

*Direzione Didattica e Formazione*  
*Settore Formazione di terzo livello e relazioni internazionali*  
*Ufficio Dottorati di ricerca e scuole di specializzazione*

**SCHEDA LINEA DI RICERCA**  
**(Ex DM 1061 del 10.08.2021)**

**Corso di dottorato:**

**Tematica**

“Dottorati e contratti di ricerca su tematiche dell'innovazione”

“Dottorati su tematiche Green”

**Linea di ricerca** (coerente con le traiettorie definite dalla Strategia Nazionale di Specializzazione Intelligente (SNSI) 2014-2020, dal Programma nazionale per la ricerca (PNR) 2021-2027. (art. 3, comma 2 del DM 1061/2021).

1. Il perimetro dell'area Salute nel **SNSI** fa riferimento all'applicazione di tecnologie chiave abilitanti a produrre un miglioramento della qualità della vita. Il cambio demografico, l'invecchiamento della popolazione, l'aumentata incidenza di molte patologie croniche degenerative costituiscono problematiche rilevanti per la sanità pubblica. Il progetto si propone di sviluppare pipeline per studiare un'epigenetica green con la quale si possa poi intervenire e prevenire stili di vita e comportamenti che possano determinare l'insorgenza di patologie non geneticamente determinate. Questo limiterebbe anche la spesa sanitaria dovuta alla cura di patologie croniche come l'obesità, la demenza e le malattie cardiovascolari strettamente correlate a variazioni epigenetiche. Il progetto si integra molto bene anche con altre aree del SNSI quali l'“Agrifood” e la “Chimica verde”.

2. Per quanto riguarda il perimetro del **PRN 2021-2027** il progetto è in linea con l'Articolazione 5 (Valutazione dell'impatto dell'ambiente sugli outcomes di patologie acute e cronicodegenerative) e con l'Articolazione 7 (Promozione della salute, prevenzione delle malattie e accesso al Servizio Sanitario Nazionale). Lo stato di salute nel Paese è minacciato dal progressivo invecchiamento della popolazione, dall'aumento dei fattori di rischio comportamentali (fumo, abuso di alcool) che producono malattie cronicodegenerative a carico degli apparati cardiovascolare e del sistema nervoso centrale e del metabolismo. Recentemente il quadro è diventato ancora più articolato e complesso a causa della crisi dovuta alla pandemia del SARS-CoV-2. Non è sufficiente studiare i meccanismi patogenetici ma è necessario studiare le interconnessioni tra le patologie, gli stili di vita e l'ambiente per poter affrontare non solo le domande cliniche ma anche quelle sulle conseguenze economiche, sociali e ambientali che si ripercuotono sulla salute. Il progetto EpiGreen produrrà dati sul rapporto tra epigenetica e

patologie, stili di vita, dieta e comportamento.

**Docente referente della linea di ricerca: Prof. Gianvincenzo Zuccotti**

**Impresa**

Denominazione: A<sup>3</sup>Cube

N. mesi da svolgere in impresa (minimo 6 mesi massimo 12 mesi): 6 mesi

Riferimento aziendale (numero telefono e indirizzo e-mail):

**Periodo di studio e di ricerca all'estero (facoltativo)**

N. mesi da svolgere all'estero (minimo 6 mesi massimo 12 mesi): 6 mesi

Indicare se già disponibile la sede di soggiorno all'estero: A<sup>3</sup>Cube Sede di San Jose - California

**Titolo:** Influenza dell'ambiente e dello stile di vita sull'epigenetica: nuovi approcci di studio dell'Epigenoma nell'era del green (**EpiGreen**)

**-Title:** Influence of the environment and lifestyle on epigenetics: new approaches to studying the Epigenome in the green period

**Target (bando):** La piattaforma **EpiGreen** si innesta nel target dello studio della Biodiversità. Infatti l'epigenetica attualmente risulta essere il meccanismo cellulare e dell'organismo che maggiormente spiega la biodiversità. Gemelli omozigoti aventi la stessa sequenza di DNA ma sottoposti a stili di vita, diete e ambienti differenti possono sviluppare patologie differenti ed addirittura modalità di invecchiamento diverse. Lo studio della biodiversità su base epigenetica è attualmente una dei settori più innovativi della genomica ed è molto utilizzato per lo sviluppo di una società green in termini di ambiente, stile di vita e dieta.

**Descrizione (3000 caratteri minimo) in italiano**  
(parole chiave: epigenoma, stili di vita, ambiente)

### **Background**

Vari fattori come l'ambiente, la dieta e lo stile di vita possono influenzare l'epigenoma. L'importanza dell'epigenetica nelle risposte ambientali è ben consolidata nelle piante ma ultimamente stanno emergendo sempre più dati del rapporto tra variazioni epigenetiche e stimoli ambientali in molte specie animali e nell'uomo. Nell'uomo è diventata evidente l'influenza dell'ambiente nello sviluppo e nella fisiologia delle cellule umane mettendo in luce il rapporto tra genoma ed epigenoma. Le interazioni "gene x ambiente" determinano come gli individui con lo stesso risponderanno alla variazione ambientale: ammalandosi o restando sani, invecchiando fisiologicamente o con segni di fragilità così come i diversi genotipi possano essere o fattori di rischio oppure protettivi per malattie. Il fatto che lo stesso DNA possa corrispondere a diversi fenotipi ereditabili è stato ora presentato come prova che la sequenza del DNA non determina il destino di un individuo, ma grande importanza assume la struttura epigenetica del DNA. I fenotipi dipendono quindi da combinazioni specifiche di composizione del genoma, componenti epigenetiche variabili da input ambientali.

L'epigenetica riguarda cambiamenti che si verificano a vari livelli sul DNA e la cromatina senza modificare la sequenza del DNA stesso: livelli ridotti degli istoni centrali, modelli alterati di modificazioni post-traduzionali dell'istone, metilazione del DNA, sostituzione di istoni canonici con varianti di istoni ed espressione alterata degli RNA (long non coding RNA e Small RNA) non codificante capaci di regolare la trascrizione.

Molte delle modificazioni epigenetiche del DNA e della cromatina sono causate da prodotti metabolici cellulari. Ad esempio, la S-adenosilmetionina (SAM) è il donatore di metile nel DNA e nella metilazione dell'istone; il folato e le vitamine B6 e B12 inducono la produzione di SAM; L' $\alpha$ -chetoglutarato ( $\alpha$ KG) è necessario per la demetilazione del DNA e dell'istone; succinato e fumarato inibiscono il DNA e le demetilasi dell'istone; l'acetil-coenzima A è il donatore di acetile per l'acetilazione dell'istone; Il  $\beta$ -idrossibutirrato inibisce le deacetilasi istoniche di classe I; e il rapporto NAD<sup>+</sup>/NADH regola le sirtuine (istone deacetilasi di classe III). Per questo alterazioni metaboliche possono indurre variazioni dell'epigenoma fissando a livello del DNA una possibile modificazione (Intlekofer A. et al. 2018; Ferrari A. et al. 2019). Inoltre, i cambiamenti metabolici cellulari sono in grado di influenzare la regolazione di loci specifici del DNA variando così il trascrittoma cellulare ed inducendo RNA regolatori (long coding RNA o Small RNA). E' noto ormai che modificazioni epigenetiche possono fissarsi nel DNA definitivamente o avere una di lunga durata dando origine all'ereditarietà epigenetica intergenerazionale (Lu T. et al. 2018; Öst, A. et al. 2014; Sharma U. et al. 2016).

L'epigenoma di un individuo può variare per differenti interazioni con l'ambiente, con lo stile di vita e con la dieta. I risultati dell'epigenetica ambientale hanno gettato nuova luce sui ruoli delle regolazioni epigenetiche nelle tossicità indotte dall'esposizione ambientale o nella suscettibilità alle malattie. Attualmente, i contaminanti ambientali emergenti (CE) sono oggetto di ulteriori indagini a causa delle prove dell'esposizione umana oltre ai loro eventi ambientali. Tuttavia, gli effetti negativi di queste CE

ambientali sulla salute attraverso meccanismi epigenetici sono ancora poco affrontati in molti aspetti. La recente scoperta che l'istone demetilasi KDM5A e KDM6A (noto anche come UTX) possa rilevare le concentrazioni di ossigeno e quindi modulare i livelli di H3K4me3 e H3K27me3 è un primo passo in questa direzione (Batie M. et al. 2019; Chakraborty A. et al. 2019).

Per quanto riguarda la variazione epigenetica in relazione alla dieta si è dimostrato che una dieta di restrizione proteica nei topi può indurre metilazione del DNA e la repressione di un cluster di geni del DNA ribosomiale (rDNA), sebbene il meccanismo e i responsabili di questo effetto restino da identificare (Holland M. L. et al. 2016). Anche nutrienti alimentari come le vitamine possono regolare i processi sia fisiologici che patologici attraverso il loro impatto diretto sull'epigenoma. La vitamina A agisce come un importante regolatore dei meccanismi epigenetici. Anche le vitamine del gruppo B, tra cui biotina, niacina e acido pantotenico, partecipano alla modulazione di vari epigenomi agendo sulla metilazione del DNA. Inoltre, la vitamina C ha dimostrato di modulare sia la metilazione del DNA che le modificazioni istoniche, ed alcuni dati hanno anche supportato il suo ruolo nelle vie mediate dai miRNA. Allo stesso modo, la vitamina D influenza anche varie modificazioni epigenetiche sia del DNA che dell'istone controllando i meccanismi di regolazione. Le Vitamine stanno assumendo sempre più un ruolo come modificatori epigenetici mettendo in evidenza le prove chiave per lo sviluppo di vitamine come potenziali epifarmaci (Suza M. et al. 2021).

Anche gli stili di vita e le abitudini come fumo ed alcol, stress, ed esercizio fisico sono dei modulatori epigenetici (Kremens K. Et al. 2012). L'invecchiamento è correlato a cambiamenti epigenetici quale l'accessibilità alterata al materiale genetico, che porta a un'espressione genica aberrante, alla riattivazione degli elementi trasponibili e all'instabilità genomica. Sorprendentemente, alcuni tipi di informazioni epigenetiche possono funzionare in modo transgenerazionale per influenzare la durata della vita della prole.

Da questi studi emergono diverse importanti conclusioni: piuttosto che essere geneticamente predeterminate, la nostra durata della vita è in gran parte determinata epigeneticamente dalla dieta e altre influenze ambientali possono influenzare la nostra vita modificando le informazioni epigenetiche (Sangita P. et al. 2016).

Nonostante le informazioni che vari fattori ambientali possono modulare l'epigenoma, i meccanismi dettagliati delle regolazioni epigenetiche mediati dall'ambiente, dalla dieta e dagli stili di vita sono ancora del tutto ignoti e quindi sono necessari studi dettagliati per decifrare il loro ruolo a livello dell'epigenoma sia nello sviluppo fisiologico sia in relazione a patologie. Infatti, gli effetti confondenti, l'impatto dell'esposizione multifattoriale, l'accesso a tessuti appropriati e la valutazione della causalità per la sequenza del DNA rispetto alla variazione epigenetica rimangono sfide importanti, in particolare nell'uomo. C'è un urgente bisogno di identificare i collegamenti diretti tra i cambiamenti ambientali, i cambiamenti metabolici e le componenti epigenetiche per poter identificare e quantificare il ruolo protettivo di una green life.

Obiettivo finale del progetto sarà studiare l'epigenoma (metilazione del DNA, modificazioni dell'istoni, espressione di RNA non codificante) per indagare l'ambiente e gli stili di vita sugli esiti della salute attraverso una piattaforma di intelligenza artificiale. Il progetto produrrà una grande quantità di dati inerenti alle variazioni epigenetiche analizzabili con sistemi di intelligenza artificiale per evidenziare connessioni con gli stili di vita. Saranno definite popolazioni strettamente caratterizzate da un punto di vista di stili di vita, dieta, abitudini, ambienti frequentati ed età da indagare per metilazione del DNA, variazioni post-traduzionali degli istoni ed espressione degli RNA non codificanti. I processi di intelligenza artificiale saranno in grado di definire pattern di variazioni correlandole alle caratteristiche di popolazione descritte. Tale progetto sarà effettuato con le nuove tecnologie per lo studio delle variazioni epigenetiche quali piattaforme di Next Generation Sequencing, Alpha express technology ed intelligenza artificiale. Il progetto produrrà prodotti innovativi:

- relativi a nuove conoscenze di come i cambiamenti della società possano interagire con il genoma umano
- relativi a processi innovativi di analisi (pipeline) dei dati prodotti dalle tecnologie per la genomica quali il Next generation sequencing
- relativi a nuove applicazioni e servizi da utilizzare nel mondo della sanità per una nuova frontiera di diagnosi e cura utilizzando le variazioni epigenetiche come biomarcatori predittivi di patologia.

- relativi a nuovi prodotti poiché la conoscenza delle variazioni epigenetiche relative all'azione di nutrienti potranno portare a nuove classi di farmaci definite "epidrug"

### **Disegno Sperimentale**

L'obiettivo finale sarà raggiunto secondo i seguenti Work Packages:

**WP1.** Definizione e raccolta delle popolazioni studiate.

Il progetto prenderà in considerazione differenti popolazioni che copriranno differenti età, diverse localizzazioni geografiche di abitazione e differenti stili di vita

Neonati sani: in collaborazione con il Laboratorio di Screening di Regione Lombardia saranno selezionati 100 neonati che non hanno presentato "anomalie" metaboliche allo Screening Neonatale Esteso

Adolescenti: 100 adolescenti sani (tra 12-15 anni di età) in collaborazione con la Pediatria dell'Istituto Buzzi e alcune scuole superiori milanesi

Popolazione anziana (tra 70-80 anni di età): 100 persone identificate come sane raccolte in collaborazione con il Centro per lo studio sull'Invecchiamento (Fondazione Golgi-Cenci; Abbiategrasso)

Popolazione Ultra-centenaria (oltre i 100 anni): 50 persone ultracentenarie sane. In collaborazione con il Centro per lo studio sull'Invecchiamento (Fondazione Golgi-Cenci)

Dopo aver ottenuto l'approvazione del Comitato Etico e aver definito i Consensi informati adeguati un questionario inerente agli stili di vita (utilizzo di alcol, fumo), la dieta (utilizzo di frutta, verdura, the, caffè) e l'esercizio fisico sarà sottoposto a tutte le persone adulte ed agli adolescenti. A tutti i partecipanti sarà richiesto un prelievo di sangue venoso per le analisi epigenetiche.

Le analisi epigenetiche per i neonati saranno messe a punto sul cartoncino di Guthrie favorendo lo sviluppo di nuovi protocolli di analisi che permettano di abbandonare il prelievo venoso. Anche per i neonati sarà prodotto un consenso ed un'informativa da somministrare ai genitori.

**WP2.** Studio delle variazioni epigenetiche

Le variazioni epigenetiche saranno considerate a più livelli:

Task 2.a: Trascrittoma ed analisi degli RNA non codificanti: L'RNA cellulare sarà estratto da cellule mononucleate del sangue utilizzando il RNeasy Mini Kit (Qiagen, Hilden, Germany) secondo il protocollo del produttore e i campioni sono quindi trattati con il RNase-Free DNase Set (Qiagen, Hilden, Germany). La concentrazione e la qualità dell'RNA estratto saranno determinate con spettrofotometro NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE, USA) e l'integrità dell'RNA è stata verificata tramite Tape station (Agilent). Le librerie RNA-seq sono state preparate con il CORALL Total RNA-Seq Library Prep Kit (Lexogen, Vienna, Austria) utilizzando 150 ng di RNA totali. Il RiboCop rRNA Depletion Kit (Lexogen, Vienna, Austria) sarà utilizzato per rimuovere l'rRNA. Le qualità delle librerie di sequenziamento saranno valutate con il saggio D1000 ScreenTape utilizzando il sistema TapeStation 4200 (Agilent, Santa Clara, CA, USA) e quantificate con il kit Qubit™ dsDNA HS Assay (Invitrogen, Carlsbad, CA, USA). L'elaborazione dell'RNA sarà eseguita utilizzando il sequenziamento NextSeq 500 Illumina. I file FastQ saranno generati tramite Illumina bcl2fastq2 (v. 2.17.1.14; <https://support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html>, ultimo accesso il 15 febbraio 2021) a partire da letture di sequenziamento non elaborate prodotte dal sequenziatore NextSeq Illumina. La qualità delle singole sequenze sarà valutata utilizzando il software FastQC dopo il taglio dell'adattatore con il software cutadapt. Le intensità dei geni e dei trascritti sono state calcolate utilizzando il software STAR/RSEM (Li B. et al. 2011), utilizzando Gencode Release h38 (GRCh38) come riferimento, utilizzando l'opzione "-strandness forward". L'abbondanza della trascrizione è stata ottenuta utilizzando la piattaforma di genomica BlueBee® (Lexogen, Vienna, Austria).

Task 2.b: Metilazione delle isole CpG del DNA: Nell'era dell'innovazione del sequenziamento di nuova generazione (NGS), questa metodologia può essere utilizzata per lo studio della metilazione del DNA su un intero genoma. Il DNA viene trattato mediante un protocollo di bisulfito e quindi avviene l'arricchimento per regioni di interesse o regioni genomiche contenenti CpG sono eseguite NGS. Tale arricchimento è ottenuto mediante ibridazione con oligonucleotidi immobilizzati utilizzando SureSelectXT Methyl-Seq (Agilent). Dopo la conversione del bisulfito utilizzando il sistema SeqCap Epi di Roche, è possibile eseguire l'ibridazione per l'arricchimento.

Task 3.c: variazioni post-traduzionali degli istoni: modificazioni dell'istone, ovvero acetilazione, metilazione e fosforilazione saranno analizzate mediante i saggi AlphaLISA (PerkinElmer).

### **WP3. Analisi informatica**

Task 3.a L'analisi dell'espressione differenziale per l'RNA sarà eseguita utilizzando il pacchetto R DESeq2 (Leng N. et al. 2013). I geni saranno considerati differenzialmente espressi con  $|\log_2(\text{campione di condizioni}/\text{campione di controllo})| \geq 1$  e un tasso di false scoperte (FDR)  $\leq 0,1$ . Si imporrà il minimo  $|\log_2FC|$  di 1 e un FDR inferiore a 0,1 come soglie per geni differenzialmente espressi. I dati grezzi ottenuti dall'analisi RNA-seq sono depositati in repository di Gene Expression. L'analisi dell'arricchimento genico è stata eseguita sui geni codificanti. Analisi del percorso dell'Enciclopedia dei geni e dei genomi (KEGG) di Kyoto (<http://www.genome.ad.jp/kegg>, ultimo accesso il 15 febbraio 2021) e analisi di WikiPathways (<https://www.wikipathways.org/index.php/WikiPathways>) di geni codificanti espressi in modo differenziale tramite lo strumento web ArricchirR. Inoltre, è sarà eseguita l'analisi di Gene Ontology (GO) per i processi biologici, i componenti cellulari e la funzione molecolare. Il software R sarà utilizzato per generare mappe di calore (funzione heatmap.2 dal pacchetto R ggplots), plot PCA (funzione prcomp dal pacchetto R ggplots), grafici vulcano, grafici dotplot (libreria ggplot2) e grafici Pathview (libreria Pathview).

Task3.b Attualmente sono disponibili poche pipeline per l'analisi dei dati di metilomica poiché la tecnologia è molto innovativa. Il progetto EpiGreen avrà anche il compito di sviluppare nuove pipeline in accordo con le tecnologie ed i protocolli utilizzati per l'analisi della metilazione del DNA.

### **WP4. Processi di analisi di intelligenza artificiale (in collaborazione con l'azienda A3Cube)**

L'azienda A3Cube sviluppa una nuova architettura denominata "Hardware accelerated Software Defined System" (HSDS™), attraverso sistemi "ultra"- computing per potenziare in un'unica piattaforma integrata e bilanciata i livelli di calcolo, di stoccaggio e di comunicazione.

La disponibilità di una grande quantità di dati prodotte dal progetto EpiGreen e la capacità di organizzarli, esaminarli e connetterli sarà ottenuto mediante tecniche di intelligenza artificiale che possono essere applicate direttamente nel campo delle scienze "omiche" (genomica, trascrittomica, proteomica...). L'utilizzo in specifico dell'intelligenza artificiale in ambito genomico può dare origine a due importanti risultati che a loro volta possono aprire nuove strade alla medicina ed alla farmaceutica in generale:

- 1) Generare nuove informazioni da set di dati su larga scala, migliorando la comprensione della variazione epigenomica in relazione a salute e malattia
- 2) Semplificare i problemi analitici chiave nell'analisi epigenomica, aiutando a focalizzare la ricerca di varianti che causano malattie e riducendo i tempi di analisi clinica

Bisogna infine tenere presente che quasi ogni fase della "pipeline" dei dati epigenomici può trarre vantaggio dagli sviluppi e dalla implementazione dell'intelligenza artificiale, facilitando l'analisi di set di dati grandi e complessi, attraverso l'apprendimento automatico delle relazioni in essi contenute che avrà come effetto fondamentale un'accelerazione di nuove scoperte nella medicina genomica.

**- Lettera di intenti dell'azienda per 1) ospitare, 2) eventualmente cofinanziare il dottorando target: nostri uffici**

### **- Descrizione del progetto da pubblicare sul sito (Dottorato o Ateneo) in Italiano e inglese**

Il cibo, l'aria, l'acqua, le sostanze chimiche, i farmaci, l'inquinamento, i campi elettromagnetici, l'energia che non vediamo come le onde, i pensieri stessi, le nostre scelte possono influenzare il nostro fenotipo senza alterare il genotipo agendo a livello epigenetico. Le modificazioni epigenetiche fanno parte del normale funzionamento delle cellule, sono coinvolte nei processi di differenziazione cellulare, che avviene durante lo sviluppo embrionale e finisce poco prima della nascita. Con la stessa ottica possiamo guardare le fasi di accrescimento, di mantenimento e d'invecchiamento, dove, però le modificazioni epigenetiche diventano in gran parte tessuto-specifiche. Perfino il pattern di invecchiamento e di patologia è definito dalle modificazioni epigenetiche che avvengono nel corso della nostra vita. In un'era di grandi cambiamenti economico-sociali e climatici, in cui vengono drasticamente modificati i sistemi ambientali, gli stili di vita e la variabilità alimentare occorre comprendere in modo preciso ed accurato quali sono le basi e le correlazioni tra questi fattori e l'epigenetica. Questo permetterà di poter agire sull'epigenoma utilizzando stili di vita, ambienti e diete che possiamo definire "green" per il loro effetto protettivo sulla vita dell'individuo. Il progetto **EpiGreen** si

propone definire **un'epigenetica green** quale strumento di prevenzione, raccogliendo dati di popolazioni strettamente selezionate, utilizzando tecnologie genomiche innovative ed analizzati con sistemi di intelligenza artificiale.

Food, air, water, chemicals, drugs, pollution, thoughts themselves, our choices can influence our phenotype without altering the genotype by acting on an epigenetic level.

Epigenetic modifications are part of the normal functioning of cells: they are involved in cell differentiation processes, from embryonic development to aging. With the same perspective, we can look at the growth, maintenance and aging phases of an individual, where epigenetic modifications become largely tissue-specific. Even the pattern of aging and pathology can be defined by the epigenetic changes that occur throughout our life. In an era of great economic, social and climatic changes, in which environmental systems, lifestyles and nutrition are drastically modified, it is necessary to understand precisely and accurately what are the bases and the correlations between these factors and the epigenetics. This will allow us to act on the epigenome using lifestyles, environments and diets that we can define as "green" due to their protective effect on the life of the individual. The **EpiGreen** project aims to define green epigenetics as a prevention tool, collecting data from strictly selected populations, using innovative genomic technologies and analyzed with artificial intelligence systems

**Target candidato:** Il candidato ideale dovrebbe avere conoscenze di base nel campo dell'epigenetica con particolare interesse per la trascrittomico e la metilomica. Lo studio della metilazione del DNA e delle variazioni introniche nonché l'analisi di trascritti quali mRNAs, lncRNAs e microRNAs saranno parte integrante del progetto. Sono necessarie conoscenze tecniche nell'ambito delle tecnologie di base di biologia molecolare e genetica. Particolarmente apprezzate saranno conoscenze nel campo del sequenziamento massivo parallelo, della bioinformatica e di interazione trascrittoma-epigenoma.