

Le scienze “omiche”, ovvero: dal dogma centrale della biologia alle scienze omiche. La biologia molecolare al servizio del clinico

Martina Collotta*

Candiolo Cancer Institute - FPO, IRCCS, Candiolo, Torino

Un percorso a puntate per capire dove sta andando la ricerca in campo biomedico

Inizia con questo primo numero dell'anno la collaborazione con la dottoressa Martina Collotta che ci introdurrà e guiderà alla conoscenza delle scienze omiche dalla loro origine fino alle sempre più frequenti e importanti applicazioni nella pratica clinica. Questo ci consentirà di leggere con maggiore consapevolezza la letteratura scientifica orientandoci meglio con termini come: *next generation sequencing*, trascrittoma, metabolomica. Non solo per comprendere i progressi della ricerca di base, ma anche i suoi risvolti applicativi “al servizio del clinico”: dallo studio dei fattori di rischio alla risposta ai farmaci, passando attraverso l'impiego dei marcatori molecolari in diagnostica e nella valutazione prognostica delle patologie. Le target therapy e le terapie del futuro, sempre più personalizzate, affondano le loro radici nello studio del genoma del singolo paziente e nella caratterizzazione molecolare della sua malattia. Nelle prossime puntate: polimorfismi del DNA e patologia; epigenomica: un “nuovo codice” oltre la sequenza di basi; il *genome editing* e le sue applicazioni ... e molto altro.

Giuseppe Ventriglia

Da Watson e Crick alle scienze omiche

Dagli anni cinquanta a oggi, la biologia molecolare ha fatto enormi progressi. Watson e Crick, con la loro scoperta del DNA quale molecola portatrice dell'informazione genetica, hanno dato il via a una serie di progressi scientifici che hanno ora l'estensione *omica*.

Il termine *omics* (scienze omiche) si riferisce a tutte le discipline che hanno per oggetto di studio la caratterizzazione e la quantificazione di *pool* di molecole biologiche, al fine di delineare struttura, funzioni e dinamiche di un organismo.

Genomica, trascrittomica, proteomica: sono questi, forse, i termini che più frequentemente troviamo non solo nella letteratura,

ma anche sul Web e che rispecchiano il cosiddetto *dogma centrale della biologia molecolare* (Fig. 1): dal DNA alle proteine attraverso l'RNA.

E in effetti genomica, trascrittomica e proteomica studiano rispettivamente l'insieme dell'informazione genetica (DNA), degli RNA messaggeri (mRNA) e delle proteine della cellula.

Lo sguardo del ricercatore si è oggi allargato dalla singola molecola alla totalità delle molecole, ampliando le conoscenze scientifiche in modo impressionante, con non poche implicazioni per la clinica.

La biologia molecolare, infatti, proprio attraverso le scienze omiche, è destinata sempre più a dare alla clinica l'impronta che cerca: *l'impronta del paziente*, quella data dalla *sua* peculiare biologia e dalla biologia,

unica, della *sua* malattia, a cui la clinica è chiamata a rispondere con una terapia sempre più personalizzata. Le scienze omiche, infatti, permettono di passare da un approccio “generalistico” a quell'approccio “individualizzato” che la medicina ha sempre ricercato e che risponde quindi in modo specifico alle esigenze del paziente e della *sua* patologia.

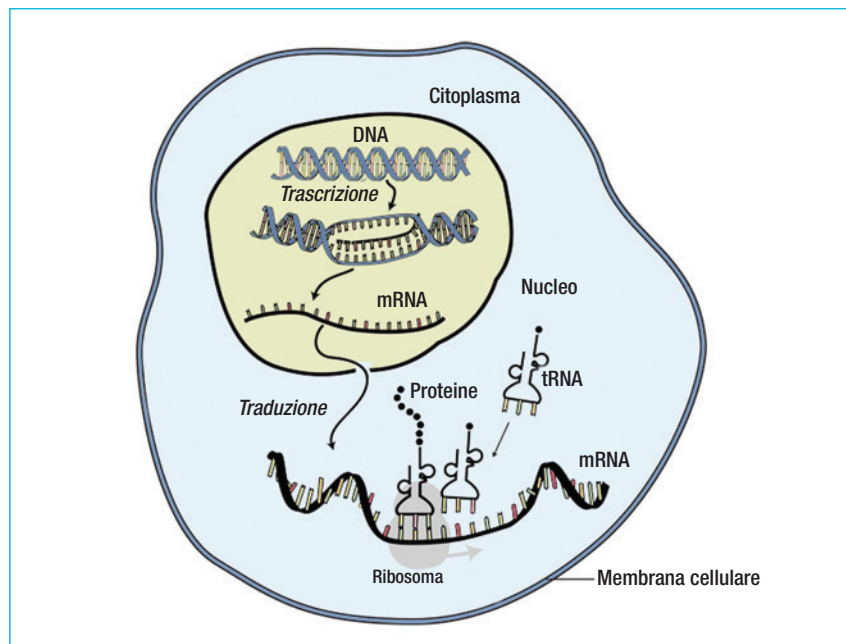
L'unicità del paziente e, dunque, la necessaria calibrazione della cura sul singolo, trova nelle scienze omiche il fondamento scientifico e gli strumenti per una gestione clinica sempre più appropriata.

Al clinico, dunque, è richiesto di conoscere almeno i lineamenti di queste nuove scienze, prendere dimestichezza con il loro linguaggio, sapere come attingere al patrimonio di informazioni necessarie sia per la

* Martina Collotta, laureata in Medicina e Chirurgia con lode e menzione speciale presso l'Università di Milano con una tesi sperimentale nel settore dell'epigenetica, si dedica allo studio della biologia molecolare. Attualmente, dopo l'esperienza acquisita nel campo della genetica delle malattie neuromuscolari, frequenta il Master in Molecular Biotechnology presso l'Università di Torino e svolge attività di ricerca in oncologia molecolare presso l'IRCCS FPO di Candiolo (Torino).

FIGURA 1.

Il dogma centrale della biologia molecolare. DNA → RNA → Proteine. L'informazione contenuta nel DNA, trascritta in RNA messaggero (mRNA) e tradotta attraverso il codice genetico che associa a ciascuna tripletta di nucleotidi gli aminoacidi, dirige la sintesi delle proteine strutturali e funzionali della cellula.



comprensione della letteratura scientifica, sia per le sempre più numerose applicazioni cliniche a cui sarà egli stesso chiamato. Partiamo dunque dalla letteratura.

Tutto ciò che è *omico* è spesso richiamato nel linguaggio anglosassone non solo dal suffisso *omics*, ma anche dalla specificazione *wide* o *whole*: *genome-wide association studies* (studi di associazione “su scala genomica”), *whole exome sequencing* (sequenziamento dell'esoma), sono solo due degli esempi che, implicitamente, riferiscono alle scienze omiche e che non raramente si trovano negli articoli con cui il clinico è chiamato a confrontarsi.

La presenza di questi termini ci richiede lo sforzo di ampliare il nostro sguardo, la nostra prospettiva di osservazione. La complessità della biologia della cellula non sta più unicamente nella rete di interazioni intercellulari che fanno dell'organismo un tutto organizzato, ma anche nella complessità insita in quanto accade all'interno della cellula stessa che, grazie alle scienze omiche, possiamo comprendere sempre meglio. Il “tutto” a cui allude il suffisso *omics*, infatti, non è solo la totalità del patrimonio di informazioni che DNA, RNA e

proteine contengono, ma è anche la *totalità delle interazioni funzionali* stanno alla base dei processi metabolici cellulari e che determinano la differenza tra cellula e cellula o tra organismo e organismo.

Il dogma non è più un dogma?

Sappiamo oggi che non è il solo DNA a determinare tale differenza: è l'espressione genica, è la sua regolazione, sono le modificazioni post-trascrizionali e le modifiche post-traduzionali a cui le proteine vanno incontro. Sono reti (*network*) di interazioni che si vengono a creare tra le tre componenti del dogma prima richiamato: DNA, RNA, proteine.

Il dinamismo intrinseco della cellula è scritto nel DNA, ma non possiamo più pensare al *dogma* in maniera “lineare” e come a un processo obbligato e necessario: il *dogma*, oggi, non è più realmente tale!

Se è vero che esso resta valido in linea generale, dobbiamo ricordare, ad esempio, che non tutti gli RNA codificano per proteine, ma molti di essi svolgono una funzione senza mai venire tradotti (pensiamo ai tRNA trasportatori di aminoacidi o, ai cosiddetti *non-coding RNA* regolatori di processi cel-

lulari, primo tra tutti, l'espressione genica). Se anche i dogmi, dunque, non sono più tali, significa che dobbiamo cominciare a pensare con quella flessibilità propria dei sistemi complessi in cui non c'è spazio per il determinismo, e le interazioni sono fatte di *feedback*, di fini regolazioni *nello spazio e nel tempo* della cellula e dell'organismo, in cui entrano in gioco una molteplicità di fattori.

Per comprendere tutto questo, sono necessari molti, sono necessarie informazioni *wide*, sono necessarie le scienze omiche: per dare spiegazione a quel che accade nella singola cellula, non possiamo focalizzarci su un singolo gene, ma dobbiamo guardare all'intero genoma e, ancora, andare *oltre* il genoma e guardare ai prodotti funzionali che dalla trascrizione e dalla traduzione derivano.

Infatti, parallelamente alla crescente comprensione dei dettagli molecolari del dogma (funzionamento dei meccanismi di replicazione del DNA, trascrizione e traduzione), è cresciuta anche la mole di informazioni a disposizione riguardante ciascuno degli *step* del dogma stesso.

Gli step del dogma su scala omica

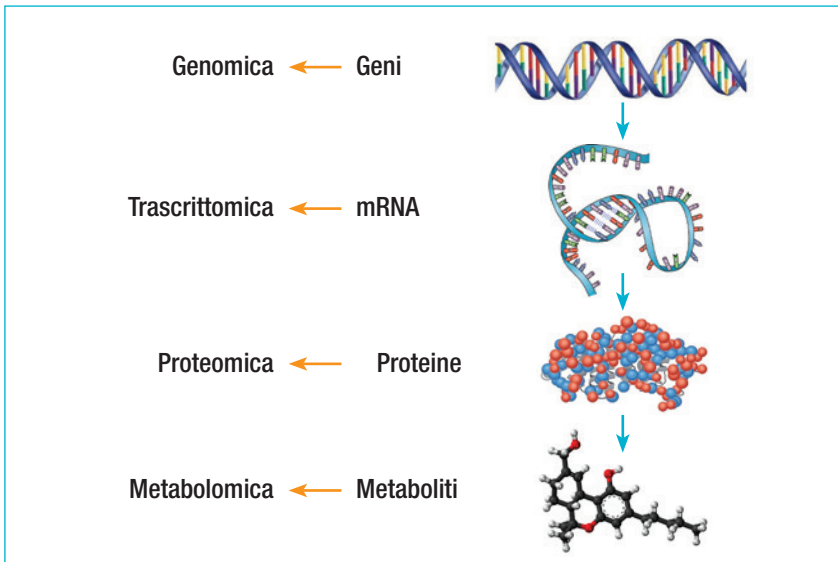
Il primo di questi *step* (Fig. 2), relativo al DNA, è la **genomica**.

Genetica e genomica non sono la medesima cosa: dobbiamo infatti sempre tenere a mente che tutto ciò che è *omico* è globale, comprensivo, abbraccia una grande quantità di informazioni. La genetica guarda al singolo gene, la genomica a *tutti* i geni dell'organismo, confronta il patrimonio genetico di organismi appartenenti a specie diverse (*genomica comparativa*), descrive le funzioni e le interazioni dei geni e dei loro prodotti (*genomica funzionale*) e rivolge il suo sguardo anche a quelle sequenze di DNA che non contengono geni e che sembrano essere “spazzatura” e sul cui significato ancora ci si sta interrogando.

Il secondo *step*, la **trascrittomica**, si occupa dello studio degli RNA messaggeri (mRNA) della cellula, frutto della trascrizione del DNA. Il trascrittoma (insieme di tutti gli mRNA) racconta le peculiarità della singola cellula, ci parla della sua differenziazione, delle sue specifiche funzioni e riflette in tempo reale gli stimoli a cui la cellula è

FIGURA 2.

Le scienze omiche. A ciascuno degli step del dogma corrisponde una delle tre omiche descritte: genomica, trascrittomica, proteomica, a cui si aggiunge la metabolomica che analizza i metaboliti frutto dei processi cellulari (vedi Glossario).



sottoposta, dai messaggi provenienti dalle altre cellule, fino alla risposta allo stress. L'importanza dello studio del *pool* degli mRNA cellulari in patologia, infatti, risiede nella loro capacità di descriverci le condizioni in cui si trova quella precisa cellula in quel determinato contesto: gli mRNA trascritti da una cellula sana, ad esempio, non sono gli stessi di una cellula tumorale. È dal confronto dei diversi trascrittomi tra loro, che possiamo apprendere molto di quanto accade nelle condizioni di malattia, andando alle radici cellulari di ogni cambiamento macroscopico che notiamo nell'organismo. Dagli mRNA tradotti, otteniamo le proteine: dello studio dell'insieme di esse, della loro struttura e funzione, si occupa la **proteomica**, terzo e ultimo *step* del dogma su scala omica.

Le proteine cellulari possono avere un ruolo strutturale oppure funzionale. Sono queste ultime che sono di particolare interesse: messaggeri intra- e intercellulari, recettori, enzimi... responsabili delle attività della cellula. La caratterizzazione molecolare della cellula stessa può avvenire tramite le proteine di membrana: pensiamo ad esempio al caso dei linfociti, tipizzati in base ai *Cluster of Differentiation* (CD) presenti sulla loro superficie. Il fatto che una proteina venga espressa unicamente da un certo tipo di cellule o solo in determinate condizioni patologiche, per-

mette di utilizzarla come *target* per terapie mirate con il beneficio di ridurre gli effetti collaterali a livello sistemico.

Perché la bioinformatica

Omica, come detto, fa riferimento a un concetto di *totalità*. Diventa allora facile intuire che dietro le scienze omiche si nasconde anche una grande quantità di *dati* che non possono essere gestiti e interpretati se non attraverso una nuova scienza: la bioinformatica. Questa nasce con le nuove tecnologie che permettono di *sequenziare* in modo più veloce ed economico i pool di acidi nucleici cellulari, moltiplicando così le analisi effettuate e dunque i dati a disposizione. Non si tratta però solo di raccogliere e conservare dati, ma anche e soprattutto di *trovare in essi correlazioni* attraverso analisi statistiche-matematiche che dal *caos* di milioni di lettere (i nucleotidi degli acidi nucleici) si trasformino in un *cosmos* più o meno ordinato di *pathway* (percorsi) intracellulari che dal DNA, attraverso l'RNA, conducono alle proteine.

È dietro questi *pathway*, dietro le tracce lasciate dai processi metabolici (oggetto di studio della metabolomica), che la complessità biologica si rivela e ci permette di cogliere il significato di dati che, altrimenti, avrebbero unicamente un fine compilativo.

Al dettaglio molecolare, patrimonio di conoscenza delle scienze omiche, dunque, si giunge solo attraverso l'analisi delle reti di interazioni dietro cui esso si nasconde.

È sempre attraverso gli strumenti forniti dalla bioinformatica che possiamo confrontare i dati molecolari relativi a un individuo, con quelli di altri, per poter così cogliere quelle sottili differenze (a volte di un solo nucleotide su milioni!) che ci fanno passare dalla normalità alla patologia, che ci predispongono a sviluppare una malattia o che predicono la nostra risposta a un'eventuale terapia.

È palese che tutto questo non si possa apprendere se non attraverso un confronto *ad extra*, capace cioè di guardare ad altri individui per trovare una *normalità informativa* sulla base della quale parlare di predisposizione o di patologia dando un "semplice" sguardo al genoma. Se, dunque, è per certi versi vero che nel DNA è scritto *tutto* circa l'organismo-paziente, è sempre più vero che per comprendere questo *tutto* che vi è scritto, è necessario il confronto con molti altri.

Come per i valori normali del laboratorio, divenuti ora valori di riferimento, il concetto di *normalità* o di *genoma di riferimento*, è ovviamente statistico, e necessita di continui aggiornamenti. La *profondità* della lettura del genoma, infatti, non è più solo quella della definizione della sequenza di basi, ma è anche quella che è capace di indicarne le modificazioni che non intaccano la sequenza, eppure cambiano l'espressione genica (*epigenoma*), il che rende tutto estremamente più problematico, ma, proprio per questo, molto più capace di rispecchiare fedelmente la complessità di quanto accade a livello di cellula e di organismo.

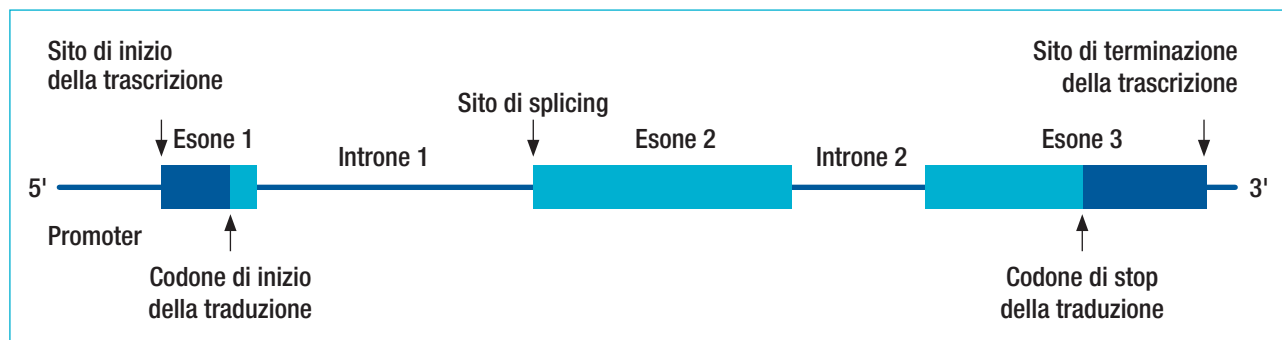
Ripensare il dogma

Il dogma dovrebbe essere ripensato *in quattro dimensioni*: non solo le tre determinate dalle corrispondenti omiche (trascrittomica, metabolomica, proteomica), ma anche una quarta: il *tempo*.

Le interazioni, infatti, sia quelle lineari che rispecchiano più fedelmente il dogma, sia quelle a *network*, con *feedback* (risposte) e modulazioni quali- e quantitative, mutano nel tempo: il tempo del ciclo cellulare, il tempo dato dalle diverse fasi di sviluppo dell'organismo, il tempo segnato dall'in-

FIGURA 3.

Struttura di un gene. I geni sono costituiti da segmenti codificanti (esoni) e non (introni), separati tra loro da siti di splicing. Hanno un sito di inizio e uno di terminazione della trascrizione, caratterizzati da specifiche triplette nucleotidiche. Sequenze presenti a monte o a valle di tali siti (ad es. promoter), hanno funzione di inibire o attivare la trascrizione e, dunque, di regolare l'espressione genica.



terazione con un ambiente a sua volta mutevole.

Trovare le tracce di questi mutamenti nel tempo, così come ricostruire quella tridimensionalità non lineare, ma fatta di *loop* (percorsi circolari) e *feedback* regolatori, è possibile solo quando i dati a disposizione sono molti e riflettono l'organizzata variabilità dei processi molecolari.

Le scienze omiche permettono proprio questo *consapevole sguardo di insieme* che va oltre la staticità unidimensionale del dogma, per dare una lettura sempre più *veritiera* e, dunque, *predittiva*, dei processi cellulari.

Dal laboratorio al letto del malato

Questa forma di conoscenza, sempre più in grado di descrivere la complessità dell'organismo, diventa allora un utile strumento non solo per le scienze di base, ma per la clinica.

La *medicina traslazionale* attinge dalle scienze omiche un patrimonio di conoscenze che permetterà di calibrare sempre più e sempre meglio la terapia sul singolo paziente, a partire da una più precisa conoscenza della sua patologia e della sua risposta clinica a livello molecolare.

Farmacogenomica, tossicogenomica, geno-

mica nutrizionale: sono solo alcuni degli esempi di *"spin off"* delle scienze omiche (oltre che omiche esse stesse) che avranno per il clinico un'importanza crescente.

Quale vantaggio, dunque, dalle scienze omiche, per la pratica clinica? A che pro conoscerne almeno i fondamenti? La risposta viene dal fine stesso della medicina: la cura del *singolo* paziente, che proprio perché *unico* richiede una risposta sempre più calibrata su *di lui* e sulla peculiare forma in cui la patologia *in lui* si manifesta.

La medicina personalizzata, le *target therapy*, sono il futuro della pratica clinica, ma per conoscere quell'unicità che caratterizza il singolo paziente, è necessario avere a disposizione quello sguardo d'insieme che ci dà il riferimento in cui collocare i concetti di normalità e di patologia.

Omico deve diventare anche il nostro modo di pensare: abbracciare in uno sguardo la totalità biologica del paziente, a partire dal più piccolo dettaglio molecolare fino allo sguardo epidemiologico che lo colloca nella realtà.

Sitografia

Il clinico può trovarsi nella necessità di dover sapere di più circa un gene o la proteina da esso codificata: pensiamo al caso

dei referti di anatomia patologica relativi a una neoplasia, alle analisi del cariotipo fetale o alle malattie rare.

In questi siti – oltre che nella ben nota banca dati *PubMed* – è possibile trovare informazioni circa il genoma umano: molte sono le informazioni contenute in essi (spesso molto tecniche!), ma non manca al clinico la capacità di selezionare quello che è utile per la sua pratica e per crescere nelle conoscenze professionali.

UCSC *University of California Santa Cruz genome browser*: raccolta delle informazioni frutto del progetto genoma umano.

OMIM *On-line Mendelian Inheritance in Man*: particolarmente utile per lo studio delle malattie rare.

TCGA *The Cancer Genome Atlas*: utile strumento in oncologia.

Bibliografia di riferimento

Alberts B, Johnson A, Lewis J, et al. *Molecular Biology of the Cell*. New York: Garland Science 2002.

Hogan RP, Kenny LC. "Omic" technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist* 2011;13:189-95.

Omics gateway (dal sito della rivista *Nature* alcune informazioni base sulle scienze omiche).

GLOSSARIO

Esoni	Sequenze di DNA all'interno di un gene che vengono trascritte in RNA messaggeri (mRNA). È la parte di DNA codificante le proteine. Considerati nel loro insieme, gli esoni di un intero genoma, formano l' <i>esoma</i> (Fig. 3).
Epigenetica	Studio dei cambiamenti ereditabili dell'espressione genica o del fenotipo cellulare, causati da meccanismi differenti rispetto ai cambiamenti della sequenza del DNA (ad es. metilazione di basi azotate, acetilazione di istoni).
Espressione genica	Insieme di eventi che, dalla trascrizione di un gene, arrivano alla produzione della macromolecola funzionale (proteina) corrispondente. L'espressione di un gene può essere finemente regolata attraverso uno o più meccanismi che intervengono a livello pre- o post- trascrizionale o traduzionale.
Genoma	Insieme dell'informazione genetica (DNA) propria di una cellula o organismo.
Introni	Segmenti di DNA che vengono inizialmente trascritti, ma che non sono presenti nell'RNA messaggero maturo e che, dunque, non contribuiscono alla codifica della proteina (Fig. 3).
Metaboloma	Insieme dei metaboliti presenti in una cellula (o in un organismo) che partecipano alle reazioni metaboliche cellulari.
Metabolomica	Studio sistematico dei metaboliti prodotti durante i vari processi cellulari. Si occupa anche del confronto dei diversi metabolomi presenti nella stessa cellula in tempi diversi (ad es. prima e dopo l'esposizione a un farmaco) o in condizioni diverse (normalità vs patologia).
mRNA	Vedi Trascrizione.
non-coding RNA	Trascritto genico che non va incontro a traduzione e, dunque, non codificante proteine. I non-coding RNA hanno diversi ruoli funzionali, particolarmente nella regolazione dell'espressione genica.
Proteoma	Insieme delle proteine presenti in una cellula o organismo, aventi ruolo strutturale o funzionale.
Sequenziamento	Processo attraverso cui viene determinato l'ordine dei nucleotidi che costituiscono l'acido nucleico (DNA o RNA).
Traduzione	Processo attraverso cui l'informazione genetica contenuta nell'RNA messaggero viene convertita in proteine. A ciascuna tripletta di nucleotidi, una molecola di RNA <i>transfer</i> (tRNA), accoppia un amminoacido secondo precise corrispondenze tripletta-amminoacido (<i>codice genetico</i>).
Trascrizione	Processo attraverso cui a partire da una molecola di DNA viene sintetizzata una molecola complementare di RNA. Nel caso si tratti di RNA messaggero (mRNA), il processo che fa seguito alla trascrizione è la traduzione (vedi).
tRNA	Vedi Traduzione.