Valeria Destro, 325939

Titolo tesi: LA METILAZIONE DEL DNA DA TAMPONI BUCCALI COME NUOVO BIOMARCATORE NON-INVASIVO PER IL DIABETE MELLITO GESTAZIONALE (GDM) IN DONNE INCINTE

Relatore: Prof. Marco Morselli

Tirocinio presso il dipartimento di Biochimica e Biologia Molecolare dall'università di Parma

Secondo le stime della Fondazione internazionale del diabete (IDF) il 14 % delle donne in gravidanza sviluppa il diabete mellito gestazionale (GDM), ossia all'incirca 20 milioni di nascite all'anno. Esso viene diagnosticato con il test OGTT, la curva di tolleranza al glucosio, essenziale per apportare cambiamenti nello stile di vita della gestante per evitare l'insorgenza di complicazioni sia nel bambino che nella madre. Tuttavia, questo metodo di diagnosi risulta molto invasivo e stressante per le donne incinte, motivo per il quale negli ultimi anni si è cercato di sviluppare nuovi metodi non invasivi per la identificazione del GDM. In questo lavoro di tesi si attua uno studio pilota per la ricerca di nuovi biomarcatori non invasivi per diagnosticare il GDM, sfruttando la metilazione del DNA estratto da tamponi buccali. Questo studio, in collaborazione con l'Università della California – Los Angeles (UCLA) ha identificato regioni differenzialmente metilate (DMRs) tra donne con e senza il GDM. L'analisi funzionale di queste regioni ha portato all'identificazione di geni associati, come TNXIP precedentemente correlati al diabete insulino-resistente, e a fattori trascrizionali, come Myc e Max noti regolatori di geni coinvolti nel metabolismo del glucosio. Le citosine comuni a tutti i campioni sono state poi utilizzate per la creazione di 5 modelli prediditivi di classificazione binaria con repeated LOOCV (leave-one-out cross-validation). Nonostante il numero di campioni sia insufficiente per la costruzione di modelli generalizzabili a dataset esterni si è osservato come i risultati migliori siano stati riscontrati con il decision tree e il SVM (support vector machine). Buoni risultati sono stati ottenuti anche con le due metodologie di regressione logistica (con e senza trasformazione dei dati in Principal Components). Il modello con le peggiori statistiche è il random forest, che richiede sicuramente una migliore ottimizzazione prima di essere definitivamente scartato. Un'analisi di intersezione fra le citosine usate nella costruzione dei modelli e le regioni identificate nell'analisi DMR ha evidenziato 6 geni associati comuni a tutti i metodi, come SORBS1, gene associato con lo sviluppo della insulino-resistenza.