

UNIVERSITÀ POLITECNICA DELLE MARCHE
Corso di Laurea Triennale in Scienze Biologiche

Relazione finale

***ELEMENTI TRASPONIBILI: una fonte di
variabilità genetica con effetti sulla
regolazione del genoma e sulla salute umana***



Candidata: Martina Silvestrini

Relatore: Prof. Marco Barucca

Anno accademico 2019/2020

Riassunto

Gli elementi trasponibili hanno una notevole importanza per la variabilità genetica umana, in quanto generano polimorfismi.

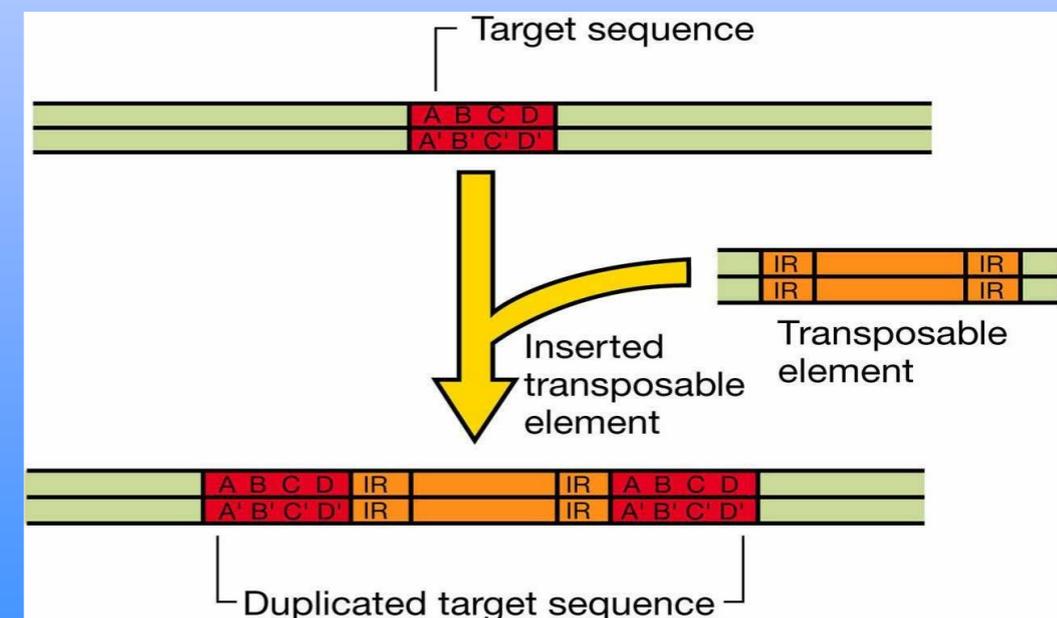
Per valutare l'impatto fenotipico di questi polimorfismi sono stati effettuati degli studi di associazione, utilizzando i set di dati genomici di popolazione, scoprendo così centinaia di associazioni tra inserzioni di elementi trasponibili e i livelli di espressione genica. Inoltre analizzando gli schemi di Linkage Disequilibrium tra polimorfismi generati dalle inserzioni e SNP precedentemente caratterizzati, hanno scoperto che essi hanno degli effetti sulle malattie umane.

Che cosa sono gli elementi trasponibili?

- Sono sequenze di DNA definite mobili, capaci di “trasporre”. Il 50% delle sequenze del genoma derivano da essi.
- Il loro successo evolutivo richiede una spiegazione ma la teoria del DNA egoista li definisce parassiti genomici, senza alcun ruolo funzionale.
- Numerosi studi hanno rivelato casi di domesticazione molecolare in cui perdono la capacità di trasporre e si convertono in elementi regolatori (enhancer/ silencer).
- Fino a questo momento gli studi si sono concentrati su elementi regolatori che derivano da inserzioni antiche. Si tratta di inserimenti fissi che si trovano nella stessa posizione all'interno di tutti i genomi degli individui umani.
- Negli ultimi anni, con le nuove tecnologie, gli studi si sono concentrati sull'attività in corso degli elementi trasponibili umani. Alcuni di essi infatti, mantengono la capacità di trasporre: Alu, SVA, L1 e il retrovirus HERV-K. Quando questi traspongono generano polimorfismi.



Fonte: google immagini



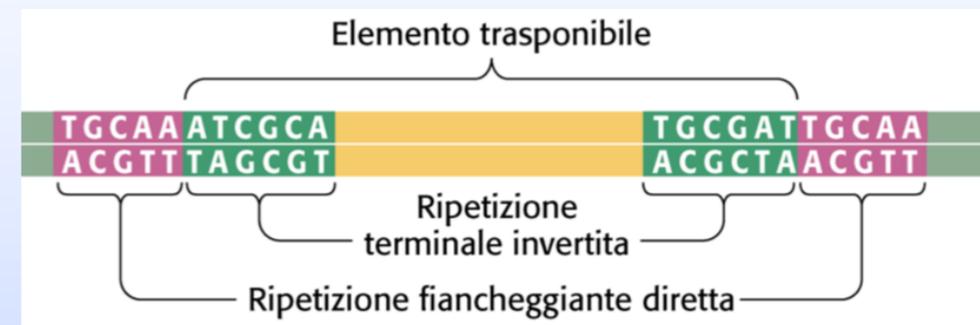
Fonte: <https://www.pinterest.it/pin/825636544157673195/>

Caratteristiche degli elementi trasponibili

□ Meccanismi di trasposizione: possono codificare per proteine che muovono il DNA dell'elemento stesso o replicare il DNA producendo un nuovo elemento che si inserisce in una nuova posizione.

□ Comprendono:

- **Sequenze di inserzione:** sono gli elementi trasponibili più semplici, contengono solo il gene per la trasposasi che riconosce le sequenze ripetute e invertite (IR) alle estremità di esse.
- **Trasposoni:** segmenti di DNA che contengono geni per la trasposizione con sequenze ripetute invertite alle estremità. Inoltre, entrambi i lati portano sequenze ripetute dirette che non fanno parte del trasposone, ma vengono generate nel sito bersaglio che viene tagliato in maniera sfalsata per permettere l'inserimento del trasposone stesso.



Schema di un trasposone

(Fonte:

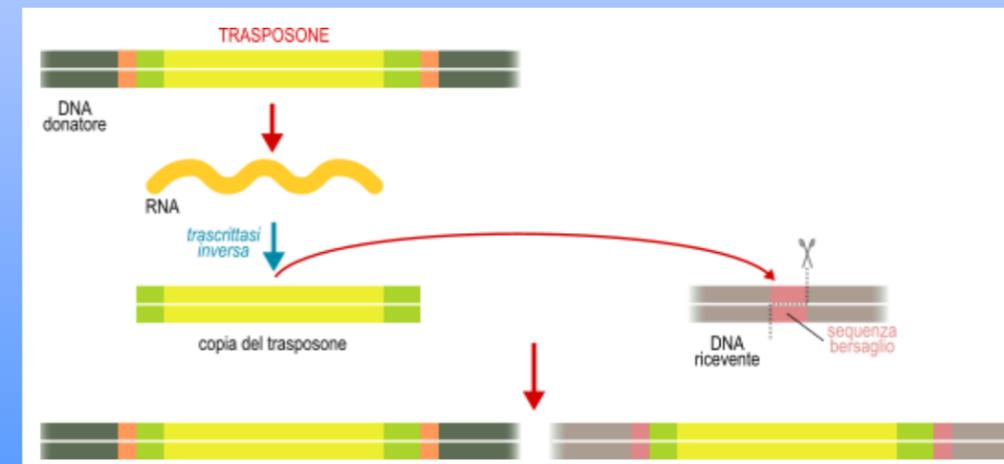
<http://profs.scienze.univr.it/delledonne/Insegnamenti/Genetica%2003.pdf>)

□ I retrotrasposoni (i trasposoni più studiati nell'uomo) attraverso retrotrascrizione convertono l'RNA in cDNA che poi si inserisce nel genoma. Si dividono in due tipologie:

1. **Elementi LTR:** derivano da virus e possiedono alle estremità sequenze ripetute.
2. **Elementi non LTR:** non possiedono sequenze ripetute e presentano una coda di poliA all'estremità 3'.

□ Gli elementi non LTR comprendono:

- **LINE:** (**L**ong **I**nterspersed **N**uclear **E**lements) lunghe sequenze di DNA che contengono geni per la trasposizione (trascrittasi inversa ed endonucleasi).
- **SINE** (**S**hort **I**nterspersed **N**uclear **E**lements) brevi sequenze di DNA che comprendono Alu e SVA, che non codificano per trascrittasi inversa ma necessitano di proteine, codificate da altre sequenze, per trasportare (LINE).



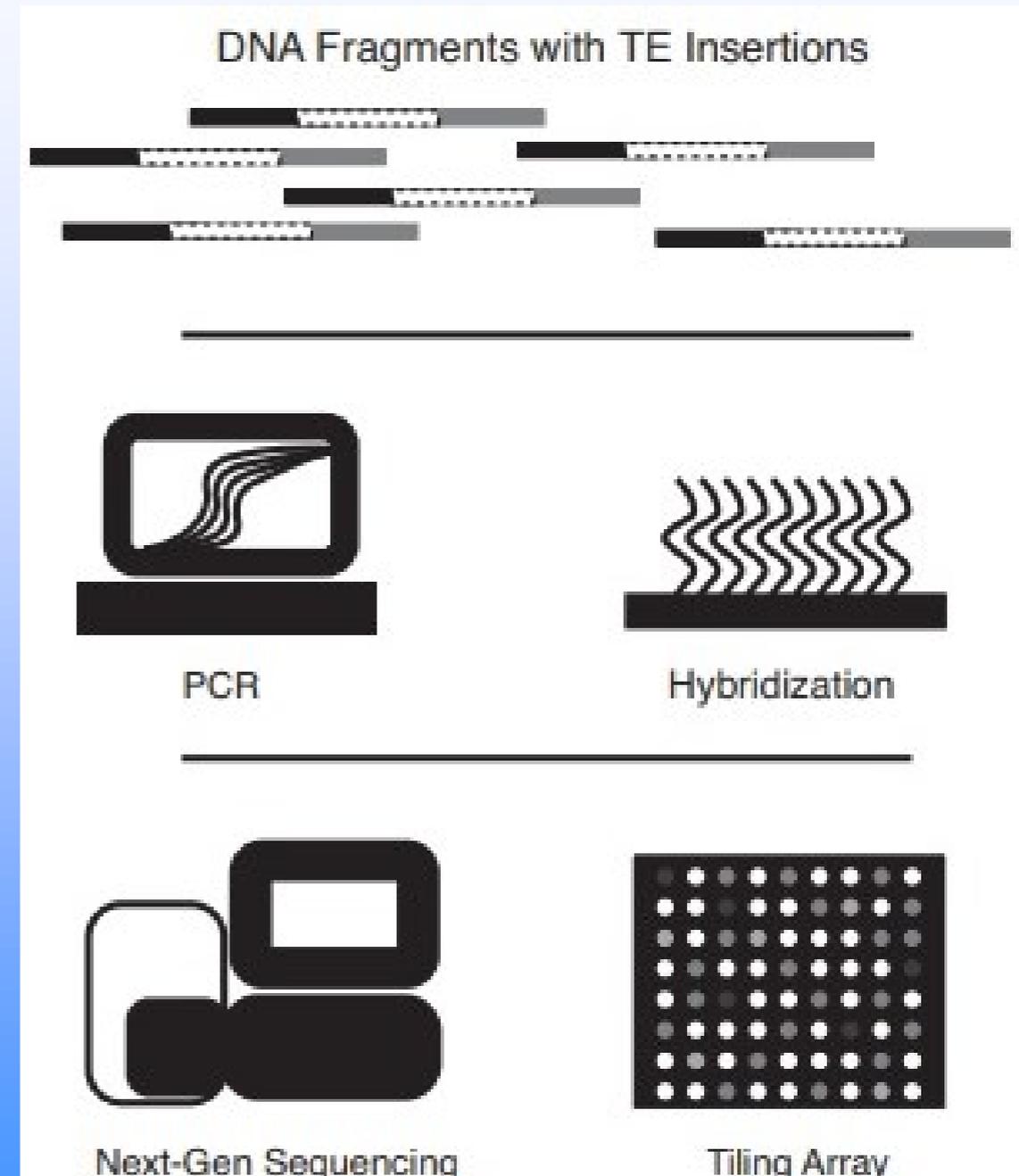
Meccanismo di trasposizione di un retrotrasposone

(Fonte: <http://www.gmpe.it/biologia/cromosoma>)

Metodi per definire le inserzioni

1. **Metodi bioinformatici:** si basano sull'elaborazione elettronica dell'intera sequenza del genoma che differisce dalla sequenza di riferimento.
2. **Metodi sperimentali ad alto rendimento** che si basano su:
 - Frammentazione
 - Arricchimento
 - Calling

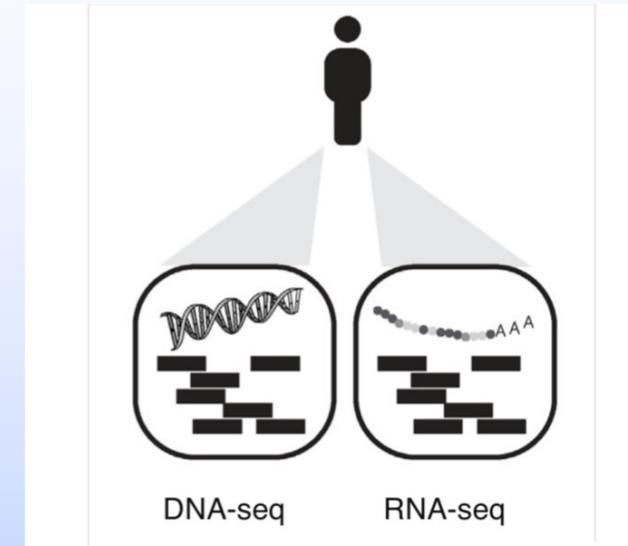
Questi approcci sperimentali sono stati applicati allo studio di varianti somatiche. Questo studio invece si concentra su mutazioni germinali che producono differenze inter-individuali (polimorfismi).



Metodi sperimentali ad alto rendimento

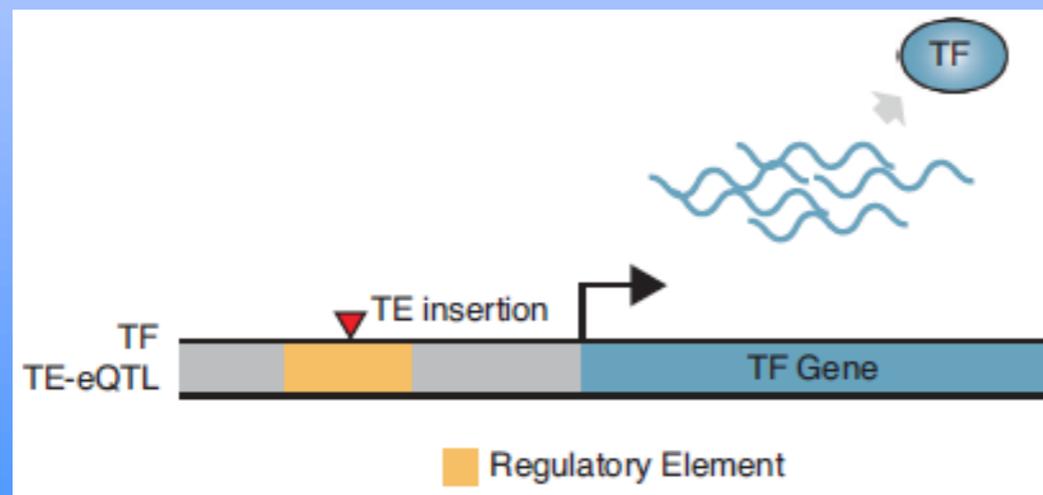
Polimorfismi generati da elementi trasponibili e regolazione del genoma

- Per valutare le conseguenze dell'attività in corso degli elementi trasponibili, a livello di popolazione, è stato effettuato uno studio di associazione considerando gli eQTL ovvero varianti associate a cambiamenti nei livelli di espressione genica.
- I genotipi di un locus possono essere:
 - 0 (omozigote) —> inserimento assente
 - 1 (eterozigote) —> un inserimento
 - 2 (omozigote) —> due inserimenti
- Questo studio di associazione richiede che i singoli campioni siano caratterizzati sia a livello genomico (DNA-seq) che a livello trascrittomico (RNA-seq).
- La fusione dei dati di due progetti ha permesso di confrontare i genotipi del sito di inserimento con i livelli di espressione genica degli stessi individui:
 1. il progetto 1KGP (1000 Genomes Project) ha definito i genotipi di inserzione di 2504 individui per 26 popolazioni umane.
 2. Il progetto GEUVADIS (**G**enetic **E**uropean **V**ariation in **D**isease) ha misurato l'espressione genica dei linfociti B di 445 individui (una popolazione africana e quattro popolazioni europee) del progetto 1KGP.

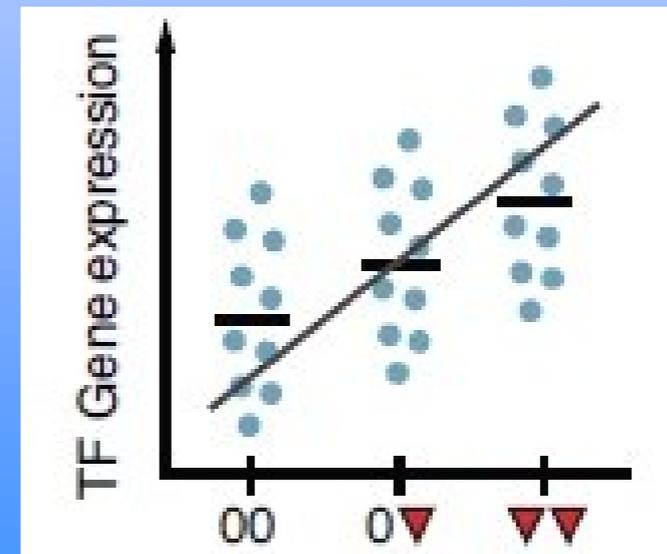


Per ogni individuo è stato analizzato il genoma intero e il trascrittoma corrispondente.

- Grazie a questo studio sono state scoperte centinaia di associazioni tra livello di espressione genica e polimorfismi generati da elementi trasponibili, sia all'interno che tra le popolazioni africane ed europee.
- Inoltre è stato osservato che le inserzioni hanno effetti sulla regolazione dell'espressione genica sia in cis che in trans.
- Un esempio: il gene PAX5 nei linfociti B, in cui l'inserzione di Alu nella sequenza di regolazione è associata a una maggiore espressione del fattore di trascrizione (effetto cis).

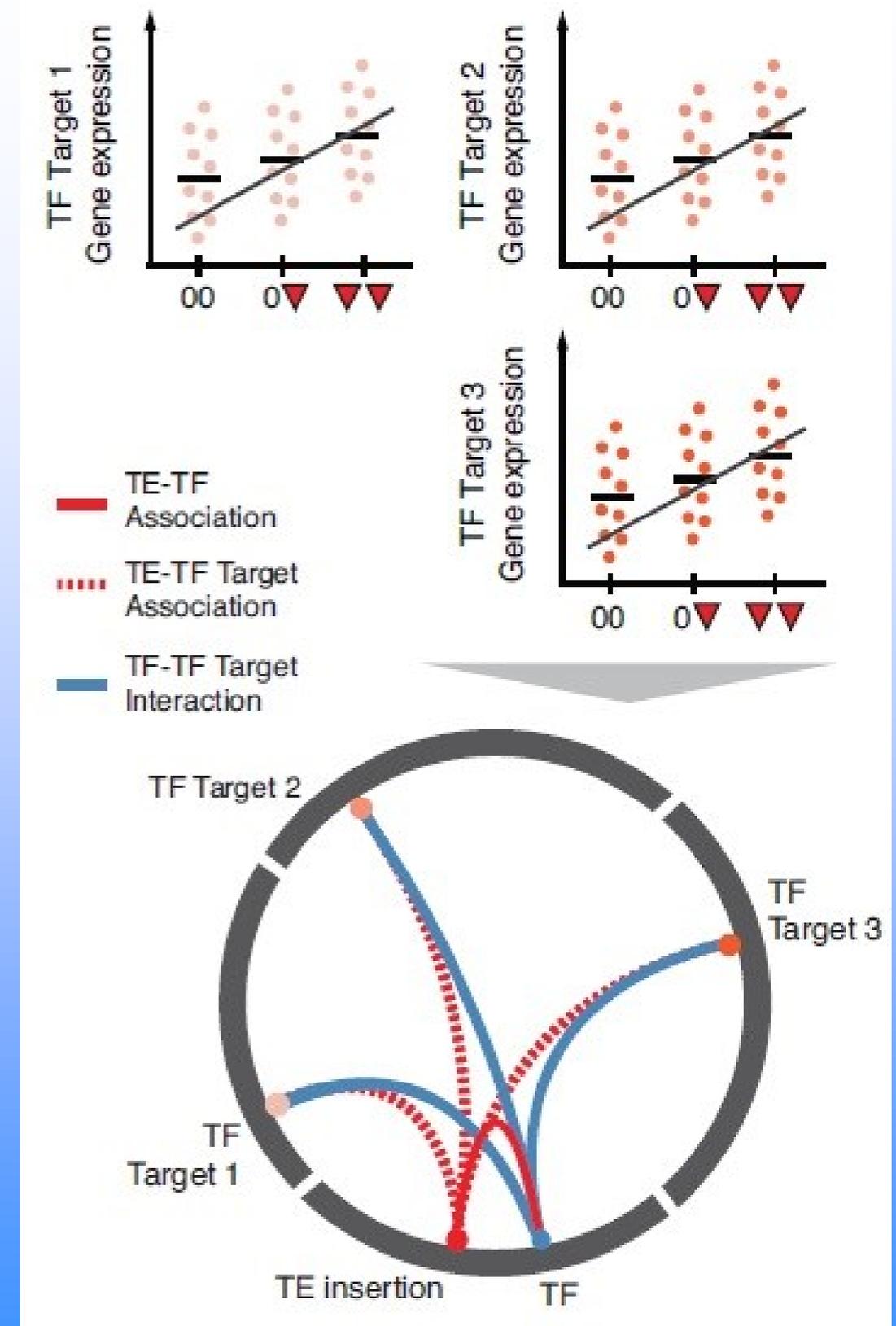


L'inserzione di Alu nella sequenza regolatrice



I genotipi del locus

- L'aumento dell'espressione del fattore di trascrizione, porta a sua volta a una maggiore espressione dei geni targets a valle (effetto trans).
- Questo esempio sottolinea come i singoli loci in cui abbiamo una inserzione, possono modificare intere reti regolatorie.
- È il primo e unico studio di questo genere sull'uomo, analoghi approcci sono stati utilizzati per scoprire le associazioni tra gli elementi trasponibili e livelli di espressione genica in Arabidopsis e mais.

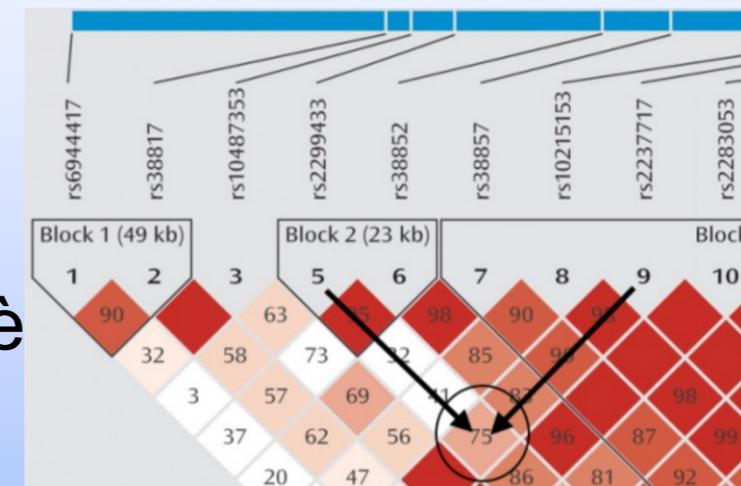


L'impatto dei polimorfismi sulle reti regolatorie

Impatto sulla salute e sulle malattie umane

Sempre a livello di popolazione sono stati effettuati degli studi che si basano sul Linkage Disequilibrium per scoprire i polimorfismi generati da elementi trasponibili collegati a SNP precedentemente associati a fenotipi di salute o malattia (GWAS).

- Il termine Linkage Disequilibrium indica che due geni sono fisicamente collegati. Lo scopo di una mappa LD è quello di dirci se due SNP vengono ereditati insieme.



Esempio di una mappa LD

- GWAS (Genome-Wide Association Study) è uno studio che si basa sul confronto di due gruppi, uno con una particolare malattia e uno di controllo, al fine di individuare gli SNP presenti solo nel primo gruppo per poi associarli alla malattia.
- Il gruppo di Kathleen Burns ha trovato 44 inserzioni Alu in LD con SNP precedentemente associati a malattie.

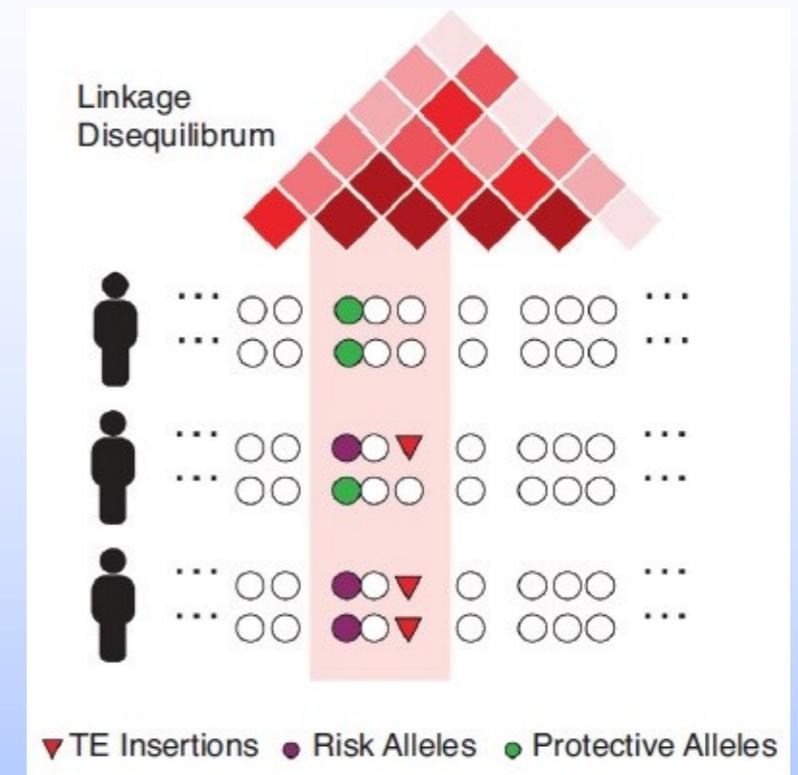
Questo studio vuole mettere in relazione la regolazione del genoma mediata da elementi trasponibili e gli effetti fenotipici correlati alla malattia

Per raggiungere questo obiettivo sono state considerate le inserzioni che generano polimorfismi che sono:

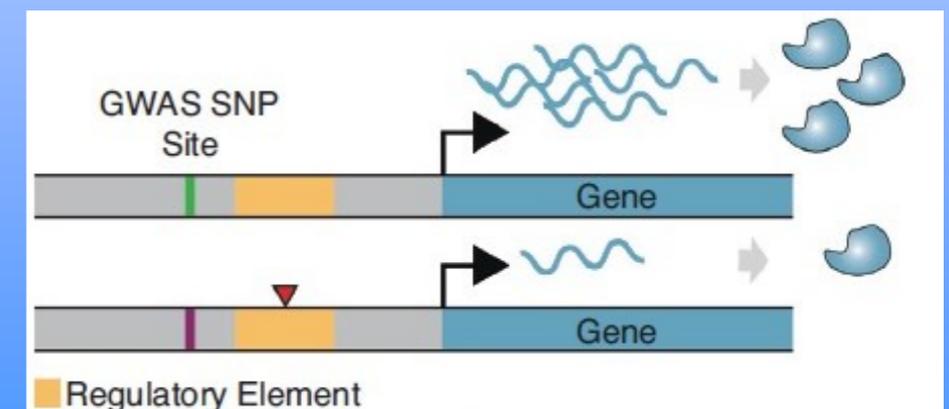
- In LD con alleli di rischio SNP precedentemente caratterizzati (GWAS)
- Situati all'interno di enhancer specifici del tessuto
- Associati a livelli di espressione genica

La ricerca è stata ristretta a geni che hanno funzioni ematiche o immunitarie, in quanto i dati dell'espressione genica che sono stati analizzati provengono da linfociti B.

Grazie a questo studio sono stati scoperti 6 polimorfismi generati da inserzioni che sono associati a fenotipi della malattia in virtù dei loro effetti regolatori. Questi includono sia elementi Alu che SVA.



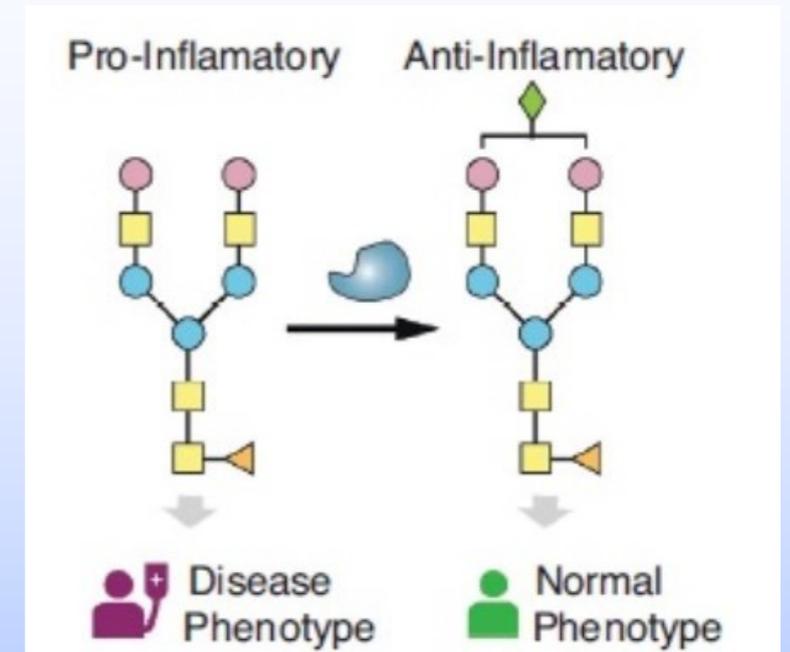
LD tra alleli a rischio e inserzioni



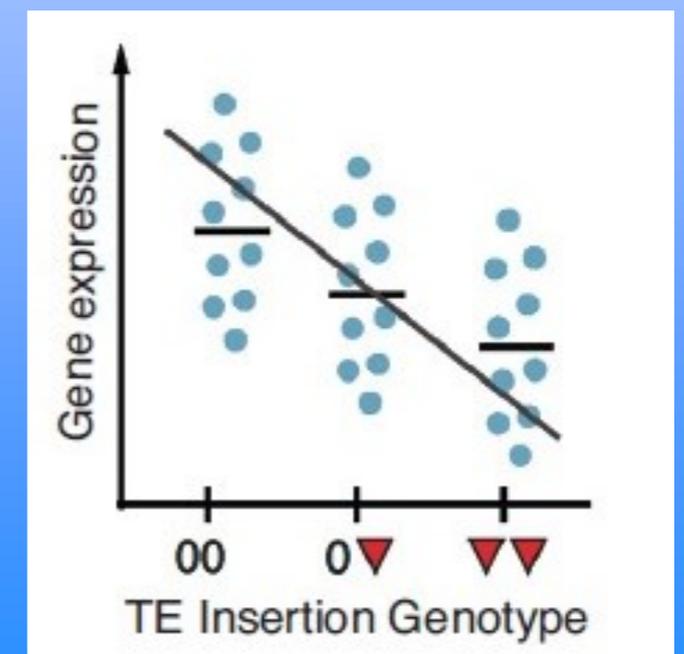
Inserzioni all'interno di enhancer associate a una ridotta espressione genica

Un esempio: inserimento di SVA nell'enhancer del gene B4GALT1

- ✦ Questo gene quando normalmente trascritto, agisce per convertire l'anticorpo immunologico IgG da una forma pro-infiammatoria a una anti-infiammatoria.
- ✦ L'inserimento di SVA nell'enhancer è associato a una minore trascrizione del gene B4GALT1.
- ✦ Nelle persone malate con la minore espressione del gene B4GALT1 abbiamo che l'anticorpo rimane nella forma pro-infiammatoria, portando ad un aumento dell'infiammazione e della malattia patologica correlata.



L'enzima converte l'anticorpo nella forma anti-infiammatoria



I genotipi del locus

Conclusioni

- Le indagini sugli effetti che i polimorfismi generati da elementi trasponibili hanno sulla regolazione del genoma e sulle malattie si basano su studi di associazione, sono quindi delle ipotesi che devono essere approfondite mediante studi sperimentali, per chiarirne la causalità e il meccanismo.
- Possiamo dire che gli elementi trasponibili sono una fonte di variabilità genetica naturale, con effetti sulla regolazione del genoma e sulla salute umana. È probabile che i polimorfismi da essi generati forniscano materiale di base cruciale per l'evoluzione umana in corso.

Bibliografia

- *Lu Wang and I King Jordan: Transposable element activity, genome regulation and human health. Current Opinion in Genetics & Development 2018, 49:25-33.*
- *Linkage disequilibrium:*
https://estrip.org/articles/read/tinypliny/44920/Linkage_Disequilibrium_Blocks_Triangles.html
- *GWAS: <https://knowgenetics.org/genome-wide-association-studies-gwas/>*