



POLITECNICO
MILANO 1863

Uno studio del Politecnico di Milano svela la “grammatica” del DNA

Su *Genome Biology* una ricerca sulle regole che governano la forma del DNA nello spazio

[Link all'articolo “Spatial patterns of CTCF sites define the anatomy of TADs and their boundaries”](#)

Milano, 27 agosto 2020 – La struttura tridimensionale del DNA è determinata da una serie di regole spaziali basate sulla presenza di particolari sequenze proteiche e sul loro ordine: è il risultato di uno studio recentemente pubblicato su **Genome Biology** da Luca Nanni, dottorando in Computer Science and Engineering al Politecnico di Milano, congiuntamente ai Professori Stefano Ceri dello stesso Ateneo e Colin Logie dell'Università di Nijmegen.

“La **maggior innovazione** portata dal nostro studio risiede nell’aver per la prima volta identificato delle **precise regole di disposizione** di alcune particolari proteine chiamate CTCF. La bellezza e la semplicità della grammatica di CTCF ci dimostra come la natura e l’evoluzione producano regolarità e sistemi incredibilmente ingegnosi e funzionali” spiega Luca Nanni, primo autore dello studio. “La conoscenza di queste regole ci potrà permettere in futuro di ingegnerizzare le sequenze di CTCF in modo da ottenere la desiderata struttura tridimensionale del DNA, ad esempio per **far interagire due geni che altrimenti non sarebbero in contatto**. Plasmare la struttura del DNA aprirà porte ancora sconosciute alla creazione di **farmaci per il trattamento di malattie** come il cancro”.

La molecola del DNA, che sarebbe lunga circa due metri se completamente srotolata, per risiedere nel nucleo delle nostre cellule deve avvilupparsi secondo un complesso sistema che ne mantenga l’accessibilità e la corretta lettura. Di particolare importanza nello studio della struttura tridimensionale del genoma sono i cosiddetti **domini topologici**, che si pensa abbiano una funzione di **aggregazione di zone di DNA con ruoli e comportamento simili**. Ad esempio, geni con funzione analoga hanno alta probabilità di risiedere nello stesso dominio topologico. “Ci siamo concentrati su alcune specifiche sequenze di DNA che codificano per la **proteina chiamata CTCF**” prosegue Nanni. “Questa proteina ha la funzione di isolare porzioni di DNA e quindi creare le barriere fra i vari domini topologici. Tramite l’ausilio di simulazioni al computer e l’ideazione di un modello di classificazione di queste proteine in base alla loro orientazione, siamo riusciti a **individuare una regolarità sorprendente nella loro disposizione lungo la sequenza del DNA**”. Lo studio ha evidenziato che l’orientazione e l’ordine di queste

Ufficio Relazioni con i Media
Politecnico di Milano
Piazza Leonardo da Vinci 32
20133 Milano

T +39 02 2399 2443
C. +39 334 1085172
relazionimedia@polimi.it
www.polimi.it



POLITECNICO
MILANO 1863

sequenze di DNA consente di ricostruire i domini topologici. Il genoma umano quindi si compatta seguendo una logica dettata da una “grammatica” i cui elementi sono le sequenze di CTCF, il loro orientamento e la distanza fra di loro.

Ufficio Relazioni con i Media
Politecnico di Milano
Piazza Leonardo da Vinci 32
20133 Milano

T +39 02 2399 2443
C. +39 334 1085172
relazionimedia@polimi.it
www.polimi.it



POLITECNICO
MILANO 1863

A Politecnico di Milano study reveals DNA "grammar"

Genome Biology published a research on the rules governing the shape of DNA in space

[Link to article "Spatial patterns of CTCF sites define the anatomy of TADs and their boundaries"](#)

Milan, 27 August 2020 - DNA three-dimensional structure is determined by a series of spatial rules based on particular protein sequences and their order. This was the finding of a study recently published in **Genome Biology** by Luca Nanni, PhD student in Computer Science and Engineering at Politecnico di Milano, together with Professors Stefano Ceri of the same University and Colin Logie of the University of Nijmegen.

The first author of the study Luca Nanni said: "Our study's **greatest innovation** lies in having identified **precise rules for the disposition** of CTCF proteins. The beauty and simplicity of CTCF's grammar shows us how nature and evolution produce regularity and incredibly ingenious and functional systems." "Knowing these rules allow CTCF sequences to be engineered to obtain the desired DNA three-dimensional structure. For example, it should be possible **to make two disconnected genes interact**. Moulding DNA structure will open doors to the creation of **pharmaceuticals for the treatment of diseases** such as cancer."

The DNA molecule, which would be about two metres long if completely unrolled, wraps itself based on a complex system that maintains its accessibility and correct reading to reside in the cell's nucleus. Crucial in the study of the three-dimensional structure of the genome are **topological domains**, which are thought to **aggregate DNA zones with similar roles and behaviour**. For example, genes with similar function are likely to reside in the same topological domain. Nanni continued: "We focused on some specific DNA sequences that encode for the **CTCF protein**." "This protein isolates portions of DNA creating barriers between the various topological domains. With the help of computer simulations and the creation of a model for classifying these proteins according to their orientation, we **identified a surprising regularity in their arrangement along the DNA sequence**." The study showed that the orientation and order of these DNA sequences makes it possible to reconstruct topological domains. The human genome compresses following a "grammar" logic comprising CTCF sequences, orientation, and the distance between them.