



Università politecnica delle Marche
Dipartimento di scienze della vita e dell'ambiente

Tesi di Laurea in scienze biologiche
a.a 2020/2021

**«Due facce della stessa medaglia: le mutazioni non codificanti
rivelano nuovi meccanismi patologici e una maggiore comprensione
della regolazione dell'espressione genica»**

Relatrice:
Anna la Teana


Candidato:
Francesca Longo
S1082951

“ Le Mutazioni ”


La sequenza del DNA può essere modificata in diversi modi quali cambiamenti spontanei, errori nella replicazione, azione delle radiazioni o di particolari sostanze chimiche.

Qualsiasi alterazione delle sequenze delle basi nel DNA viene definita **Mutazione**

Le mutazioni possono avvenire sia all'interno del DNA codificante che nel DNA non codificante .

 Per DNA codificante si intende invece ogni sequenza di DNA che porta le informazioni per la codifica delle proteine.



 Per DNA non codificante o anche detto **noncoding DNA** si intende ogni sequenza di DNA in un genoma che non porta informazioni per proteine, quindi che non verrà tradotto.

Tali sequenze possiedono varie funzioni (es. vengono trascritte in rRNA, tRNA, ecc.). In passato, ignorata la funzione, era identificato come DNA spazzatura (**junk DNA**)

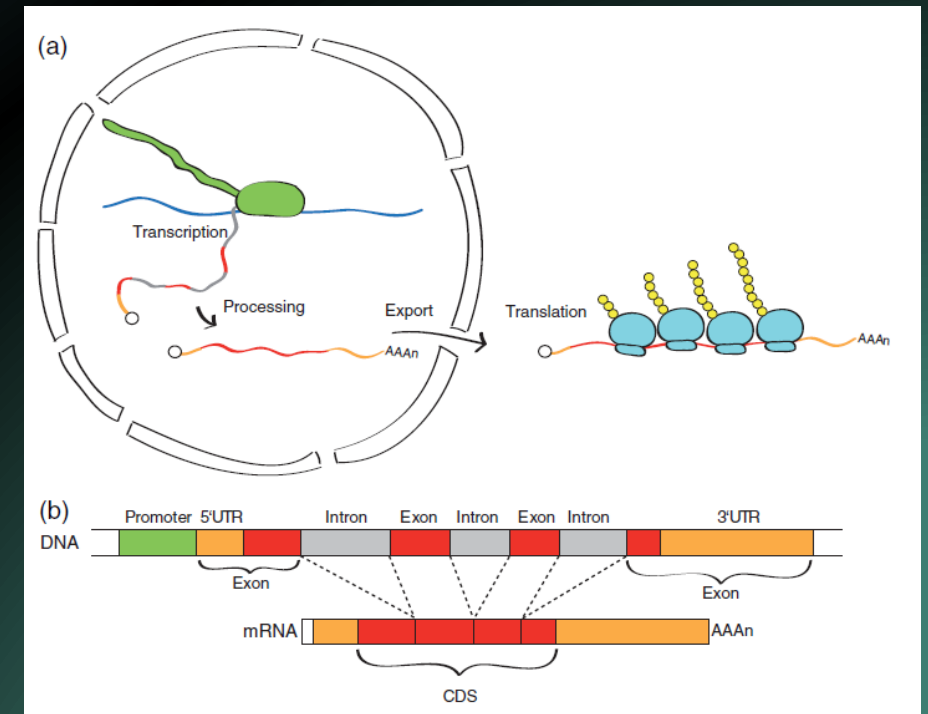


Introduzione

Le sequenze non codificanti costituiscono la maggior parte del Genoma umano e anche dei pre-mRNA. Le varianti a singolo nucleotide in queste regioni sono spesso trascurate ma possono essere responsabili di gran parte delle variazioni dei fenotipi osservati.

Meccanismo patologico

Sappiamo che il numero dei geni non è correlato alla complessità dell'organismo ma la complessità nasce piuttosto dalla regolazione e dall'interazione fra i diversi geni.



Tipo	Specie	Complessità di sviluppo	Dimensioni del genoma	Numero di geni stimato
Batterio	<i>Escherichia coli</i>	1 cellula procariotica	$4,6 \times 10^6$ bp	4200
Archeobattero	<i>Halobacterium sp. NRC-1</i>	1 cellula	$2,5 \times 10^6$ bp	2630 (2411)
Lievito	<i>Saccharomyces cerevisiae</i>	1 cellula eucariotica	$1,2 \times 10^7$ bp	5800
Verme	<i>Caenorabditis elegans</i>	~ 1000 cellule	1×10^8 bp	19099
Moscerino	<i>Drosophila melanogaster</i>	~ 50000 cellule	$1,8 \times 10^8$ bp	13601
Arabidopsis	<i>Arabidopsis thaliana</i>	10^{10} cellule	$1,2 \times 10^8$ bp	25000
Pesce palla	<i>Fugu rubripes</i>	10^{12} cellule	$3,8 \times 10^8$ bp	38000
Riso	<i>Oryza sativa</i>	5×10^{10} cellule	~ $4,6 \times 10^8$ bp	46000-56000
Topo	<i>Mus musculus</i>	10^{11} cellule	$3,2 \times 10^9$ bp	40000
Uomo	<i>Homo sapiens</i>	10^{14} cellule	$3,2 \times 10^9$ bp	35000

Dimensioni dei diversi genomi messi a confronto

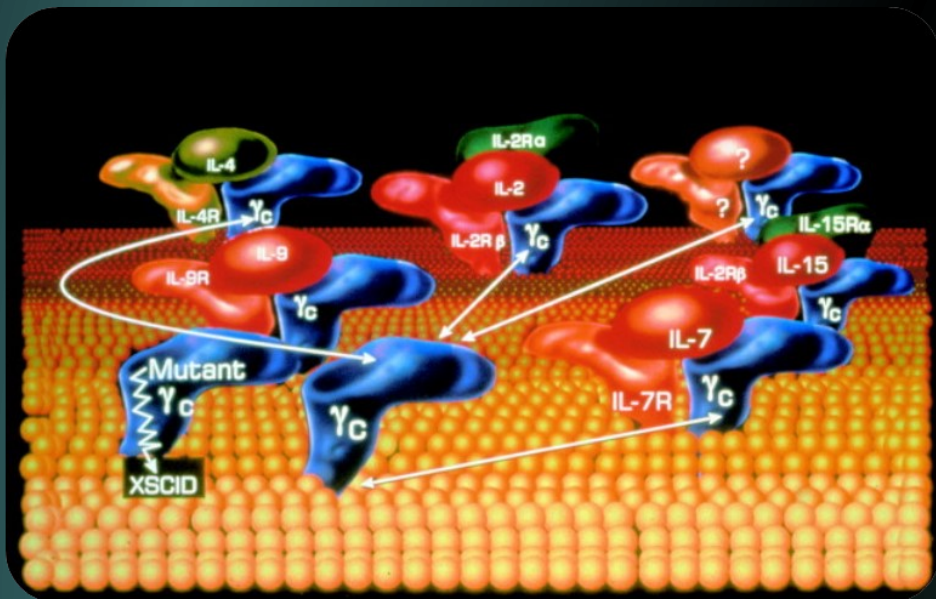


Mutazioni prese in considerazione

Immunodeficienza combinata grave (X-SCID)



- Malattia recessiva legata all'X
- Causata da una mutazione del gene IL2RG
- Il gene IL2RG è situato sul cromosoma Xq13 e codifica per la catena gamma comune (IL2R γ)
- Nei pazienti X-SCID le cellule T e cellule NK sono ASSENTI mentre le cellule B sono presenti ma non funzionanti.



	Gene	Mutation/variant	Disease	Effect	References
Mutations affecting splicing	<i>F9</i>	c.278-3A>G	Hemophilia B	New 3' SS, inclusion of 2 intronic nucleotides, frameshift, premature stop codon	Rogaev, Grigorenko, Faskhutdinova, Kittler, & Moliaka, 2009
	<i>IL10RA</i>	c.368-10C>G	VEO-IBD	New 3'SS, inclusion of 9 intronic nucleotides = 3 new amino acids, mislocalization of IL10RA protein	Murugan et al., 2014
	<i>CFTR</i>	c.3718-2477C>T (3,849 + 10 kb C>T)	Cystic fibrosis	New 5'SS, inclusion of 84 intronic nucleotides, pseudo-exon with in-frame stop codon	Highsmith et al., 1994
	<i>CTLA4</i>	c.458-1G>C, c.567 + 5G>C	CTLA-4 haploinsufficiency with autoimmune infiltration (CHAI) disease	Disruption of 3'SS or 5'SS, exon 3 skipping, expression of full-length protein reduced	Kuehn et al., 2014
	<i>IL2RG</i>	c.468 + 3A>C	Atypical X-SCID	Inactivation of 5'SS, partial usage of alternate 5'SS = insertion of 20 new amino acids	Ginn et al., 2004

	Gene	Mutation/variant	Disease	Effect	References
	<i>IL2RG</i>	c.87delG	Chronic parasitic disease, IL-21R-like deficiency	Single nucleotide deletion, frameshift, alternative 5'SS prevents premature termination and rescues <i>IL2RG</i> expression	Illig et al., 2019
	<i>LMNA</i>	c.1824C>T	Hutchinson-Gilford Progeria syndrome	Activation of alternative 5'SS due to altered RNA sequence and secondary structure	Shilo, Tosto, Rausch, Le Grice, & Misteli, 2019

IL2R γ è una componente dei recettori interleuchinici IL2R, IL4R, IL7R, IL9R, IL15R, IL21R.

Mutazioni che influenzano il sito di splicing

FENOTIPO X-SCID ATIPICO

TRASVERSIONE DA A → C DELLA TERZA BASE NELL'INTRONE 3

TUTTAVIA SONO STATI TROVATI ALCUNI mRNA correttamente maturati

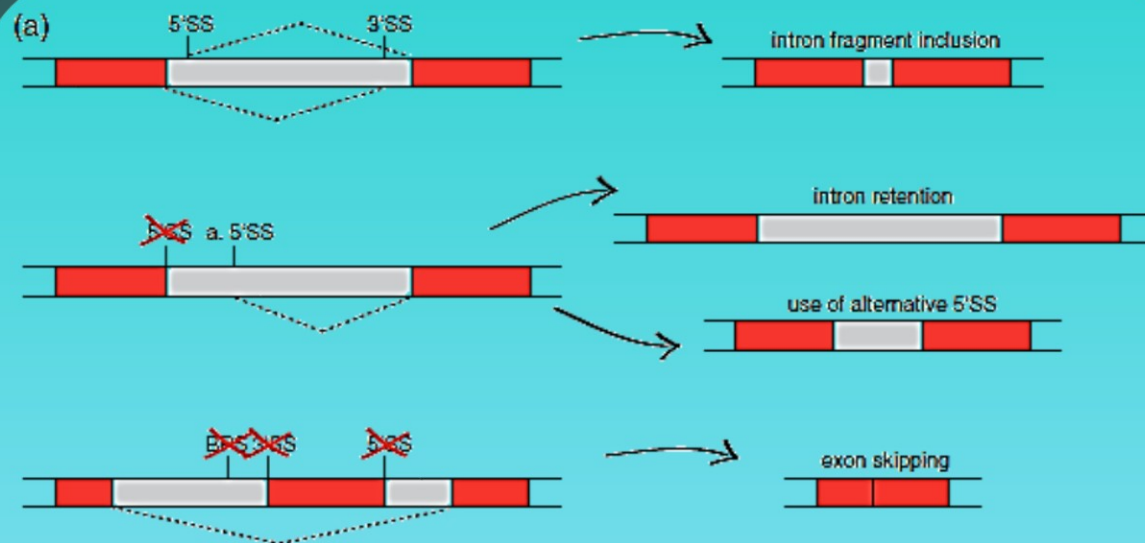
FENOTIPO X-SCID ATIPICO:

- ASSENZA DI CELLULE T
- CELLULE NK PRESENTI
- CELLULE B PRESENTI

LA MUTAZIONE DI UN SINGOLO NUCLEOTIDE ABBASSA LA COMPLEMENTARITA' DI UI_{sn}RNA E RIDUCE FORTEMENTE IL SITO 5'ss .

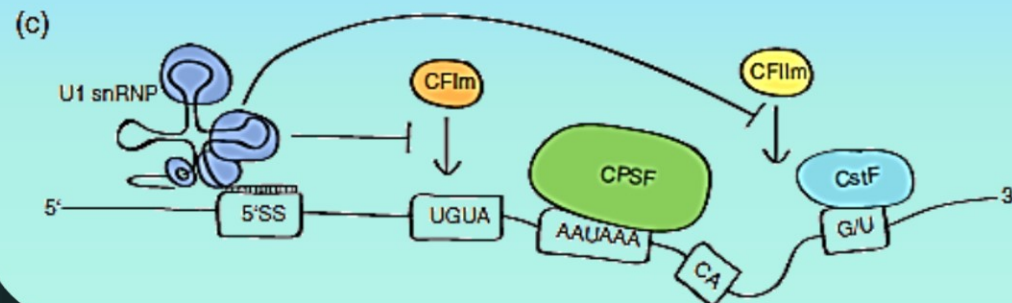
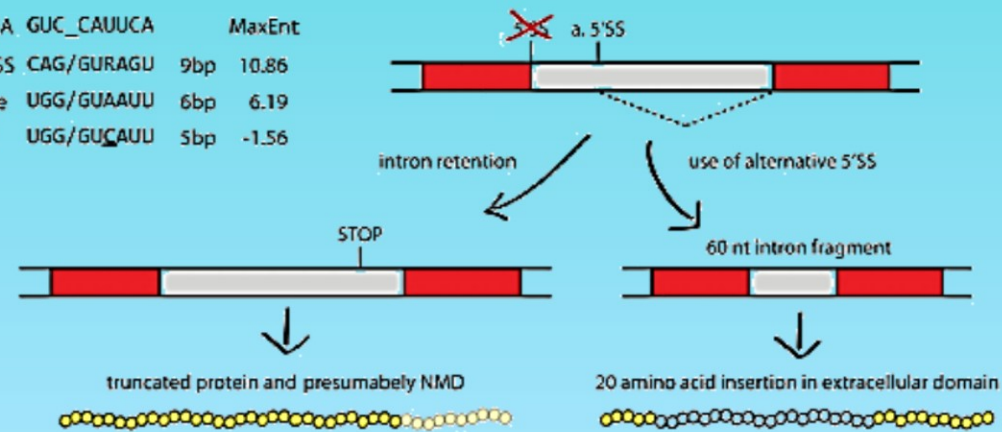
LA PROTEINA TRADOTTA DA QUESTO mRNA ha un'inserzione di 20 AA nel dominio extracellulare

LA RIDUZIONE DELLO SPLICING IN QUESTO SITO PORTA ALLA RITENZIONE DELL'INTRONE 3 NELLA MAGGIOR PARTE DEI TRASCritti E ALLA FORMAZIONE DI UN CODONE DI STOP IN-FRAME.



(b)

U1 snRNA	GUC_CAUUCA	MaxEnt
Cons. 5'SS	CAG/GURAGU	9bp 10.86
wild type	UGG/GUAAUU	6bp 6.19
patient	UGG/GUCAUU	5bp -1.56



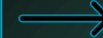
MUTAZIONI STRUTTURALI DELL'RNA

“UNA MUTAZIONE SINONIMA IN IL2RG PUO' FUNZIONARE ATTRAVERSO UNA STRUTTURA”

Due fratelli soffrono di una malattia parassitaria cronica e hanno dimostrato avere un'alterata risposta e segnalazione di IL-21



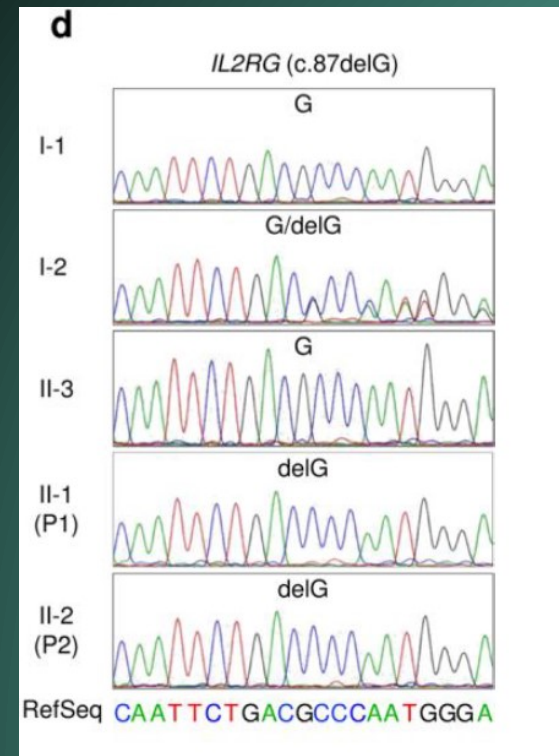
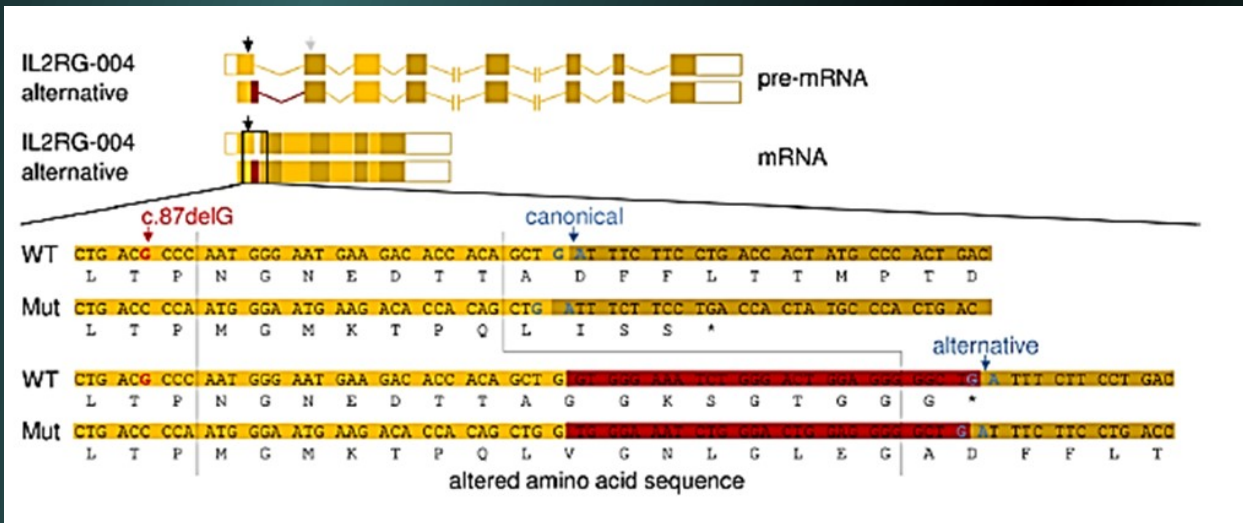
Questa risposta alterata è stata ricondotta ad una mutazione che riguarda la delezione di un singolo nucleotide dell'esone 1 di IL2RG.



Il frameshift risultante causerebbe la terminazione prematura della traduzione nell'esone 2, tuttavia una certa espressione di IL2R γ è mantenuta nei pazienti è questo è dovuto all'uso di un 5'ss alternativo nell'introne 1.

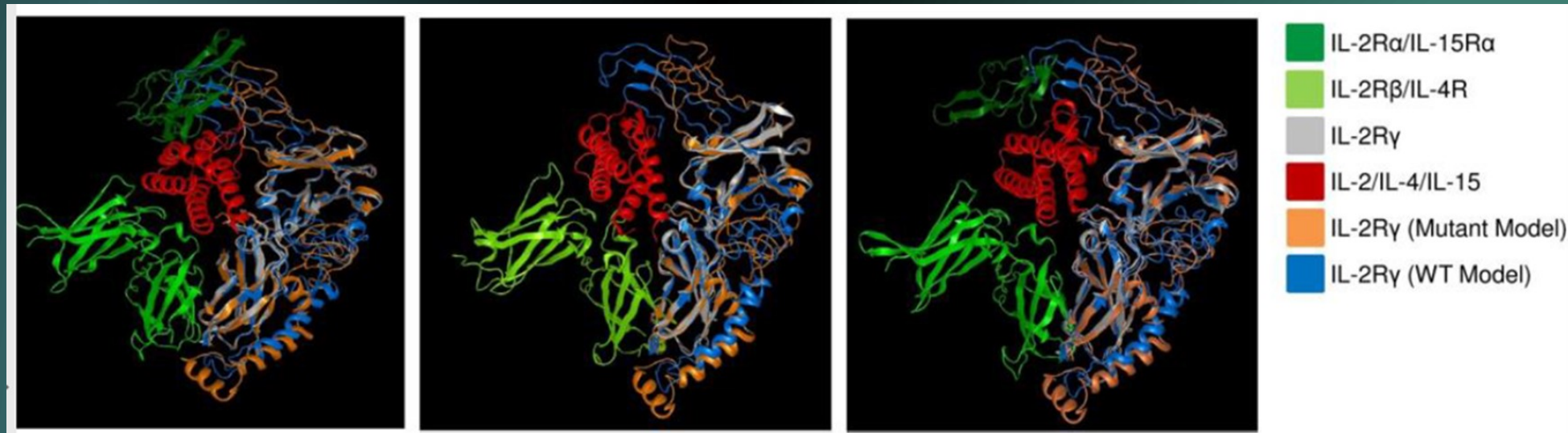


Aggiunta di 9 AA .



Si ipotizza che la delezione altera la struttura dell'RNA nell'introne alternativo 5'ss e permette di usarlo poco frequentemente.

I pazienti hanno mostrato principalmente una segnalazione difettosa di IL-4 e IL-21, mentre la segnalazione di IL-2, IL-7 e IL-15 era solo leggermente colpita. La segnalazione anormale di IL-4 e IL-21 è in linea con un'alterata differenziazione delle cellule B, e la segnalazione residua di IL-2, IL-7 e IL-15 potrebbe spiegare la presenza di cellule T e NK nei pazienti. Questi risultati sono stati inaspettati sulla base degli effetti deleteri previsti della delezione della coppia di basi in IL2RG. Gli esperimenti funzionali hanno mostrato che un meccanismo di splicing alternativo in IL2RG ha aggirato la perdita di espressione della proteina e ha portato all'attività di segnalazione residua.



CONCLUSIONE

Alcune delle varianti non codificanti mostrano un potenziale patogeno simile a scambi di amminoacidi o codoni di stop prematuri. Tuttavia, la maggior parte di loro funziona come modificatori fenotipici.

L'avvento del sequenziamento dell'intero genoma umano ha generato terabyte di dati che permettono di identificare tutti gli SNPs nel genoma di una persona. Tuttavia, collegarli con fenotipi è spesso impossibile.

I nuovi database che raccolgono i dati del genoma per consentire ai ricercatori di determinare la frequenza della loro particolare mutazione è un primo passo molto importante.

Nonostante tutti questi progressi, crediamo che sia importante capire meccanicamente come certe mutazioni causano patologie. Ci sono due ragioni principali per questo.

In primo luogo, un meccanismo patologico in buona fede permette ad altri ricercatori di cercare simili mutazioni non codificanti nel loro collettivo di pazienti. In secondo luogo, le mutazioni ci indirizzano a motivi e regioni di sequenza, che sono importanti, per esempio, per la regolazione della biogenesi dell'RNA. In questo modo si ottengono anche informazioni sulla regolazione dell'espressione genica cellulare. E come Gideon Dreyfuss ha dichiarato alla conferenza annuale della società RNA a Cracovia:

"alla fine, dobbiamo ancora sapere come le cose funzionano".

Inoltre, la comprensione dei meccanismi patologici delle mutazioni non codificanti offrirà anche nuove opzioni terapeutiche.

Il futuro introdurrà l'intelligenza artificiale e l'apprendimento automatico nel campo delle malattie genetiche. L'aumento della potenza di calcolo quantistico potrebbe essere in grado di simulare il comportamento di un'intera cellula in tutti gli stati possibili. Il risultato di una variante non codificante può essere prevedibile da questi strumenti. Queste simulazioni saranno così potenti da guidare i ricercatori e i clinici nella valutazione delle varianti e dare una prognosi o una classificazione della malattia.

RIFERIMENTI:

- Fox-Walsh, K. L., & Hertel, K. J. (2009). Splice-site pairing is an intrinsically high fidelity process. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 1766–1771.
- Ganser, L. R., Kelly, M. L., Herschlag, D., & Al-Hashimi, H. M. (2019). The roles of structural dynamics in the cellular functions of RNAs. *Nature Reviews. Molecular Cell Biology*, 20, 474–489.
- Ginn, S. L., Smyth, C., Wong, M., Bennetts, B., Rowe, P. B., & Alexander, I. E. (2004). A novel splice-site mutation in the common gamma chain (gammac) gene IL2RG results in X-linked severe combined immunodeficiency with an atypical NK+ phenotype. *Human Mutation*, 23, 522–523.
- Hsu, A. P., Fleisher, T. A., & Niemela, J. E. (2009). Mutation analysis in primary immunodeficiency diseases: Case studies. *Current Opinion in Allergy and Clinical Immunology*, 9, 517–524.
- Illig, D., Navratil, M., Kelecic, J., Conca, R., Hojsak, I., Jadresin, O., ... Kotlarz, D. (2019). Alternative splicing rescues loss of common gammachain function and results in IL-21R-like deficiency. *Journal of Clinical Immunology*, 39, 207–215.
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., ... International Human Genome Sequencing Consortium.
- Lai, D., Proctor, J. R., & Meyer, I. M. (2013). On the importance of cotranscriptional RNA structure formation. *RNA*, 19, 1461–1473
- Lim, C. K., Abolhassani, H., Appelberg, S. K., Sundin, M., & Hammarstrom, L. (2019). IL2RG hypomorphic mutation: Identification of a novel pathogenic mutation in exon 8 and a review of the literature. *Allergy, Asthma and Clinical Immunology*, 15, 2.
- Noguchi, M., Yi, H., Rosenblatt, H. M., Filipovich, A. H., Adelstein, S., Modi, W. S., ... Leonard, W. J. (1993). Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans. *Cell*, 73, 147–157.
- (2001). Initial sequencing and analysis of the human genome. *Nature*, 409, 860–921.
- Will, C. L., & Luhrmann, R. (2011). Spliceosome structure and function. *Cold Spring Harbor Perspectives in Biology*, 3, 1–25.
- Rochman, Y., Spolski, R., & Leonard, W. J. (2009). New insights into the regulation of T cells by gamma (c) family cytokines. *Nature Reviews. Immunology*, 9, 480–490.
- Shabalina, S. A., Spiridonov, N. A., & Kashina, A. (2013). Sounds of silence: Synonymous nucleotides as a key to biological regulation and complexity. *Nucleic Acids Research*, 41, 2073–2094.
- Solaymani-Mohammadi, S., Eckmann, L., & Singer, S. M. (2019). Interleukin (IL)-21 in inflammation and immunity during parasitic diseases. *Frontiers in Cellular and Infection Microbiology*, 9, 401.

RIASSUNTO:

Le sequenze non codificanti costituiscono la maggior parte del Genoma umano e anche dei pre-mRNA. Le varianti a singolo nucleotide in queste regioni sono spesso trascurate ma possono essere responsabili di gran parte delle variazioni dei fenotipi osservati.

In questo articolo i ricercatori hanno volto l'attenzione nello studio del meccanismo patologico per comprendere in maniera approfondita la malattia e i percorsi cellulari effettuati.

In particolare farò due esempi di Mutazione che avvengono a livello del gene IL2RG il quale codifica per la catena gamma (IL2R γ) una componente dei recettori interleuchinici IL2R , IL4R ,IL7R,IL9R,IL15R,IL21R.

La mutazione del gene IL2RG porta ad una malattia detta : Immunodeficienza combinata grave (X-SCID).

Nei pazienti X-SCID le cellule T e cellule NK sono ASSENTI mentre le cellule B sono presenti ma non funzionanti.

Tuttavia esporrò due diverse mutazione che possono avvenire a livello di questo gene che portano entrambi a fenotipi atipici. Nel primo caso la mutazione è una trasversione da A a C della terza base dell'introne 3 che porta ad un fenotipo X-SCID atipico con un numero quasi normale di cellule NK sebbene funzionalmente carenti (T-,NK+,B+). La mutazione (trasversione) in questo caso abbassa la complementarità di U1snRNA e la riduzione quindi dello splicing nel sito canonico 5'ss la cui forza si riduce drasticamente e inoltre si assiste alla ritenzione dell'introne nei trascritti di IL2RG non solo c'è anche la formazione di un codone di stop in frame. Sono stati trovati diversi mRNA di IL2RG diversi fra loro e questo è sicuramente dovuto allo splicing alternativo , alcuni trascritti contenevano solo i primi 60 nucleotidi dell'introne 3 indicando quindi l'uso di un 5'ss criptico situato nell'introne e la proteina tradotta quindi conterrà 20 AA in frame nel dominio extracellulare . Tuttavia sono stati trovati anche livelli minori di mRNA di IL2RG correttamente elaborati e quindi tracce di IL2R γ potrebbero essere rilevate su alcune cellule B e questo potrebbe spiegare il fenotipo atipico.

Il secondo caso riguarda sempre il gene IL2RG questa volta però avviene la delezione di un singolo nucleotide G che porterebbe a un frameshift e una terminazione prematura. Tuttavia l'uso di 5'ss alternativo impedisce la terminazione prematura e salva l'espressione di IL2RG se pur in piccole quantità. I dati suggeriscono che la mutazione IL2RG colpisce piuttosto selettivamente IL-4 e IL-21. Questo è in linea con l'osservazione che le cellule T e NK sono presenti in numero normale, mentre le cellule B sono notevolmente ridotte. . La segnalazione anormale di IL-4 e IL-21 è in linea con un'alterata differenziazione delle cellule B, e la segnalazione residua di IL-2, IL-7 e IL-15 potrebbe spiegare la presenza di cellule T e NK nei pazienti. Questi risultati sono stati inaspettati sulla base degli effetti deleteri previsti della delezione della coppia di basi in IL2RG. Gli esperimenti funzionali hanno mostrato che un meccanismo di splicing alternativo in IL2RG ha aggirato la perdita di espressione della proteina e ha portato all'attività di segnalazione residua. Il frameshift provocato dalla delezione è risolto da meccanismi di splicing alternativo nell'introne 1 di IL2RG .

L'analisi bioinformatica ha suggerito che il sito di splicing alternativo crea un codone di stop prematuro in IL2RG wild-type ma porta al ripristino di una versione produttiva dell'mRNA di IL2RG a partire dall'esone 2 . Si prevede che la proteina abbia 16 aminoacidi mutati a causa del frameshift e dello splicing a valle dell'esone 1, mentre la sequenza amminoacidica del peptide di segnale, il dominio transmembrana e il dominio di segnalazione di IL2R γ rimangono invariati e questo porta ancora una volta ad un fenotipo X-SCID atipico.



Grazie per
l'attenzione!

FRANCESCA LONGO