Qual è la relazione fra un gene e i tumori? Capire quanto un oncogene sia legato all'insorgenza del cancro è la missione di OncoScore, l'innovativo strumento informatico in grado di classificare i geni in base alla loro associazione con il cancro e di aggiornarsi continuamente da solo sulla base della letteratura biomedica. Il nuovo strumento per la ricerca sui tumori è stato sviluppato da un gruppo di ricerca guidato dal professor Carlo Gambacorti-Passerini, docente di Malattie del sangue all'Università di Milano-Bicocca.

L'articolo è stato pubblicato sulla rivista *Nature Scientific Reports* (R. Piazza, D. Ramazzotti, R. Spinelli, A. Pirola, L. De Sano, P. Ferrari, V. Magistroni, N. Cordani, N. Sharma, C. Gambacorti-Passerini, *OncoScore: a novel, Internet-based tool to assess the oncogenic potential of genes* ).

Il database di riferimento scandagliato da OncoScore è PubMed , costituito da circa 27 milioni di citazioni di ricerche scientifiche registrate dal 1949 ad oggi. Il panorama in costante evoluzione delle mutazioni tumorali rappresenta una sfida estremamente impegnativa per identificare i geni cancerogeni e la capacità di dare un ordine di priorità a queste variabili diventa quindi di fondamentale importanza: è proprio per risolvere questo problema che è stato sviluppato OncoScore.

Un oncogene è un gene o una serie di nucleotidi che codificano una proteina, che aumenta le possibilità di sviluppo in senso tumorale delle cellule. OncoScore è in grado di includere nelle ricerche anche i nuovi oncogeni e attribuisce – utilizzando i big data – un punteggio ai geni, distinguendo fra gli oltre 20 mila geni che compongono il genoma umano il 70 per cento che non ha una relazione conosciuta con l'insorgenza del cancro dal restante 30 per cento che invece può legarsi a questo tipo di malattie. L'indice costruito dal software è basato sul grado di oncogenicità e permette di scoprire immediatamente, sulla base delle ricerche scientifiche pubblicate fino ad oggi, la "pericolosità" di un determinato agente.

Le basi sequenziate dei genomi vanno da 20 milioni a 3 miliardi per ogni paziente analizzato: un aggiornamento manuale sarebbe quindi impensabile. Attualmente, l'aggiornamento automatico per includere tutte le ricerche più recenti incluse da PubMed è attivato ogni sei mesi e impiega circa un mese per completarsi, ma è possibile anche un aggiornamento su base mensile. Questo meccanismo, secondo i ricercatori, permette di ottenere serie storiche con una precisione e un livello di attualizzazione dei dati senza precedenti.

Se un fattore, un gene, è citato spesso nelle ricerche scientifiche ma non ha una relazione forte e riconosciuta con il cancro, OncoScore è in grado di riconoscerlo e gli attribuirà quindi un punteggio basso – anche se compare con una frequenza notevole – probabilmente inferiore alla soglia di cut-off (pari a 21.09). Un esempio potrebbe essere quello di GUSB, un gene spesso usato come elemento di controllo che ha un OncoScore di 20. Noti oncogeni, invece, raggiungono valori che con questo indice superano anche quota 100.

«L'utilizzo di OncoScore, lungi dall'essere uno strumento diagnostico – spiega Carlo Gambacorti-Passerini, professore all'Università di Milano-Bicocca – è molto utile per analizzare velocemente lunghi elenchi di geni come quelli presenti in cromosomi, o parti di essi, che

vengono persi o al contrario amplificati in diversi tumori, oppure per distinguere mutazioni “passenger” e “driver” fra quelle identificate in un determinato genoma tumorale».