



*Grazie a un approccio innovativo all'analisi dei Big Data, un gruppo di ricercatori del Centro della Complessità e dei Biosistemi (CC&B) dell'Università di Milano è riuscito a identificare una sorta di firma genetica comune a obesità, cancro alla mammella e diabete*



Milano, 10 luglio 2017 - L'obesità è sempre più diffusa nel mondo, come dimostrano i dati secondo i quali circa il 10% dei bambini statunitensi ed europei sono obesi o sovrappeso. Dal punto di vista sanitario, l'obesità sta superando il fumo come principale causa di morte prematura; è, infatti, responsabile di più del 70% dei casi di diabete ed è stata associata ad alcuni tipi di tumore, fra cui quello alla mammella.

Prove cliniche ed epidemiologiche hanno evidenziato il legame fra obesità, cancro alla mammella e diabete, ma non era ancora stata ottenuta una solida conferma di questa relazione a livello di espressione genetica. Ciò è dovuto a una serie di fattori, dalla grande variabilità fra i pazienti ai limiti dei modelli di studio in vitro.

Ma soprattutto, c'è anche un grosso problema relativo ai dati: la presenza di una grande quantità di rumore di fondo, che rende difficile individuare alcuni elementi ricorrenti nei risultati delle analisi di migliaia di geni in molti individui diversi.

“Nel campo biomedico vengono effettuati moltissimi esperimenti, grazie ai quali è stato possibile raccogliere ingenti quantità di dati biologici in diversi database pubblici”, dice Caterina La Porta, membro del CC&B e professoressa di patologia generale al Dipartimento di scienze e politiche ambientali dell'Università di Milano.

“Combinare set di dati provenienti da studi diversi sarebbe molto utile per ottenere informazioni sempre più accurate e rilevanti, ma ciò comporta anche un problema, chiamato batch effect – spiega La Porta, che

ha coordinato la ricerca appena pubblicata su NPJ Systems Biology and Applications – I dati provenienti da ciascun esperimento sono infatti condizionati da cause tecniche che non hanno a che fare con i fattori biologici. Questo genera un rumore di fondo che può mascherare alcune differenze importanti dal punto di vista biologico quando si confrontano campioni appartenenti a lotti diversi”.

Un problema che i ricercatori coordinati da La Porta hanno cercato di mitigare usando un nuovo approccio, basato sulla combinazione di due diverse tecniche chiamate decomposizione ai valori singolari e analisi di deregolazione dei pathway.

In questo modo sono riusciti a individuare 38 geni che sono espressi in maniera diversa negli adipociti provenienti da soggetti obesi, confrontati con quelli provenienti da soggetti non obesi. Una sorta di firma genetica che sembra caratterizzare in maniera specifica la condizione di obesità, indipendentemente dal genere del soggetto.

Questi geni sono soprattutto associati a processi di infiammazione e risposta immunitaria, e a complicazioni note dell'obesità come il diabete di tipo 2 e l'infertilità. Essi, inoltre, sono deregolati in maniera simile nel caso di cancro alla mammella, il che sembra quindi confermare l'associazione fra questo tipo di tumore e l'obesità. Alcuni di questi geni potrebbero quindi rappresentare degli interessanti marcatori biologici, utili non solo per ulteriori ricerche su questi temi, ma, eventualmente, anche per possibili scopi diagnostici.

“La forza del nostro lavoro deriva dall'uso di metodi di filtraggio e riduzione del rumore particolarmente appropriati, grazie ai quali siamo riusciti a mitigare il batch effect. Questa strategia di analisi potrebbe venir utilizzata anche per studiare altre patologie, consentendo di sfruttare con maggior accuratezza l'enorme quantità di dati accumulati nella letteratura biomedica – conclude La Porta – Grazie a questo approccio, siamo riusciti a identificare una lista di geni caratteristici dell'obesità, che sono anche associati al diabete di tipo 2 e al cancro alla mammella. Il tutto con un grado di precisione simile a quello usato per identificare il Bosone di Higgs”.