

Visto e sentito a ASHG 2021

Alessandra Renieri & Team Genetica Medica Siena

A causa delle differenze in LD (Linkage Disequilibrium) non è possibile trasferire un PRS (Polygenic Risk Score) calcolato sulla popolazione Europea a quella Africana.

L'accuratezza di predizione crolla a meno del 25% rispetto a quella europea di partenza (al 60% rispetto a quella americana e a quella dell'Asia del sud e al 50% a quella dell'Asia dell'est). Vi sono differenze di PRS anche all'interno dell'Africa stessa, come ad esempio tra l'Africa dell'ovest (Burkina faso), Africa dell'est (Kenya) e Africa del sud.

Un'analisi genetica su 50.000 soggetti della UK biobank ha identificato non solo varianti di rischio associate a malattie complesse, ma anche varianti protettive (di potenziale uso terapeutico), ad esempio per vene varicose la p.Pro2510Leu in *PIEZO1* (OR=0.6), un meccanosensore che determina la formazione di valvole in vasi linfatici e vene.

Sempre dalle analisi di UK biobank si confermano molti geni protettivi per tratti quantitativi (es. *SLC22A12* -> diminuzione urato) e corrispondente malattia (gotta) e si identificano 2 nuovi geni associati all'ipertensione (*SLC9A3R2* -> diminuzione SPB, proteina renale scaffold per sodio -> protezione da ipertensione) e diabete (*MAP3K15* -> diminuzione dell'emoglobina glicata -> protezione da diabete tipo 2)

Utilizzando i dati di sequenziamento dell'intero esoma (WES) provenienti da cinque diverse coorti di origine europea ed un innovativo modello di Machine Learning, è stato possibile identificare un insieme di varianti ultra-rare, rare, a bassa frequenza e comuni implicate nella variabilità clinica del COVID-19 (incluse quelle in linkage disequilibrium con i loci individuati in studi di associazione genome-wide (GWAS)). Il modello ha una buona predittività (circa il 75% di sensibilità e specificità) ed è facilmente interpretabile per la costruzione di una piattaforma di trials clinici basato su marcatori genetici.

Il cancro dell'endometrio è il tumore ginecologico più comune nei paesi sviluppati. Un studio GWAS ha identificato tre nuovi loci (7q22.1, 8q24.3 e 16q12.2) che influenzano il rischio di cancro dell'endometrio attraverso livelli alterati di ormoni sessuali.

Il carcinoma basocellulare (BCC) è uno dei tumori maligni più comuni. Una meta-analisi GWAS multietnica ha identificato 78 loci associati a BCC. I nuovi loci identificati sono implicati nella normale funzione cellulare (*SENPT7*), nell'adesione delle piastrine e nella formazione della matrice extracellulare (*ITGA2*) o nella riparazione del DNA (*ATM*).

Oltre 150 loci di suscettibilità genetica sono stati identificati per il cancro del colon-retto (CRC) attraverso uno studio GWAS. Questo ha rivelato un gran numero di presunti geni di suscettibilità per CRC, inclusi geni precedentemente non riconosciuti che forniscono nuove informazioni sui meccanismi sottostanti i loci di rischio genetico CRC.

Per quanto concerne disturbi quali Megalencefalia, Emimegalencefalia e Displasia Corticale Focale, sono state identificate molteplici mutazioni (LoF e GoF) nel pathway di *PIK3-AKT-MTOR*. Dall'analisi di una coorte di 58 pazienti, su cui sono stati eseguiti esperimenti di Droplet Digital PCR, si è raggiunta una resa diagnostica del 90% nei casi di megalencefalia (per cui si sono riscontrate per di più varianti a carico di *PIK3CA*), e del 24% nei casi di displasia corticale focale (varianti a carico di *MTOR*). Sono stati creati modelli su organoidi per studiare la patogenesi di queste malattie in vitro,

e si sono testati tre diversi inibitori del pathway di *MTOR* per valutare eventuali riduzioni dell'ipertrofia e iperplasia riscontrate nelle varie patologie.

Bisogna valutare il processo di *skewed inactivation* per guidare un'analisi più attenta degli esomi per disturbi del neurosviluppo X-linked. Questo ha portato all'identificazione di varianti in geni come *ATRX* (mental retardation-hypotonic facies syndrome), *RLIM* (Tonne-Kalsheuer syndrome), *OTUD5* (sindrome da anomalie congenite X-Linked) e *RBM10* (Tar syndrome).

È stato scoperto un nuovo disordine del neurosviluppo dato da mutazioni nel gene *ZMYND8*, caratterizzato da disabilità intellettiva da moderata a profonda, caratteristiche faciali peculiari e anomalie cardiovascolari, oftalmologiche e scheletriche (quest'ultime in misura minore). Una collaborazione internazionale ha identificato 12 varianti missenso e tre varianti di stop, di cui una ricorrente.

Un'analisi WES su trios condotta su individui con alto grado di consanguineità ha permesso di identificare nuove varianti in geni precedentemente candidati ad autismo in forma autosomica recessiva, e di ampliare la casistica dei geni candidati. Le varianti riscontrate sono 93, tra cui 57 bialleliche, 24 de novo/dominanti e le restanti segreganti in forma X-Linked. Il 52% di queste determinano perdita di funzione genica, e il 47% sono non-sinonime. Si sono identificate infine anche CNVs bialleliche in quattro trios, causative di perdita di esoni in geni quali *SEMG1*, *DNAH7*, *DHRS42L2*, *SIRPB1*.

Attraverso un attento fine mapping dei loci di suscettibilità al melanoma sono stati identificati due nuovi geni di suscettibilità tra cui uno, localizzato nella regione enhancer, che regola l'espressione di *CDKAL1*

Case report: femmina di 12 anni con una delezione de novo di 3.18 Mb in 7q36.1-q36.2 presenta clinicamente ADHD, DI, tremori, bassa statura e malformazione Chiari 1. Tale fenotipo è associato alla perdita del gene *KMT2C*.

Malattia di Hirschsprung: ruolo di SNP nella funzionalità del gene *RET* (Rearranged during Transfection). Riscontrati 38 SNP localizzati in regioni non codificanti del gene *RET*, di cui alcuni danneggiano l'attività di specifici enhancers e altri provocano alterata espressione del gene *RET*.

Il modello di previsione del rischio di cancro al seno (BC) BOADICEA è stato modificato per incorporare fattori di rischio basati su questionari (QRF), mammographic density (MD) e storia familiare (FH).

Per caratterizzare la suscettibilità genetica alle neoplasie multiple è stato condotto uno studio di sequenziamento dell'intero esoma. Sono state trovate mutazioni nei geni *PLSCR1*, *MYC* e *ACAA1*, mai implicati nello sviluppo di tumori secondari.

In 5 differenti famiglie con disordini del neurosviluppo e vari gradi di dismorfismi facciali, grazie a studi di WES, sono state identificate varianti bi-alleliche in *ATG7*, enzima chiave nel processo di autofagia; in tutti i casi tranne uno si trattava di varianti missenso.

CAPRIN1 interagisce con proteine coinvolte nel neurosviluppo, inclusa FMRP ed è emerso come possibile gene candidato per ASD. Sono stati riportati nove casi di loss-of function

in *CAPRIN1* con disturbi del linguaggio (100%), ritardo del linguaggio (89%), disabilità intellettiva (88%), ADHD (86%), ASD (68%), ritardo dello sviluppo (33%) e convulsioni (22%).

Mediante studi GWAS sono stati individuati centinaia di marcatori genetici per il diabete di tipo 2 (T2D), localizzati nella maggior parte (~ 90%) in regioni non codificanti del genoma, suggerendo che i circuiti trascrizionali disregolati sono un meccanismo primario che contribuisce all'eziologia del T2D.

E' stato sviluppato un modello statistico per le varianti introniche, di difficile interpretazione, implicate in oltre il 15% dei disordini genetici rari.

Questo modello si basa su uno score che valuta automaticamente lo splicing intronico (scores SpliceAI intronici), previsioni di patogenicità delle regioni codificanti e modelli mutazionali sottostanti per valutare il significato di varianti in eterozigosi composta osservate nei casi recessivi.

Il Genome-to-Treatment (GTRx) è un sistema virtuale automatizzato per la diagnosi di malattie genetiche e una guida alla gestione di bambini malati nelle unità di terapia intensiva. Questo sistema ha prestazioni analitiche per SNP, inserzioni-delezioni, varianti strutturali e CNV. (<http://gtrx.rbsapp.net/>).

L'ASD è una condizione che comprende una combinazione degli effetti di varianti rare e comuni, rischio poligenico e sesso. I fattori presi in considerazione sono varianti loss of function de novo (dnLoF), mutazioni missenso (dnMIS), varianti ereditarie rare (inhLoF) e scores poligenici per ASD (PSASD). È stato dimostrato un "effetto protettivo femminile" per cui le femmine tollerano un maggiore carico genetico. Allo stesso modo viene applicato alle varianti comuni e rare.

Lo sviluppo di un algoritmo e l'implementazione in un pacchetto software, chiamato Cameleolyser, consente di analizzare regioni genomiche omologhe.

I disturbi complessi, caratterizzati da fenotipi variabili, come autismo e disabilità intellettiva, sembrerebbero essere associati alla variazione del numero di copie di specifiche regioni genomiche. Inoltre, alcuni specifici quadri clinici sembrerebbero essere associati ad alterazioni di determinate regioni cromosomiche, come per esempio la delezione del cromosoma 16p12.1 che è associata a ritardo psicomotorio e autismo nei bambini, mentre negli adulti è associata a maggiore rischio di sviluppare schizofrenia e ridotto quoziente intellettivo.

Gli enzimi del metabolismo e del trasporto dei farmaci, implicati nella metabolizzazione delle molecole xenobiotiche, presentano espressività variabile sesso dipendente, che coinvolge in particolare 16 geni. Ad esempio, difetti nel gene *UGT2B17*, coinvolto nel metabolismo del testosterone, determina ipercolesterolemia nell'uomo, mentre nella donna non si evidenziano gli stessi effetti fenotipici.

È stato eseguito il sequenziamento dell'intero trascrittoma (RNA-Seq) con WGS per migliorare l'interpretazione clinica. Un nuovo metodo per migliorare la copertura dei geni a bassa espressione, è il metodo CRISPRclean che utilizza la nucleasi Cas9 e 360.000 RNA guida per rimuovere in modo specifico i frammenti della libreria RNA-Seq da oltre 4.000 geni mirati, con un aumento medio di copertura pari a sei volte.

Le variazioni strutturali (SV) hanno ruoli chiari nelle malattie e nei fenotipi. Parliament2 è un metodo

di rilevamento delle SV, che armonizza le chiamate da più software in un singolo set e mostra una specificità migliore. Il metodo di genotipizzazione articolare, muCNV, riduce i falsi positivi che si accumulano tra metodi e campioni di grandi dimensioni. La pipeline Parliament2-muCNV ha generato un catalogo completo di SV per analisi di associazione genetica.

Il metodo GIPSeqcluster consente il rilevamento di diverse condizioni fisiologiche e patologiche e apre strade alternative per l'indagine sul cell free DNA.

I Pattern dell'mRNA di *CELSR-3* in embrioni di topo hanno evidenziato uno spiccato livello di espressione nel sistema nervoso e urinario in via di sviluppo. In tutti i pazienti sono state riscontrate varianti bialleliche in *CELSR3*. *CELSR3* viene valutato, dunque, come potenziale gene alla base di anomalie congenite del rene e delle vie urinarie (CAKUT), anomalie anorettali e del SNC.

I geni metabolici dei farmaci più comuni, come *CYP3A4*, *CYP2C19*, *CYP1A2* mostrano un'espressione differenziata nei due sessi. Sono state inoltre identificate, 71 varianti a singolo nucleotide (SNP) sui geni DMET analizzati, correlati ad aspetti genotipici e fenotipici distinti nei due sessi.

E' stato applicato il modello di prevenzione e trattamento della medicina di precisione (PreEMPT) per stimare i vantaggi e il rapporto costo-efficacia dei test a cascata sui fratelli di neonati con varianti patogeniche. È stato utilizzato un pannello di 11 geni associati a sindromi che predispongono a carcinoma pediatrico (*RET*, *RB1*, *TP53*, *DICER1*, *SUFU*, *PTCH1*, *SMARCB1*, *WT1*, *APC*, *ALK* o *PHOX2B*). Se questi fratelli fossero sottoposti a sorveglianza per una diagnosi precoce, si stima una riduzione dei decessi del 52%.

Il gene *HDAC9* contiene sequenze che regolano la trascrizione di *TWIST1*. Saggi epigenetici in vitro e in vivo hanno dimostrato che varianti strutturali nel gene *HDAC9*, che interrompono la regolazione trascrizionale di *TWIST1*, determinano malformazioni cranio-facciali e degli arti.

Studi di GWAS su trios in famiglie africane con casi di labio-palatoschisi non sindromiche, hanno permesso di identificare 110 mutazioni de novo patogenetiche; queste includono nuove varianti LoF nei geni *TTN*, *MINK1* e *ARHGAP10* e varianti missenso in *DHR3*, *TULP4*, *SHH*, *TP63*, *FKBP10*, *ACAN*, *RECQL4* e *KMT2D*.

Studi GWAS hanno identificato 241 loci di rischio significativamente associati alle due forme più comuni di malattia infiammatoria intestinale (IBD), il morbo di Crohn e la colite ulcerosa. Tra questi i geni *NOD2*, *IL18RAP*, *IL23R*, *NCF4* e *IL2RA*, normalmente implicati nel corretto funzionamento dei linfociti T-reg, monociti e macrofagi.

Mutazioni loss of function di *RERE* contribuiscono allo sviluppo della palatoschisi in individui con delezioni prossimali 1p36 e NEDBEH (disturbi dello sviluppo neurologico con o senza anomalie del cervello, degli occhi o del cuore).

Una variante nel gene *HOXB13* rs77179853 (frequency=3.2%) è associata ad un aumentato rischio di cancro alla prostata negli uomini di origine africana. Gli uomini portatori di questa variante possono beneficiare di uno screening PSA.

Il WGS rappresenta l'analisi più completa per ASD, poichè è in grado di fornire nuove informazioni inerenti l'eterogeneità genetica. Sono state individuate varianti strutturali (SV) in *AUTS2*, *MBD5*, *NRXN1*, *SHANK1* e *PTCHD1-AS*, un'inversione che interrompe *SCN2A* e un riarrangiamento su *KCNQ2* caratterizzabili solo mediante il WGS.

Il database gnomAD, che contiene frequenze di popolazione per 10.850 varianti del mtDNA, sarà utile per valutare varianti cliniche del mtDNA di significato sconosciuto. Le varianti del mtDNA sono disponibili nella nuova versione di gnomAD v3.1 (gnomad.broadinstitute.org).

Studi GWAS hanno identificato nuovi loci, 7q22.1, 8q24.3 e 16q22.2, correlati ad un rischio aumentato di carcinoma dell'endometrio associati ad obesità o a livelli aumentati di testosterone. Il risultato dello studio è stato quello di identificare nuove opportunità terapeutiche con ormoni correlate alla riduzione del rischio di sviluppare un carcinoma endometriale.

Varianti gain-of-function (GoF) del gene MYCN sono causative di una nuova sindrome da Megalencefalia. È stata Individuata una nuova variante in MYCN in un feto con megalencefalia, ventricolomegalia, ipoplasia del corpo calloso e polidattilia.

La meta-analisi trans-etnica mediante GWAS associata alla regolazione della lunghezza dei telomeri (TL) ha individuato come rilevanti le cellule del sangue/immunitarie, un'interazione significativa dell'SNP a livello del gene TCL1 con l'aumentare dell'età e l'aumento della lunghezza dei telomeri con overespressione di KBTBD6.

L'scRNAseq su diversi tipi di cellule intestinali in combinazione con i risultati del GWAS, hanno permesso di associare alterazioni geniche nei linfociti Treg, nei monociti e nei macrofagi alla patogenesi della malattia infiammatoria intestinale, e identificare i geni RELA e CCR7 come effettori della malattia.

La schizofrenia è una patologia psichiatrica severa ed invalidante, ereditabile e poligenica, i cui meccanismi causativi rimangono ancora da chiarire. I geni coinvolti appartengono a pathways utili per il legame di piccole GTPasi, l'assemblaggio dei filopodi ed il processo catabolico dei lipidi cellulari, permettendo di ipotizzare cambiamenti nella funzione sinaptica, plasticità e morfologia delle spine dendritiche alla base della schizofrenia.

EZH1: gene coinvolto nel mantenimento della pluripotenza e della plasticità delle cellule staminali embrionali; risulta ampiamente espresso sia nelle cellule in divisione che nella fase post-mitotica, anche nel sistema nervoso in fase di sviluppo e dell'adulto.

Modello patogenico del CCM sporadico: mutazioni somatiche GoF di PIK3CA inducono la comparsa di DVA (Developmental Venous Anomaly). Quando co-occorrono mutazioni somatiche LoF dei geni del complesso CCM o mutazioni GoF di MAP3K3 si ha la comparsa di CCM associata a DVA, comportando così un quadro clinico più complesso.

Le delezioni nella regione 1p36 sono le delezioni telomeriche più comuni nell'uomo e sono associate ad un aumentato rischio di schisi del distretto orofacciale. Il gene RERE si trova nella regione critica prossimale a quella interessata da sindrome da delezione 1p36 e codifica per un co-regolatore del recettore nucleare. Varianti LoF di RERE contribuisce allo sviluppo della palatoschisi in individui con delezioni prossimali 1p36 ed un quadro di NEDBEH (disturbi dello sviluppo neurologico con o senza anomalie strutturali del cervello, dell'occhio o del cuore); l'espressione di RERE nelle cellule della cresta neurale sembra pertanto essere necessaria per il normale sviluppo palatale.

Il neuroblastoma (NB) è un tumore pediatrico del sistema nervoso simpato-adrenergico.

Colpisce circa 650 bambini in un anno negli USA e circa 100 negli UK. Il 50% dei casi di bambini affetti da NB hanno età inferiore ai 2 anni. Negli ultimi 15 anni con i progressi attuati dalla genetica medica mediante l'NGS sono stati identificati i geni più frequentemente mutati nei casi di NB (TP53, PALB2, KIF1B, EZH2, NF1, SDHB, VANGL1, PDE6G, NRAS, BRIP1, BRCA1, BRCA2, PTPN11, APC). Inoltre, sono state riscontrate anche CNV della regione 16p, di TP53, di BARD1 e CHEK2. In uno studio caso-controllo basato su SNP- array sono state identificate due varianti del mtDNA, rs2853493 e rs2853499, associate al rischio di neuroblastoma.

Regioni genomiche contenenti MT-ND4 e MT-ND5 sono state riportate come regioni hotspots per diversi tipi di cancro e disordini neurologici.

I casi sporadici di CCM (Cerebral Cavernous Malformations) sono causati da mutazioni somatiche nei geni KRIT1, CCM2, PDCD10 (che compongono il CCM complex) e MAP3K3 e PIK3CA. In particolare in casi di CCM occorrono multiple mutazioni somatiche in questi geni, ad esempio LoF per i geni del complesso CCM e GoF per MAP3K3 e PIK3CA.

Il sequenziamento dell'esoma in due soggetti unrelated con agenesia pancreatica e relativi genitori consanguinei ha permesso di identificare la presenza in entrambi di varianti loss of function in omozigosi a carico del gene ZNF808. Il fenotipo è limitato al pancreas, con diabete insorgenza nel periodo neonatale e basso peso alla nascita.

Il rhabdomyosarcoma (RMS) è il più comune sarcoma pediatrico dei tessuti molli. In circa l'8% dei casi vengono identificate varianti patogenetiche o probabilmente patogenetiche (P/LP) in geni noti correlati alla predisposizione di sviluppare il cancro (CPG). Dal sequenziamento dell'esoma di 615 pazienti con recente diagnosi di RMS è stato visto che la presenza di una variante P/LP era significativamente associata ad una peggiore OS (HR=1,79, P=0,04) dopo correzione per l'età alla diagnosi, lo stadio del tumore, l'istologia del tumore e etnia.

La maggior parte delle varianti missenso dei geni responsabili di sindrome di Lynch è classificata come VUS. Uno studio funzionale del gene MLH1 evidenzia che il 22% delle varianti missenso analizzate ha un effetto negativo sulla funzione della proteina e che tali varianti coinvolgono principalmente il dominio ATP-asico N-terminale e il dominio di dimerizzazione C-terminale.