA messaggero (mRNA) e di altri RNA nucleari [14]. Tre sequenze segnale sono molto frequenti (Figura 1.2) e hanno diversa localizzazione rispetto al sito di inizio della trascrizione: sono la TATA box (–25 bp), le isole CpG (–40 e –110 bp), presenti in una regione chiamata CG box (con sequenza ripetuta tipo GGGCGG), e la CAAT box (–110 bp), che sono riconosciute rispettivamente dai fattori di trascrizione TFIID, SP1 e CTF [15-17]. La TATA box è presente in geni trascritti in modo tessuto-specifico o in alcune fasi cellulari, mentre la CG box è presente nei geni espressi in modo costitutivo (housekeeping). Ci possono essere, inoltre, altre sequenze che funzionano da **repressori (silencer)** della trascrizione.

Il **DNA mitocondriale (mtDNA)**, diversamente da quello nucleare, è una molecola circolare composta da un filamento interno (L) e uno esterno (H), senza introni e senza grandi spazi intergenici. Sebbene il mitocondrio necessiti di molte proteine per funzionare, il DNA mitocondriale codifica per poche proteine essenziali per la funzionalità del mitocondrio, tra cui le subunità del complesso NADH-CoQ reduttasi, citocromo c ossidasi, citocromo b e alcune subunità dell'ATPasi; inoltre codifica per gli rRNA per la sintesi dei ribosomi mitocondriali e per i tRNA necessari per la sintesi delle proteine mitocondriali. La maggior parte delle proteine mitocondriali è, invece, di importazione citosolica e codificata dal DNA nucleare. Diversi dati indicano che il DNA mitocondriale sia di origine batterica e che, nel corso dell'evoluzione, la maggior parte dei geni mitocondriali sia migrata nel nucleo.

Cromatina e cromosomi

La molecola del DNA è impacchettata con proteine basiche (istoni) a formare una struttura chiamata **cromatina**, la cui unità fondamentale è il **nucleosoma**, un complesso costituito da un ottamero di istoni (H2A, H2B, H3 e H4, ciascuno in due unità) intorno al quale è avvolto un rocchetto di 146 paia di basi fissato dall'istone H1. Le proteine istoniche hanno code amino-terminali flessibili che protrudono fuori dall'ottamero, ricche in lisina e arginina e dotate di carica positiva che, interagendo con i gruppi fosfato del DNA prossimale e di nucleosomi distanti, facilita la condensazione della cromatina. Quando la cromatina è attiva, cioè in fase di trascrizione, essa prende il

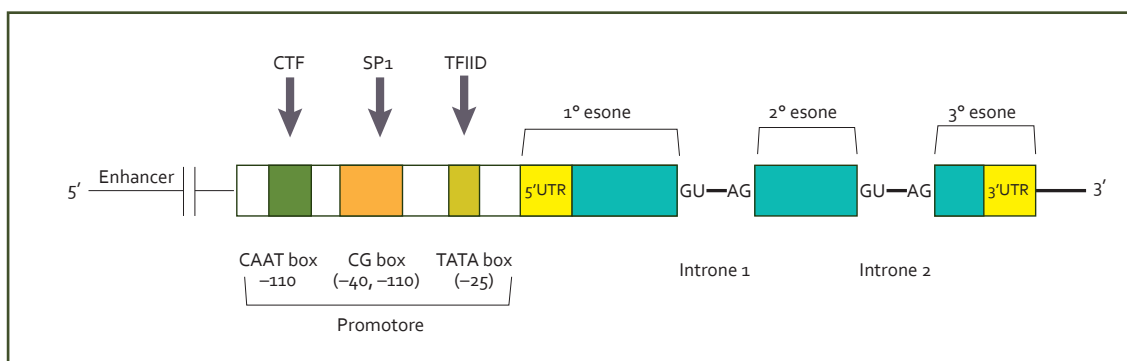


Figura 1.2 Esempio di struttura di gene con tre esoni e due introni, siti di splicing e potenziale localizzazione degli elementi del promotore. CTF: CAAT binding-Transcription Factor; SP1: Specificity Protein 1 (Transcription Factor); TFIID: Transcription Factor IID. Per la spiegazione vedi il testo.

nome di **euromatina** e si presenta al microscopio elettronico come una fibra del diametro di 10 nm con l'aspetto di una collana di perle costituite da nucleosomi. Al contrario, quando è inattiva, la cromatina viene chiamata **eterocromatina** e assume una conformazione a solenoide in cui i nucleosomi sono avvolti a spirale, originando una struttura superimpacchettata del diametro di 30 nm che risulta inaccessibile ai fattori di trascrizione [18,19]. Nella fase del ciclo cellulare chiamata metafase, successiva alla duplicazione del DNA (fase S), il DNA si trova superimpacchettato e forma corpuscoli visibili al microscopio ottico (dopo opportuna colorazione con coloranti basici), che furono chiamati **cromosomi** da H.W.G. von Waldeyer-Harts, nel 1889, secondo la terminologia di origine greca *chroma* (colore) e *soma* (corpo). Come già accennato, ci sono 23 coppie di cromosomi, di cui 22 autosomici e una coppia di tipo sessuale. **Ciascun cromosoma è composto da una sola molecola di DNA**, quindi ci sono in totale 46 molecole di DNA (44 autosomiche + 2 sessuali). Considerando che gli autosomi sono presenti in duplice copia (una materna e una paterna), ci sono al massimo 24 molecole di DNA diverse (22 autosomi + XY). Oltre alle proteine istoniche, nel cromosoma si trovano le **proteine impalcatura** e altri fattori proteici necessari al funzionamento del DNA, come la DNA polimerasi, la RNA polimerasi, le proteine di riparazione del DNA e i fattori di trascrizione. Analizzando la struttura del cromosoma si osserva una costrizione primaria detta **centromero**, più o meno centrale, che divide il cromosoma in due porzioni, di cui una è chiamata **braccio corto (p)** e l'altra **braccio lungo (q)** (Figura 1.3). In base alla posizione del centromero, il cromosoma è detto metacentrico (posizione centrale), submetacentrico (mediana), acrocentrico (periferica), telocentrico (terminale). Nella porzione terminale di alcuni cromosomi è presente una costrizione secondaria chiamata **organizzatore nucleolare**, che contiene i geni codificanti l'RNA ribosomiale. Le estremità del cromosoma sono chiamate **telomeri** e sono costituite da sequenze ripetute (TTAGGG) per circa 10 kb e da eterocromatina [20]. Durante il ciclo cellulare alcune regioni cromosomiche di eterocromatina, con i geni che devono essere trascritti, vengono convertite a euromatina e prendono il nome di **eterocromatina facoltativa**, mentre altre regioni come il centromero, i telomeri e la cromatina del secondo cromosoma sessuale X rimangono inattive e sono chiamate **eterocromatina costitutiva**. L'attività della cromatina cambia al variare delle fasi del ciclo cellulare e della specificità tessutale.

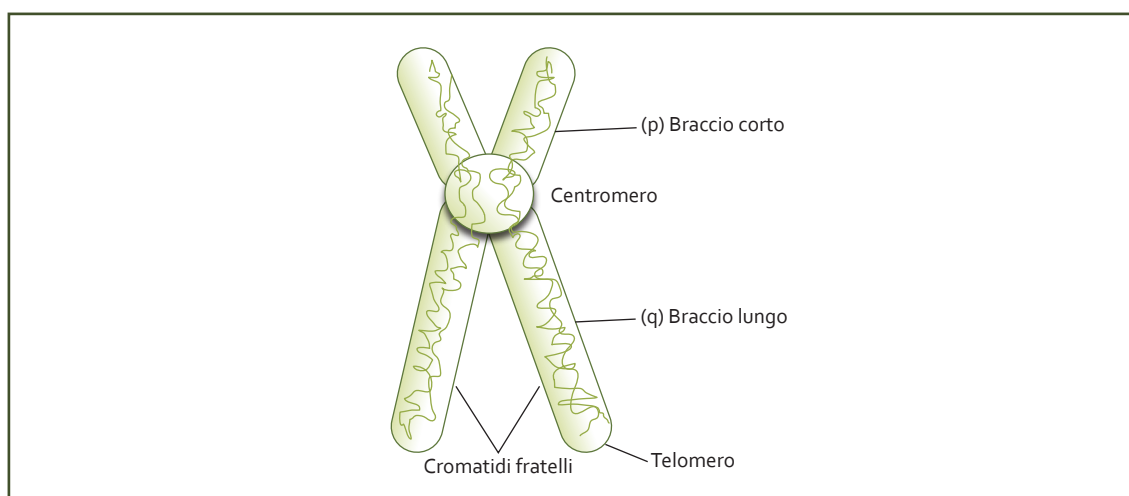


Figura 1.3 Il cromosoma metafaseico.

CAPITOLO 1 - Basi biologiche della nutrigenomica e dell'epigenetica

In metafase, dopo la duplicazione del DNA, il cromosoma è costituito da **due cromatidi fratelli** saldati dal centromero. L'impiego di specifici coloranti (ad esempio reattivo di Giemsa, idrocloreuro di quinacrina) nei cromosomi metafasici permette di ottenere il **caratteristico bandeggio** in cui si distinguono regioni e sottoregioni numerate a partire dal centromero. Le **bande G** sono quelle più scure (colorate con Giemsa), mentre le **bande R** sono quelle più chiare. Sembra che le bande G abbiano un contenuto di GC inferiore alla media, mentre le bande R abbiano un contenuto superiore. Per indicare la posizione di un gene sul cromosoma si usa un codice che esplicita il numero del cromosoma, il braccio del cromosoma, la regione, la sottoregione e la banda cromosomica. Ad esempio, il gene *ACE* ha localizzazione citogenetica 17 q23.3, che significa che esso si trova sul cromosoma 17, braccio q, regione 23, sottoregione 3. Oggi sono disponibili metodiche più avanzate (array CGH e analisi FISH con coloranti fluorescenti), caratterizzate da una risoluzione molecolare 100-1.000 volte superiore alla metodica convenzionale, che permettono di rilevare alterazioni della macrostruttura dei cromosomi e di analizzare il **cariotipo**, cioè il patrimonio cromosomico in termini di numero, dimensione e forma dei cromosomi [21].

Conservazione e trasmissione del patrimonio genetico

L'informazione genetica viene conservata mediante la **replicazione del DNA**, catalizzata principalmente dalla **DNA polimerasi** e da altri complessi enzimatici. La replicazione delle estremità telomeriche è invece operata dalla **telomerasi**, un complesso proteina-RNA dotato di stampo a RNA per la replicazione delle sequenze ripetitive del telomero e con attività analoga a quella della trascrittasi inversa. La duplicazione del DNA è **semiconservativa** e questo significa che ciascuna molecola figlia contiene un filamento nuovo e un filamento della molecola madre originale [22-25]. La conservazione del DNA, inoltre, è supportata da diversi sistemi di riparazione, tra cui un sistema che corregge gli errori di appaiamento (**mismatch repair**) assicurando la fedeltà della duplicazione.

La trasmissione del DNA alle cellule figlie somatiche avviene mediante il processo della mitosi, mentre la trasmissione alla progenie avviene attraverso la meiosi, con la formazione dei gameti. Semplificando, la **mitosi è la fase M del ciclo cellulare** (G₀, G₁, S, G₂, M) in cui la cellula madre si divide in due cellule figlie trasferendo loro il patrimonio genetico in modo conservativo. Durante l'interfase (G₁+S+G₂) la cellula diploide (2n) sintetizza proteine, raddoppia il patrimonio genetico (4n) in fase S (S = sintesi del DNA) e si prepara alla mitosi, processo in cui si distinguono quattro macrofasi chiamate profase, metafase, anafase e telofase.

In **metafase** la cellula si presenta con i cromosomi disposti sulla piastra equatoriale, ciascuno costituito dai due cromatidi fratelli. La mitosi si conclude con la segregazione bilanciata dei cromatidi fratelli (anafase) e la formazione di due cellule figlie diploidi (2n) identiche mediante il processo di citodieresi (telofase). Anche il DNA mitocondriale, presente nella matrice dei mitocondri, viene diviso tra le due cellule figlie in modo equo. Nelle cellule somatiche la telomerasi è tuttavia spenta e la replicazione dei telomeri non avviene completamente, producendo così un accorciamento a ogni ciclo cellulare. È stato proposto che la lunghezza dei telomeri funzioni come un timer (**orologio mitotico**), rendendo possibile solo un certo numero di divisioni cellulari: quando i telomeri sono consumati il ciclo cellulare si arresta e viene avviato il programma di senescenza e di morte cellulare [26,27].

La **meiosi**, invece, si verifica nei tessuti germinali preposti alla gametogenesi e, partendo da una cellula diploide (2n), genera quattro cellule gametiche aploidi (n) mediante due divisioni cel-

lulari successive, di cui la prima, detta "riduzionale" (meiosi I), riduce effettivamente il corredo genetico da $2n$ ($4n$ dopo la duplicazione) ad aploide (n); la seconda è chiamata "equazionale" (meiosi II) e produce quattro gameti aploidi. Nella meiosi I si verifica il **crossing-over**, processo in cui i cromosomi omologhi si appaiano e si scambiano materiale genetico prima di separarsi nelle due cellule figlie aploidi. Come risultato del crossing-over, un gamete prodotto da un individuo contiene informazioni genetiche di provenienza sia paterna sia materna. Questo evento è di fondamentale importanza per la riproduzione sessuale, la sopravvivenza e l'evoluzione della specie, in quanto è una delle principali fonti della variabilità genetica e biologica. In alcuni casi, lo scambio di materiale genetico tra gli omologhi non è bilanciato e si verifica un crossing-over ineguale che produce duplicazione genica in un gamete e delezione in un altro. Il gamete con duplicazione può dare dunque origine a uno zigote con triplicazione di un particolare gene o regione genica ed essere alla base di alterazioni patologiche (ad esempio trisomie). Con la riproduzione sessuale si verifica la fusione dei due corredi cromosomici aploidi dei gameti e la generazione di uno **zigote diploide**.

Si ritiene che il DNA mitocondriale, invece, sia trasmesso in modo uniparentale dalla cellula uovo e sia pertanto di origine materna. Per ogni gene del DNA nucleare viene ereditata una versione paterna e una materna. Le forme alternative dello stesso gene sono chiamate **alleli**. I due alleli posseduti da un individuo costituiscono il **genotipo**, mentre il carattere che si manifesta come risultato del genotipo prende il nome di **fenotipo**. Se l'organismo possiede due alleli identici in un locus genico, allora viene detto **omozigote** per quel locus genico; se gli alleli sono diversi, allora si parla di organismo **eterozigote**. I due alleli possono avere effetto qualitativo e quantitativo diverso. Se l'effetto di un allele prevale su quello dell'altro, questo allele è detto **dominante** e convenzionalmente si indica con la lettera maiuscola (ad esempio A); l'altro allele viene invece detto **recessivo** perché si manifesta solo in omozigosi e si indica con la lettera minuscola (a).

Il dogma centrale della biologia

Intorno agli anni Cinquanta, Watson e Crick, oltre a svelare la struttura del DNA, scoprirono anche che l'informazione genetica contenuta nel DNA viene trascritta in RNA messaggero nel nucleo e poi tradotta in catena polipeptidica nel citosol. Con riferimento ai dogmi religiosi, che non sono soggetti a discussione, Crick definì questo flusso monodirezionale dal gene alla proteina come il **dogma centrale della biologia**, anche se successivamente, come da lui sospettato, si scoprì l'esistenza di un percorso a ritroso dall'RNA al DNA catalizzato dalla trascrittasi inversa, come nel caso dei retrovirus (ad esempio HIV) oppure dei retrotrasposomi, sequenze di DNA che possono replicarsi dall'RNA da loro stessi trascritto e integrarsi in altre regioni del DNA umano [28]. Si è scoperto, inoltre, che ogni singolo gene può portare l'informazione per fabbricare diverse **isoforme proteiche**, cioè versioni diverse della stessa proteina, ma con proprietà differenti. Nel nucleo, il flusso dell'informazione genetica passa attraverso due processi fondamentali, che sono la trascrizione e la maturazione dell'RNA messaggero. Durante la **trascrizione**, l'RNA polimerasi legge il filamento stampo di DNA $3' \leftarrow 5'$ producendo un trascritto primario di RNA che cresce in direzione $5' \rightarrow 3'$ [29]. Il filamento di DNA $5' \rightarrow 3'$ che non viene trascritto ha la stessa sequenza del trascritto, eccezion fatta per la timina che è sostituita dall'uracile, e per questo motivo viene definito in inglese **sense strand**; al contrario, il filamento stampo viene chiamato **antisense strand**. Dopo la trascrizione, il trascritto primario subisce un processo di maturazione che lo trasforma in RNA messaggero (mRNA) [30]: l'estremità $5'$ viene

CAPITOLO 1 - Basi biologiche della nutrigenomica e dell'epigenetica

metilata con formazione di un cappuccio protettivo chiamato “**cap**”, mentre l'estremità 3' viene tagliata, in un sito specifico, da un'endonucleasi e poliadenilata dalla poli-A-polimerasi, che attacca una **coda di circa 200-250 adenine** con funzione stabilizzante [31]. Successivamente, l'RNA va incontro allo **splicing alternativo**, con unione degli esoni e rimozione degli introni. Questo processo è estremamente importante per la generazione di isoforme proteiche con diverse proprietà funzionali (Figura 1.4).

Dopo lo splicing, il messaggero è dunque maturo e pronto ad abbandonare il nucleo per essere tradotto in proteine. Inoltre, deve essere ricordato che ci possono essere altri processi, tra cui quello di “**editing**” del messaggero, un processo che può apportare modifiche alla sequenza nucleotidica originariamente codificata. Un esempio interessante è quello della proteina ApoB100, il cui messaggero viene editato con formazione di un segnale di terminazione (stop codon) e produzione della proteina più corta ApoB48. Analogamente al DNA, anche l'RNA è complessato con proteine, tra cui i fattori di splicing e le proteine di esportazione nucleare che trasferiscono il messaggero dal nucleo al citosol.

La regolazione epigenetica dell'espressione genica

L'espressione genica è finemente regolata da molteplici fattori che modulano l'attività dei **fattori di trascrizione**, proteine che riconoscono sequenze segnale presenti nel DNA (**elementi**) e avviano la formazione del complesso di trascrizione mediante il reclutamento di **co-attivatori**

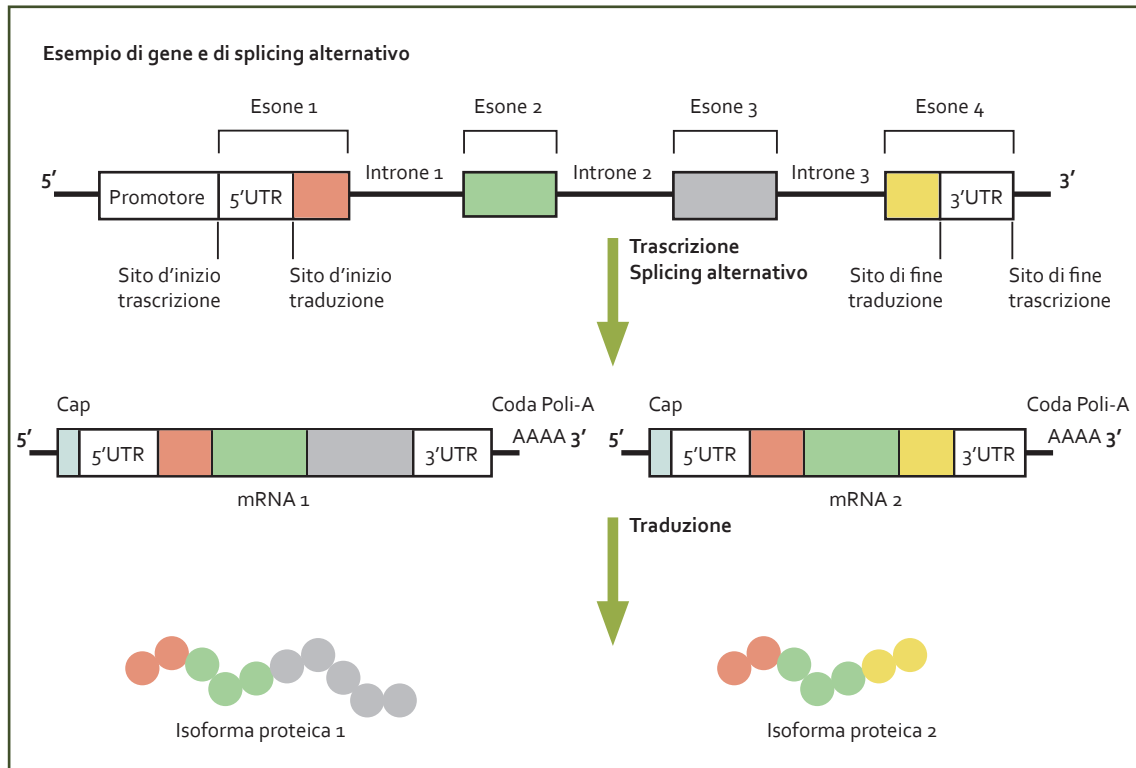


Figura 1.4 Esempio di struttura di un gene con quattro esoni e di splicing alternativo con produzione di due messaggeri diversi.

o di **co-repressori**, proteine con attività enzimatica che, senza legare il DNA, influenzano l'attività del promotore attivando o sopprimendo l'espressione. Alcuni geni sono vitali per tutte le cellule, hanno un'espressione costitutiva e sono chiamati in inglese **housekeeping genes** (ad esempio gene per DNA polimerasi oppure del metabolismo intermedio); altri geni sono invece espressi solo in alcune cellule e sono dunque detti **tessuto-specifici**.

Alcune osservazioni indicano che, oltre al codice genetico, sono necessarie altre informazioni per coordinare e regolare la trascrizione. Com'è possibile, ad esempio, che a partire da un solo oocita fertilizzato (zigote) possano derivare centinaia di cellule diverse che presentano espressione tessuto-specifica di alcuni geni? Per fare un esempio più concreto, come mai l'emoglobina viene espressa solo nei globuli rossi e la mioglobina solo nei muscoli? Il fatto che la medesima informazione genetica origini cellule diverse lascia pensare che ci debbano essere informazioni aggiuntive all'informazione genetica capaci di dirigere la trascrizione selettiva dei fattori di trascrizione, cui segue una trascrizione tessuto-specifica dei geni. Un altro esempio in cui la medesima informazione genetica non spiega completamente il fenotipo è quello dei gemelli identici monozigoti, i quali possono presentare un certo grado di discordanza per fenotipo e storia clinica. Un'altra osservazione indiziale è che le cellule differenziate mantengono un'identità cellulare nel ciclo cellulare e si dividono mitoticamente in due cellule figlie che ricordano e mantengono l'identità cellulare della cellula madre. In termini molecolari, questo significa che lo stato di attività e di condensazione della cromatina viene trasmesso alla cellula figlia in modo da tenere accesi quei geni caratteristici della linea cellulare e spenti quelli non necessari alla medesima linea. Un esempio di silenziamento genico mitoticamente ereditabile è l'**inattivazione del cromosoma X** nelle cellule somatiche femminili di mammifero, con formazione del **corpo di Barr**. Com'è noto, le cellule femminili hanno due cromosomi X, uno di origine paterna e l'altro di origine materna, mentre i maschi hanno un solo cromosoma X. Nella fase embrionale di gastrulazione, si assiste nelle cellule femminili all'inattivazione casuale di uno dei due cromosomi X al fine di compensare la dose genica, relativa ai geni posti sul cromosoma X, che altrimenti sarebbe doppia (e deleteria) rispetto a quella dei maschi (XY). Il fatto interessante è che in questa fase alcune cellule inattivano in modo casuale il cromosoma X materno e altre quello paterno, ma da questo momento in poi le cellule discendenti mantengono sempre inattivato lo stesso cromosoma che era stato inattivato nella cellula madre, dando vita a due cloni cellulari distinti e quindi a un mosaico cellulare.

L'insieme delle informazioni, in aggiunta a quelle del codice genetico, necessarie alla regolazione della trascrizione viene indicato complessivamente con il termine di **epigenetica** (dal greco *epi*, sopra, e *gennetikòs*, relativo all'eredità familiare). Il concetto di epigenetica è abbastanza recente e comincia a comparire nel 1942, quando Conrad Waddington iniziò a parlare di epigenotipo e di epigenesi, cioè di come il genotipo origini il fenotipo durante lo sviluppo [32]. Negli anni Novanta Robin Holliday introdusse il concetto di controllo spazio-temporale dell'espressione genica durante lo sviluppo di organismi complessi [33]. Nel 1996 Arthur Riggs definì il termine epigenetica come lo studio dei cambiamenti di espressione genica mitoticamente ereditabili che si verificano senza cambi nella sequenza del DNA [34]. La definizione oggi più accettata considera l'epigenetica come l'insieme di informazioni (mitoticamente e/o meioticamente ereditabili) in aggiunta a quelle genetiche che devono essere presenti per controllare l'espressione genica e consentire lo sviluppo e la differenziazione di un organismo complesso. Le parole "mitoticamente e/o meioticamente ereditabili" sono state poste tra parentesi perché alcune modifiche epigenetiche sono rapide e reversibili e non sono necessariamente ereditabili.

CAPITOLO 1 - Basi biologiche della nutrigenomica e dell'epigenetica

li. Immaginando il codice genetico come una successione di lettere (basi) senza punteggiatura, si potrebbe pensare all'epigenetica come all'insieme dei formati che permettono la lettura del codice. Senza punteggiatura sarebbe abbastanza difficile leggere la seguente frase: "Le modificazioni epigenetiche permettono di capire quando il gene inizia come sono organizzati i cromosomi e quali geni esprimere e quali silenziare". Tuttavia, con l'aggiunta della punteggiatura, degli spazi e dei paragrafi, diventa di facile lettura: "Le modificazioni epigenetiche permettono di capire:

- quando il gene inizia;
- come sono organizzati i cromosomi e
- quali geni esprimere e quali silenziare."

In termini molecolari, i principali meccanismi di regolazione epigenetica della trascrizione sono i seguenti: modificazione post-traduzionale delle code istoniche N-terminali, metilazione del DNA, codifica di varianti istoniche, rimodellamento della cromatina, attività di RNA non codificanti.

Come già accennato, le modifiche delle catene N-terminali degli istoni che protrudono fuori dal nucleosoma sono di fondamentale importanza nella regolazione epigenetica della trascrizione. Le principali modifiche post-traduzionali degli istoni sono elencate nella Tabella 1.2. Queste reazioni possono avvenire su tutti gli istoni, ma più frequentemente su H3 e H4, e la loro combinazione origina un codice chiamato **codice istonico**.

Tra queste reazioni, le più importanti e meglio investigate sono l'acetilazione e la metilazione, la funzione delle quali è stata solo in parte chiarita da varie ricerche, tra cui quelle del progetto ENCODE [35]. L'**acetilazione** dei residui di lisina degli istoni H2A, H2B, H3 e H4 maschera le cariche positive delle catene laterali aminoacidiche che, non potendo più legare i fosfati del DNA, permettono alla cromatina di assumere una conformazione chiamata **eucromatina**, meno condensata, accessibile alle proteine e attiva [36]. L'acetilazione libera i siti di legame per i fattori di trascrizione presenti sui promotori ed è associata con l'attività dei promotori e degli enhancer; i siti acetilati sono riconosciuti da proteine non istoniche con domini specifici per i gruppi acetile (bromodomain). Al contrario, la **deacetilazione** degli istoni impacchetta la cromatina formando l'**eterocromatina**, che risulta inattiva [37]. La reazione di acetilazione è mediata dalle istone acetiltransferasi (HAT), mentre la deacetilazione dalle istone deacetilasi (HDAC). A tale proposito, si ricorda che alcune istone deacetilasi, le sirtuine, sono attivate da

Tabella 1.2 Principali modificazioni epigenetiche delle code istoniche N-terminali

Modificazione	Residuo	Funzione regolativa
Metilazione (Me)	Lisina (K) e arginina (R)	Trascrizione, riparazione
Acetilazione (Ac)	Lisina (K)	Trascrizione, riparazione, replicazione
Ubiquitinazione (Ub)	Lisina (K)	Trascrizione
Sumoilazione (Su)	Lisina (K)	Trascrizione
ADP-ribosilazione	Glutamato (E)	Trascrizione
Fosforilazione (Ph)	Serina (S) e treonina (T)	Trascrizione, riparazione, condensazione
Citrullinazione (Cit)	R convertita a citrullina	Trascrizione

diverse molecole, tra cui il transresveratrolo, che potrebbero trovare un ruolo nel silenziamento genico. L'acetilazione è reversibile, è soggetta anche a regolazione circadiana e non presenta necessariamente ereditabilità mitotica. Quando interessa alcuni residui di lisina (H3K9ac e H3K27ac), l'acetilazione attiva la trascrizione in corrispondenza di promotori ed enhancer ed è un marker di cromatina attiva [35]. La **metilazione degli istoni**, invece, non altera la carica delle proteine e avviene principalmente su residui di lisina (K) e arginina (R); la trimetilazione, tuttavia, introduce nella catena laterale della lisina una carica positiva sull'atomo di azoto, che diventa quaternario (quattro legami). I residui aminoacidici possono essere mono-, di- o trimetilati e sono siti di ancoraggio per altre proteine non istoniche che possiedono domini di legame specifici per i gruppi metilici (chromodomain). Le reazioni di metilazione sono reversibili e sono operate da enzimi specifici come la lisina metiltransferasi (KMT), mentre la rimozione dei metili è invece catalizzata da lisina demetilasi specifiche (KDM). Il significato funzionale della metilazione cambia in base all'istone, al residuo aminoacidico metilato e al numero di unità metiliche aggiunte. Ad esempio, la monometilazione di H3K27me1 e di H3K9me1 è stata associata ad attivazione genica, mentre la trimetilazione degli stessi correla con la repressione genica [38]. I siti **H3K9me3/me2** sono associati alla formazione di eterocromatina costitutiva e al silenziamento di retrovirus, mentre la metilazione su **H3K27me3** sembrerebbe essere associata alla repressione di geni e all'ipermetilazione delle isole CpG all'interno del promotore, quindi potrebbe essere correlata alla formazione di eterocromatina facoltativa [35,39]. Diversamente da quanto sopra esposto, la trimetilazione **H3K4me3** indica il punto di inizio della trascrizione e recluta l'apparato di splicing; la monometilazione di H3K4me1 avviene all'estremità 5' del promotore e analogamente facilita la trascrizione promuovendo il reclutamento dell'RNA polimerasi II e sembra ostacolare il legame della DNA-metiltransferasi [35]. Nelle cellule embrionali pluripotenti, alcuni geni hanno un dominio bivalente con il sito attivatore H3K4me3 e quello repressore H3K27me3 per permettere l'espressione o la soppressione del gene nel corso della maturazione cellulare e in presenza di stimoli adeguati. Durante la differenziazione cellulare i geni che restano espressi perdono H3K27me3, mentre quelli che diventano repressi perdono H3K4me3 [40]. Dati sperimentali indicano che, comunque, il codice istonico è più complesso di quanto sino a ora descritto ed è costituito da altre modifiche a carico di altri residui di lisina e arginina, serina o treonina. La letteratura indica che ci potrebbero essere anche reazioni di propionilazione e butirrilazione di residui di lisina in cui gli acidi propionico e butirrico prodotti dalla flora batterica potrebbero svolgere un ruolo regolatorio della trascrizione, almeno nei colonociti.

La **metilazione del DNA** si verifica sui residui di citosina presenti in specifiche sequenze dinucleotidiche chiamate **CpG**, dove la lettera "p" indica il legame fosfodiesterico presente tra i due nucleotidi dello stesso filamento. La reazione di metilazione viene operata dall'enzima DNA-metiltransferasi, che utilizza S-adenosilmetionina come fonte di metile e produce 5-metil-citosina (5-metil-C), che conserva la capacità di appaiarsi con la guanina. La metilazione avviene simmetricamente sui due filamenti, grazie al fatto che il dinucleotide CpG si appaia con il GpC presente sull'altro filamento. Ci sono diversi tipi di enzimi **DNA-metiltransferasi (DNMT)**: DNMT3a e DNMT3b catalizzano la **de novo metilazione** su entrambi i filamenti; DNMT3L facilita la de novo metilazione aumentando l'affinità per il DNA, ma risulta sprovvista di attività enzimatica; DNMT1 interviene invece nel **mantenimento della metilazione**, particolarmente dopo la mitosi, quando, per effetto della duplicazione semiconservativa, il DNA si presenta emimetilato.

DNMT1 riconosce il dinucleotide ^mCpG sul filamento madre e metila il dinucleotide ($\text{GpC} \rightarrow \text{GpC}^m$) sul filamento figlio, garantendo così la conservazione e l'ereditabilità dei siti metilati nella linea cellulare. I dinucleotidi CpG sono sparsi per tutto il DNA e sono normalmente metilati nelle regioni intergeniche e in corrispondenza delle sequenze ripetitive costituite da DNA satellite, caratteristico delle regioni telomeriche e centromeriche, e da sequenze trasposomiche (ad esempio SINE, LINE). In queste regioni le sequenze ripetute CpG prendono il nome di “**non-isole CpG**” (non-CpG islands), per distinguerle dalle isole CpG presenti in prossimità del promotore genico. L'ipermetilazione di queste regioni (circa il 40-60% delle non-CpG islands) ha lo scopo di stabilizzare il genoma e mantenere la sua integrità, poiché previene inserzioni, delezioni, inversioni, traslocazioni reciproche, ma anche la proliferazione dei trasposomi e le ricombinazioni illegittime. In termini evolutivi questi siti di citosina metilati tendono, però, a deaminare spontaneamente e a mutare in timina, la quale può essere trasmessa alla progenie quando tale transizione avviene nei gameti. In prossimità o all'interno dei promotori ci sono sequenze ripetute CpG (da 500 a 2.000 circa) chiamate **isole CpG** (CpG islands) che, nei geni costitutivi (housekeeping genes) e in quelli tessuto-specifici, sono ipometilate, mentre nei geni non necessari alla linea cellulare sono ipermetilate. Una stima recente prevede che circa il 40-70% dei geni abbia CpG islands in prossimità del promotore e che il loro grado di metilazione si attesti intorno al 10-15%. La metilazione delle isole CpG nel promotore ha l'effetto di **sopprimere la trascrizione genica** [41]. Come regola generale, geni con promotori ipometilati sono attivi rispetto a geni ipermetilati. L'effetto che può avere il grado di metilazione del promotore dipende dalla funzione del gene nella cellula, per cui ad esempio l'ipometilazione di un gene attivatore del ciclo cellulare potrebbe sortire lo stesso effetto dell'ipermetilazione di un gene soppressore del ciclo cellulare.

Deve essere comunque precisato che la sola metilazione non è sufficiente a spegnere l'espressione genica e che sono necessari altri processi epigenetici, quali la deacetilazione e la metilazione degli istoni. Sono stati ipotizzati due meccanismi di **silenziamento genico** mediati dalla metilazione del DNA: da un lato, i nucleotidi metilati potrebbero impedire il legame dei fattori di trascrizione; dall'altro, faciliterebbero il legame con altre proteine dotate di un dominio di legame per il CpG metilato (Methyl Binding Protein Domain, MBD) tra cui MBD1, MeCP1 e MeCP2 (MeCpG Binding Protein), a loro volta in grado di reclutare altre proteine tra cui la istone deacetilasi e la istone metiltransferasi, che avviano la condensazione della cromatina. Infatti, la H3K9 metiltransferasi (KMT), che è un'istone metiltransferasi specifica per la lisina 9, trimetila H3K9 che lega la proteina HP1 (Heterocromatin Protein 1), una proteina non istonica in grado di legare la H3K9 metiltransferasi. Quest'ultima può dunque metilare i residui di lisina nel nucleosoma adiacente e propagare l'eterocromatina fino ad alcuni siti chiamati **isolanti (insulators)**, molto ricchi di residui acetilati e di H3K4 metilati che, come già accennato, sono indicatori di trascrizione genica. La metilazione può infatti anche favorire la trascrizione, quando riguarda gli elementi isolanti, interposti tra enhancer e promotore; in questo caso la loro metilazione ostacola il legame di un repressore della trascrizione (CTFC) e facilita l'azione attivatrice dell'enhancer sul promotore.

Come sarà spiegato nel Capitolo 11, la metilazione del DNA cambia durante alcune fasi dello sviluppo ed è inoltre alla base del meccanismo di imprinting epigenetico.

La posizione dei nucleosomi sul DNA e l'organizzazione della cromatina sono quindi di fondamentale importanza per il controllo genico. Il continuo **rimodellamento della cromatina** è sotto il controllo di complessi proteici (ad esempio SWI/SNF) che regolano l'assemblaggio o il

disassemblaggio dei nucleosomi e l'impacchettamento della cromatina. Il promotore e il DNA in trascrizione si presentano liberi da nucleosomi, a eccezione del sito di inizio che invece sembra richiedere la presenza di nucleosomi. Un altro meccanismo di controllo genico consiste nell'espressione di diverse **varianti istoniche** nelle diverse fasi del ciclo cellulare. La variante H3.3, ad esempio, è espressa in interfase e ha la funzione di mantenere stabile la struttura della cromatina durante la trascrizione. La variante H2A.Z, invece, è associata a instabilità nucleosomica ed è necessaria per l'accesso di fattori proteici al DNA [42].

Di recente si è scoperto che buona parte del DNA non codificante viene trascritta in RNA con funzione regolatoria. In particolare i **miRNA**, che sono non-coding RNA con dimensione di 17-25 nucleotidi, avrebbero la capacità di legarsi a sequenze complementari sui messaggeri di altri geni, regolandone la funzionalità, la stabilità e l'emivita. Inoltre, i miRNA potrebbero anche legarsi a sequenze complementari presenti nel promotore, competendo per siti di legame di fattori di trascrizione. Considerando i diversi meccanismi, un miRNA potrebbe regolare l'espressione genica di parecchi altri geni, anche centinaia. Almeno la metà dei miRNA viene codificata da geni con promotori contenenti le isole CpG e ciò suggerisce che la trascrizione dei miRNA possa essere regolata epigeneticamente [43].

L'attività dei fattori di trascrizione viene anche modulata da segnali extracellulari e/o intracellulari ed è sotto il controllo di **cascate intracellulari** capaci di influenzare l'espressione genica principalmente mediante l'attività chinasi di fosforilazione [14]. Un esempio di cascata intracellulare è quella della **via dell'adenilato ciclasi/AMPciclico/PKA** che, in risposta a stimoli di diversa natura (ad esempio adrenalina, glucagone, ACTH), trasdotti all'interno della cellula dalle G-proteine, determina la fosforilazione del fattore nucleare CREB (CRE-Binding protein), il quale, dimerizzando, lega sul DNA l'elemento di risposta al cAMP (CRE) e, reclutando la p300 e altri fattori di trascrizione, favorisce l'acetilazione e induce l'avvio della trascrizione. La trascrizione può essere modulata da numerose vie, tra le quali alcuni esempi sono rappresentati dalla **via dei fosfatidilinositoli**, che attiva la **PKC** e aumenta il calcio intracellulare, e da quella delle **chinasi calcio-calmodulina-dipendenti**, che interpretano le variazioni intracellulari del calcio come secondo messaggero; dalla via dei recettori accoppiati a tirosina chinasi citosoliche come la **JAK-STAT**, che trasmette il segnale delle citochine dalla membrana cellulare al DNA nucleare; dalla via delle **MAP-chinasi (Ras/Raf/MEK/MAP)** attivata dai recettori con attività tirosinchinasi intrinseca (RTK) in risposta a insulina e fattori di crescita (EGF, PDGF, NGF, M-CSFR, VEGFR ecc.), che porta a fosforilazione di fattori di trascrizione nucleari. È bene ricordare che alcuni fattori di trascrizione sono **recettori citosolici/nucleari** per diversi ligandi con funzione vitaminica e ormonale, tra cui il recettore per l'acido retinoico (RAR), il recettore per la vitamina D (VDR), per il progesterone (PR), per gli estrogeni (ER), per i glucocorticoidi (GR) e per gli ormoni tiroidei (TR). Tutti questi recettori, sebbene differenti, hanno quattro siti di legame simili, cioè quello per il ligando, per il DNA, per la proteina inibitoria e per il dominio di attivazione della trascrizione. Il legame del ligando libera la proteina inibitoria e induce cambi conformazionali che permettono il legame con coattivatori proteici che determinano l'avvio della trascrizione. I recettori attivati si legano al DNA come omodimeri oppure come eterodimeri, dirigono l'iperacetilazione degli istoni nei nucleosomi vicini e inducono una **risposta primaria** nei primi 30 minuti, caratterizzata dalla trascrizione di un piccolo set di geni, i cui prodotti proteici governano la **risposta secondaria** tardiva. È interessante notare che ci sono anche meccanismi epigenetici controllati da recettori nucleari. Numerose evidenze indicano che i glucocorticoidi, legandosi al recettore nucleare GR, possono controllare numerosi geni (>100) legandosi a particolari ele-

menti di risposta sul DNA (GRE) presenti in promotori che contengono isole CpG appartenenti all'**orologio epigenetico**, un gruppo di 353 dinucleotidi CpG che è associato all'età biologica e a numerose patologie correlate con l'aging e lo stress, tra cui le patologie cardiovascolari, la sindrome metabolica, il cancro e le patologie psichiatriche e neurodegenerative [44]. I corticosteroidi inducono l'ipometilazione della maggior parte di questi geni, probabilmente mediante escissione della base mediata da Gadd45 [45] oppure ossidazione operata da TET [46]; inoltre provocano l'ipermetilazione di una piccola parte (circa il 10%) di tali geni [47].

La traduzione del codice genetico e la sintesi proteica

Il codice genetico è costituito da **triplette** di nucleotidi, chiamate **codoni**, lette sul messaggero in successione e senza punteggiatura [48]. Cambiando il punto di inizio di un nucleotide si possono ottenere tre **moduli di lettura** completamente diversi, di cui però ne viene letto solo uno [49]. Ricordando che nell'mRNA ci sono quattro nucleotidi (A, U, C, G), la loro combinazione a gruppi di tre produce 64 triplette (4^3). Di queste triplette ben 61 codificano per aminoacidi e, visto che la maggior parte degli aminoacidi è codificata da più di un codone, si parla di **codice genetico degenerato** [50,51]. Ci sono ben sei codoni per arginina, serina e leucina, quattro codoni per alanina, glicina, prolina, treonina e valina, tre per isoleucina, due per tutti gli altri, tranne il triptofano e la metionina, che sono codificati da una sola tripletta. Il codone per la metionina è anche il **segnale di inizio** della traduzione (AUG), pertanto tutte le proteine nascono con una metionina amino-terminale che viene, in genere, successivamente rimossa. Ci sono infine **tre codoni di stop** (UAA, UGA, UAG), detti anche **segnali di arresto** in quanto indicano la terminazione della sintesi proteica (Figura 1.5).

La **traduzione** (da mRNA a proteina) richiede la presenza di ribosomi, mRNA, tRNA, aminoacidi e numerosi fattori proteici. I **ribosomi** sono complessi di rRNA e proteine e sono il sito di lettura e traduzione del messaggero. Il **ribosomi eucariotici** sono costituiti da due subunità, quella grande (60S), costituita da 50 proteine e da rRNA (28S, 5.8S e 5S), e quella piccola (40S), costituita da 33 proteine e un rRNA (18S) [52]. I ribosomi possono essere liberi nel citosol oppure adesi al reticolo endoplasmatico rugoso (RER), un complesso di vescicole cellulari citoplasmatiche che al microscopio elettronico presenta un aspetto rugoso per la presenza superficiale di ribosomi disposti a rosetta sul messaggero a formare complessi noti come poliribosomi. In realtà, la traduzione inizia sui ribosomi citosolici liberi e, se nella catena nascente compare il segnale di smistamento al RER, il complesso messaggero-poliribosomi migra sul reticolo rugoso. La decodifica delle triplette viene operata dall'RNA transfer (tRNA), che possiede un'ansa con una sequenza complementare al codone, chiamata **anticodone**, e trasporta all'estremità 3'-OH uno specifico aminoacido, formando un complesso detto **aminoacil-tRNA**. La prima base dell'anticodone può stabilire anche appaiamenti non standard con il codone e questo fenomeno, chiamato **vacillamento della prima base dell'anticodone**, permette al codice di essere decifrato da un numero di tRNA inferiore a 61 [53]. Nella traduzione si distinguono tre fasi dette di inizio, allungamento e terminazione, controllate da specifici fattori proteici. Nella **fase di inizio** la subunità piccola riconosce il CAP del messaggero e scorre fino al sito di inizio; il tRNA porta la metionina (il primo aminoacido di ogni proteina) nel sito aminoacilico (A) della subunità maggiore. Nella **fase di allungamento** il metionil-tRNA si sposta dal sito aminoacilico a quello peptidilico (P); quindi entra un altro aminoacil-tRNA nel sito A, mentre l'aminoacido presente nel sito P viene trasferito, mediante peptidil transferasi, nel sito A e legato, con legame peptidico, al gruppo $-NH_2$ dell'aminoacido presente nel sito A, formando un peptidil-tRNA.

		2ª posizione									
		U		C		A		G			
1ª posizione	U	UUU	Fenilalanina	UCU	Serina	UAU	Tirosina	UGU	Cisteina	U C A G	
		UUC	Fenilalanina	UCC	Serina	UAC	Tirosina	UGC	Cisteina		
		UUA	Leucina	UCA	Serina	UAA	Stop	UGA	Stop		
		UUG	Leucina	UCG	Serina	UAG	Stop	UGG	Triptofano		
	C	CUU	Leucina	CCU	Prolina	CAU	Istidina	CGU	Arginina	U C A G	
		CUC	Leucina	CCC	Prolina	CAC	Istidina	CGC	Arginina		
		CUA	Leucina	CCA	Prolina	CAA	Glutamina	CGA	Arginina		
		CUG	Leucina	CCG	Prolina	CAG	Glutamina	CGG	Arginina		
	A	AUU	Isoleucina	ACU	Treonina	AAU	Asparagina	AGU	Serina	U C A G	
		AUC	Isoleucina	ACC	Treonina	AAC	Asparagina	AGC	Serina		
		AUA	Isoleucina	ACA	Treonina	AAA	Lisina	AGA	Arginina		
		AUG	Metionina/Start	ACG	Treonina	AAG	Lisina	AGG	Arginina		
	G	GUU	Valina	GCU	Alanina	GAU	Ac. aspartico	GGU	Glicina	U C A G	
		GUC	Valina	GCC	Alanina	GAC	Ac. aspartico	GGC	Glicina		
		GUA	Valina	GCA	Alanina	GAA	Ac. glutammico	GGA	Glicina		
		GUG	Valina	GCG	Alanina	GAG	Ac. glutammico	GGG	Glicina		

Aminoacido apolare

Aminoacido polare

Figura 1.5 Il codice genetico del DNA nucleare.

Segue la fase di traslocazione, in cui il tRNA scarico presente nel sito P si sposta nel sito di exit (E) per lasciare il ribosoma, il peptidil-tRNA si sposta dal sito A al sito P e il ribosoma avanza di un codone sul mRNA in direzione 3'.

Questi cicli, che richiedono energia (GTP), si ripetono finché viene incontrato il **segnale di terminazione**, il quale recluta i fattori di terminazione che promuovono il taglio del peptidil-tRNA e la separazione delle unità ribosomiche. Le proteine nascono con l'estremità N-terminale e crescono in direzione C-terminale [54,55].

Gli aminoacidi

Aminoacidi e proteine: relazione struttura-funzione

Le proteine sono polimeri costituiti da una lunga catena di 20 tipi diversi di L-aminoacidi legati tra loro con un particolare legame covalente detto legame peptidico. La formula generale degli aminoacidi comprende un carbonio α centrale cui sono legati con legame covalente il gruppo

CAPITOLO 1 - Basi biologiche della nutrigenomica e dell'epigenetica

carbossilico (COOH), il gruppo aminico (NH_2) e la catena laterale (R). A pH neutro il gruppo carbossilico è dissociato (COO^-), mentre il gruppo aminico è protonato (NH_3^+) (Figura 1.6). Il nome degli aminoacidi può essere indicato per esteso oppure mediante abbreviazione a tre lettere oppure a una lettera. Il più semplice tra gli aminoacidi è la glicina, la cui catena laterale è uguale a un atomo di idrogeno ($\text{R}=\text{H}$) e viene indicata con l'abbreviazione Gly oppure G.

I 20 aminoacidi che compongono le proteine differiscono per le caratteristiche chimico-fisiche della catena laterale (R), che determinano l'ingombro sterico (cioè tridimensionale), la carica positiva o negativa, l'affinità per l'acqua (idrofilia) o per i lipidi (lipofilia) e la capacità di reagire chimicamente con altri aminoacidi o altre molecole. La conoscenza delle caratteristiche degli aminoacidi è fondamentale per comprendere l'effetto delle mutazioni sulla struttura e sulla funzionalità delle proteine (Figura 1.7).

In base alla carica elettrica presente sulla catena laterale, si distinguono aminoacidi acidi (con carica negativa), basici (con carica positiva) e neutri (con carica nulla) (Tabella 1.3).

Gli **aminoacidi acidi** sono due, l'**acido aspartico** (Asp o D) e l'**acido glutammico** (Glu o E), che a pH neutro presentano il gruppo carbossilico della catena laterale in forma dissociata, quindi con carica negativa (COO^-). Questo gruppo negativo può stabilire all'interno della proteina legami ionici, intracatena o intercatena, con gruppi positivi oppure partecipare al binding di un ligando nel caso di recettori o di un substrato nel caso di enzimi.

Gli **aminoacidi basici** sono tre e comprendono l'arginina, la lisina e l'istidina, che presentano, a pH neutro, una carica positiva su un azoto della catena laterale. L'**istidina** (His o H) contiene nella catena laterale un anello imidazolico (con due atomi di azoto) molto reattivo che può presentare una carica su uno dei due azoti in dipendenza del pH circostante. L'istidina è fondamentale all'interno della mioglobina e dell'emoglobina per il legame dell'ossigeno al gruppo eme. La **lisina** (Lys o K) ha un gruppo aminico-terminale protonato (NH_3^+) mentre l'**arginina** (Arg o R) ha una catena laterale che termina con un gruppo guanidinico protonato a pH neutro, con carica positiva stabilizzata dall'effetto di risonanza. Lisina (K) e arginina (R) sono molto abbondanti nelle proteine istoniche all'interno della cromatina, particolarmente nelle code istoniche N-terminali dove sono siti target di diverse reazioni di regolazione epigenetica della trascrizione: metilazione (K e R), acetilazione (K), ubiquitinazione e sumoilazione (K), citrullinazione (R). Con le loro cariche positive contribuiscono a impacchettare il DNA stabilendo legami con le cariche negative dei gruppi fosfato. La lisina è inoltre fondamentale per la stabilità di alcune proteine extracellulari in cui partecipa alla formazione di legami crociati (cross-link) tra catene proteiche: nel collagene, ad esempio, ci sono legami crociati intracatena di tipo aldolico formati dalla lisilossidasi (rame-dipendente) tra due residui di lisina, e legami crociati

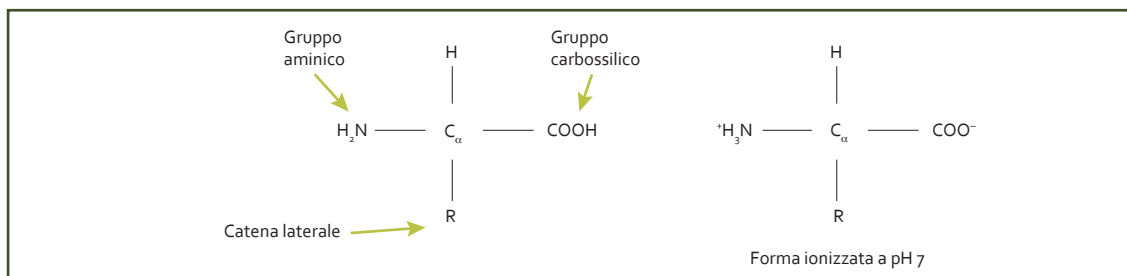


Figura 1.6 Formula generale di un aminoacido.

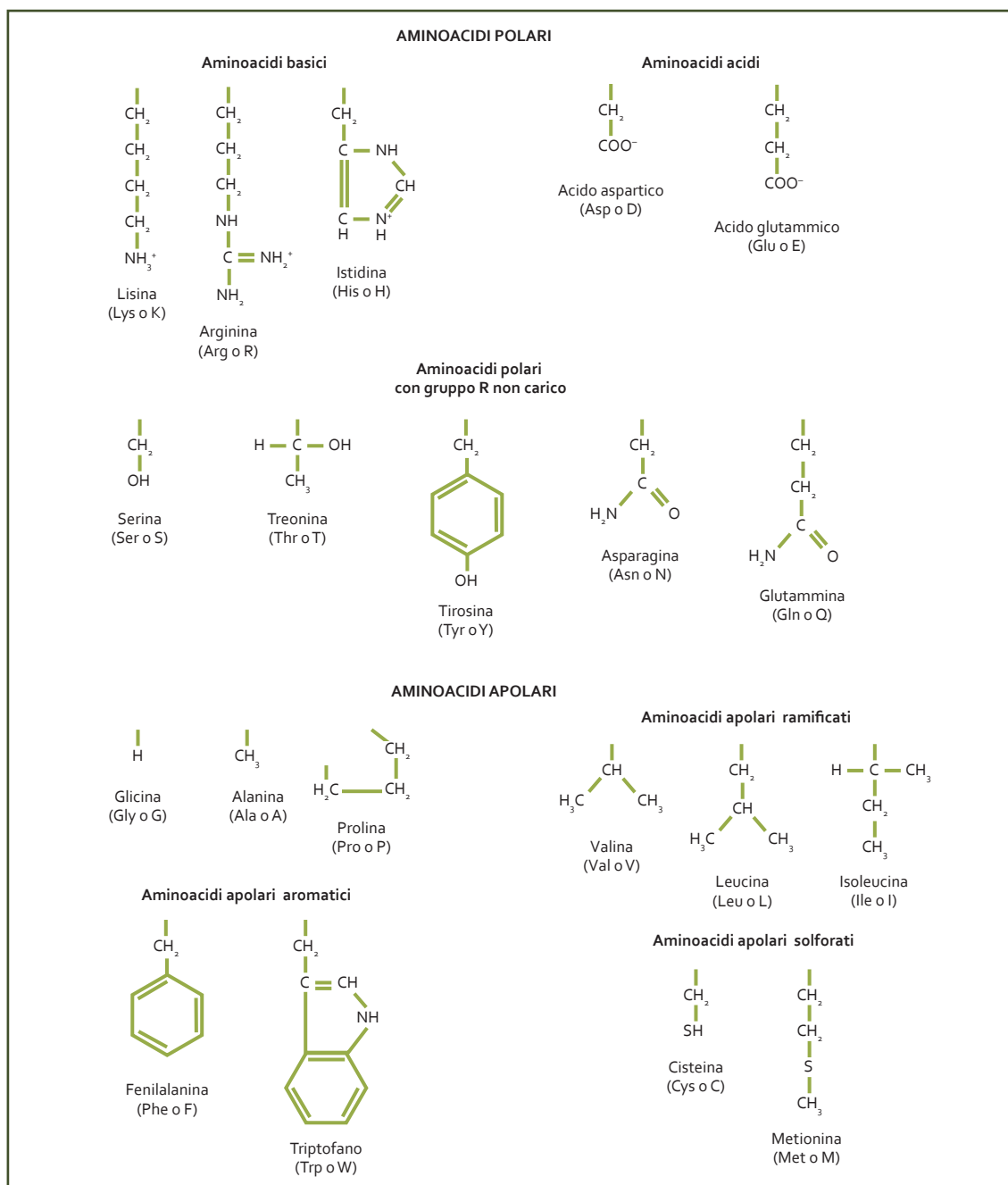


Figura 1.7 Catene laterali degli aminoacidi.

intercatena di tipo idrossipiridinico tra quattro residui di lisina. Similmente, nell'elastina sono presenti legami crociati tra due residui di lisina (lisinorleucina) o tra quattro residui di lisina (de-smosina). Questi legami crociati permettono al collagene di sviluppare resistenza meccanica e all'elastina di avere proprietà elastiche che facilitano il recupero della forma originale dopo il proprio allungamento [56,57].