



A.D. MDLXII

# **UNIVERSITÀ DEGLI STUDI DI SASSARI**

## **FACOLTÀ DI MEDICINA E CHIRURGIA**

Dipartimento di Scienze Chirurgiche, Microchirurgiche e Mediche

Direttore: Prof. Gian Vittorio Campus

Cattedra di Medicina d'Urgenza

Direttore: Prof. Nicola Glorioso

## **TESI DI LAUREA**

### **ANALISI GENOMA-WIDE BASATA SULLA ABITUDINE AL FUMO IDENTIFICA NUOVI LOCI RELATIVI ALL'OBESITÀ**

Relatore:

**Prof. Nicola Glorioso**

Correlatore:

**Dr.ssa Aurora Vecchiato**

Tesi di Laurea di:

**Amen Amer**

**ANNO ACCADEMICO 2015- 2016**

# INDICE

<b>Introduzione</b>	<b>1</b>
<b>Scopo dello studio</b>	<b>7</b>
<b>Materiali e Metodi</b>	<b>8</b>
<b>Risultati</b>	<b>20</b>
<b>Discussione</b>	<b>38</b>
<b>Conclusione</b>	<b>43</b>
<b>Bibliografia</b>	<b>44</b>

## INTRODUZIONE

L'obesità è stata definita come un'epidemia globale dall'Organizzazione mondiale della Sanità<sup>I</sup>, colpisce percentuali sempre più ampie delle popolazioni dei paesi industrializzati. Negli Stati Uniti, ove colpisce di fatto la metà della popolazione, è considerata la seconda causa di morte dopo il fumo di sigaretta<sup>II</sup>.

Lo studio oggetto di questa tesi si è data la ambiziosa finalità di determinare eventuali relazioni tra questi due fattori di rischio vascolare a livello genetico. Nello specifico, questo studio mira ad individuare se il fumo comporti delle alterazioni sui loci del genoma umano correlati con lo sviluppo dell'obesità.

Prima di trattare dei risultati emersi dallo studio, è necessario soffermarsi brevemente sulle caratteristiche d'interesse di questa patologia.

L'obesità è considerata una patologia cronica, multifattoriale, caratterizzata dall'eccesso di massa adiposa e grasso corporeo, maggiore di 35% nelle donne e di 25% negli uomini<sup>I</sup>. L'eccesso di grasso corporeo influisce negativamente sulla qualità di vita ed è associato ad un aumento del rischio di morbosità e mortalità. È ormai appurato che la frequenza dell'obesità subisce un incremento con l'avanzare dell'età, almeno fino ai 50- 60 anni<sup>III</sup>.

La malattia può essere determinata da diversi fattori: da uno squilibrio tra l'introito calorico e la spesa energetica; da quadri sindromici da mutazioni genetiche, come nella sindrome di Prader-Willy, Lauren Moon; da una predisposizione genetica in associazione con fattori ambientali e culturali; infine da fattori psicologici, come i disturbi del comportamento alimentare, il cui principale è il disturbo da alimentazione incontrollata<sup>IV</sup>. Vengono individuate anche delle cause secondarie, consistenti in traumi e lesioni a livello della regione ipotalamica, deficit di GH, ipercortisolismo e pseudoipoparatiroidismo<sup>V</sup>.

Tra i fattori ambientali possiamo identificare la dieta ad alto contenuto di lipidi, la sedentarietà, il basso livello culturale (tuttavia nei Paesi in via di sviluppo sono gli

---

individui che appartengono ai ceti sociali più elevati a presentare un tasso maggiore di obesità<sup>VI</sup>), i farmaci, come gli antidepressivi e antipsicotici od infine la sospensione del fumo.

Possiamo distinguere tre tipi di obesità: l'obesità centrale, periferica e diffusa.

Si parla di obesità centrale quando la distribuzione del grasso è prevalentemente a livello addominale e a livello viscerale intraperitoneale, altresì, quando il rapporto tra la circonferenza vita-fianchi è  $> 0.85$ .

Nell'obesità periferica (definita anche ginoide o sottocutanea) la distribuzione del grasso è soprattutto a livello sottocutaneo, tra l'ombelico e il ginocchio ed è tipico nelle donne. In questo caso il rapporto tra la circonferenza vita-fianchi è  $< 0.78$ .

Infine, l'obesità diffusa prevede un rapporto circonferenza vita-fianchi tra 0.78 e 0.85.

Per la misurazione del tessuto adiposo vengono utilizzati diversi metodi, quello più frequente (indiretto), anche se non preciso, è l'indice di massa corporea (BMI) = peso/altezza al quadrato, difatti, tale indice non dà risultati fedeli in quanto non prende in considerazione che la massa muscolare del soggetto. Precisamente, si possono classificare differenti valori. Quando l'indice è inferiore a 18,5 il soggetto è sottopeso, tra 18,5 e 24,9 ha un peso normale, tra 25 e 29,9 è sovrappeso, mentre invece se l'indice è tra 30 e 34,9 il soggetto ha un'obesità lieve, tra 35 e 39,9 un'obesità moderata e se supera i 40 è considerato obeso grave<sup>VII</sup>.

Si può optare per altri metodi, distinguibili tra: diretti, con cui i valori vengono misurati su individui considerati normali e l'obesità e/o la magrezza vengono definiti in termini statistici, oppure indiretti, i quali stimano il grasso corporeo con modalità più semplici che si correlano con i risultati ottenuti con la tecnica diretta.

Tra le tecniche dirette distinguiamo: la densitometria, la tomografia computerizzata, la risonanza magnetica (RM), la impedenzometria. Una tecnica indiretta utilizzata è, invece, la plicometria cutanea, la quale misura le pliche cutanee a livello della coscia, del bicipite, petto, ventre e della zona sottoscapolare<sup>VIII</sup>.

L'obesità ha importanti effetti negativi sulla salute, comportando un aumento del rischio di patologie sia fisiche che mentali<sup>IX</sup>, identificate come sindrome metabolica. Detta sindrome, precedentemente nota come sindrome X o sindrome da insulino-resistenza, si caratterizza per la simultanea presenza di fattori metabolici che aumentano il rischio di malattie cardiovascolari. Viene utilizzata la classificazione dell'ATPIII. Per la diagnosi è necessaria la presenza di almeno tre dei seguenti componenti:

1. circonferenza vita >102 cm nell'uomo e >88 cm nella donna
2. HDL <40 mg/dl nell'uomo e <50 mg/dl nella donna
3. pressione arteriosa >130/85 mmHg
4. trigliceridi >150 mg/dl
5. glicemia >110 mg/dl<sup>X</sup>

Dall'obesità possono derivare combinazioni di disturbi, come: il diabete mellito di tipo II, che altera la risposta del corpo all'insulina, portando ad uno stato di insulinoresistenza; l'ipertensione arteriosa, è stato dimostrato che vi è una correlazione positiva tra questa e il peso corporeo, viceversa una buona parte dei soggetti ipertesi presenta iperinsulinemia e insulinoresistenza, le patologie osteoarticolari, il carcinoma dell'endometrio, le patologie respiratorie e quelle cardiovascolari<sup>XI</sup>.

I fattori di rischio di queste ultime si ritiene che derivino da errate abitudini e stili di vita, in particolare da un'alimentazione non salutare, inattività fisica e abitudine al fumo, altresì, da fattori ambientali e biologici. Giocano un'importante ruolo nell'aumento degli eventi cardiovascolari l'ipertensione arteriosa, con pressione arteriosa sistolica > 140 o diastolica > 90 mmHg, l'ipercolesterolemia, con colesterolemia >240 mg/dl, l'HDL < 40 mg/dl nel maschio e <50 nelle femmine, l'LDL >115 mg/dl, l'iperglicemia, con un valore tra 110 e 125 mg/dl, l'ipertrigliceridemia, con trigliceridemia >150 mg/dl, infine il diabete, con glicemia a digiuno > 126mg/dl.

Si ritiene che il fumo di tabacco consista in una delle principali cause di mortalità della popolazione dei paesi industrializzati e non solo: in media l'aspettativa di vita di un fumatore si riduce di 13.2 anni per l'uomo e di 14.5 per le donne. Il fumo aumenta la contrattilità miocardica acuta e la rigidità delle grandi arterie periferiche, determinando così un netto aumento del carico di lavoro miocardico. Si consideri inoltre che il trasporto dell'ossigeno risulta ridotto a causa del monossido di carbonio presente nel fumo, il quale si lega all'emoglobina formando carbossiemoglobina<sup>2</sup>, in un contesto al contrario di aumentata necessità<sup>XII</sup>.

Il rischio legato al fumo è, però, proporzionale al numero di sigarette fumate.

I principali settori colpiti riguardano il sistema respiratorio e cardiovascolare. In merito a quest'ultimo si rileva che la nicotina è in grado di aumentare la frequenza cardiaca e la pressione arteriosa attraverso l'attivazione del sistema ortosimpatico. Altresì, determina modificazioni sull'umore, l'attenzione e sull'emotività, riduce l'ansia e la tensione, le quali si associano ad un miglioramento dello stato di allerta e di attenzione, della capacità di concentrazione e anche dell'attività motoria<sup>XIII</sup>. Favorisce la produzione di LDL, importante fattore di promozione e sviluppo dei processi aterosclerotici. In modo quindi lineare, aumenta il rischio di coronaropatia, ictus ed impotenza su base vascolare. E' stato stimato come i forti fumatori (40 sigarette al giorno) abbiano un rischio 10 volte maggiore di avere l'infarto miocardico rispetto ai non fumatori e ai fumatori moderati<sup>XIV</sup>.

Anche il fumo passivo comporta un significativo aumento del rischio cardiovascolare<sup>XV</sup>. È stato dimostrato che i rischi collegati al fumo si riducono sostanzialmente entro due anni dalla sua cessazione<sup>XVI</sup>.

Alcuni studi di genoma wide-association (GWSA) tengono conto dell'esposizione ambientale, come quella data del tabacco, il quale ha un verosimile largo impatto sulle variazioni genetiche che possono predisporre all'insorgere dell'obesità.

Nello studio della relazione tra questi due importanti fattori di rischio, dai nostri risultati emerge come vi sia una stretta correlazione: è stato riscontrato che il

---

tabacco aumenti la suscettibilità genetica sulla totalità del tessuto adiposo, ma tuttavia riduca gli effetti genetici sulla distribuzione del grasso.

In questo studio sono stati utilizzati i dati del genoma WAS dal GIANT (investigazione genetica delle caratteristiche antropometriche), con la volontà di identificare i loci genetici che influenzano l'obesità, valutata mediante la circonferenza della vita, e il rapporto di quest'ultima e i fianchi.

Lo studio identificava 23 nuovi loci genetici correlati all'obesità, di questi, solamente 9 presentavano un'interazione con l'abitudine al fumo.

Fra questi 23, ce ne sono 6 per il BMI, 11 per la circonferenza vita adattata per il BMI e 6 per il rapporto vita- fianchi adattato al BMI<sup>XVII</sup>.

Molti tra i loci identificati per il BMI evidenziavano nuove funzioni biologiche, come la risposta allo stress ossidativo, alla dipendenza e alla funzione regolatoria.

Si nota spesso che i fumatori abbiano un minor peso corporeo e BMI, e tuttavia una maggiore circonferenza della vita rispetto ai soggetti che non fumano. I primi hanno anche una ridotta fluttuazione di peso rispetto a coloro che non hanno mai fumato o che hanno smesso di fumare da più di vent'anni.

È altresì noto che i grandi fumatori (più di 20 sigarette al giorno) e coloro che hanno fumato per più di venti anni sono maggiormente a rischio di obesità rispetto ai moderati fumatori (meno di 20 sigarette al giorno) o ai non fumatori<sup>XVIII</sup>.

Si ritiene, inoltre, che sia gli uomini che le donne acquistino peso molto rapidamente nel momento in cui smettono di fumare. La loro fluttuazione di peso converge con quella dei non fumatori, entro i primi anni dal momento in cui smettono di fumare.

Non è totalmente chiaro il motivo per cui la cessazione del fumo conduca all'aumento del peso e per quale ragione i fumatori a lungo termine lo mantengano invariato. Si suppone che il tabacco sopprima l'appetito<sup>XIX</sup>, ma la prova che i fumatori assumano un numero maggiore di calorie, rispetto ai non fumatori, sembra contraddire questa idea. È possibile, quindi, che il motivo

consista nel fatto che il fumo, e in particolare i sopracitati effetti nicotinici, aumentino l'indice metabolico<sup>XX</sup>.

Molte persone, specialmente le donne, utilizzano impropriamente il fumo come una tecnica per gestire il peso corporeo<sup>XXI</sup>.

L'identificazione dei geni che influenzano l'obesità e la sua interazione con l'abitudine al fumo, ci può quindi finalmente aiutare a chiarire le vie metaboliche che sottostanno ai sopracitati riscontri epidemiologici di una determinata popolazione ad alto rischio cardiovascolare, e alla sua futura terapia.

### SCOPO DELLO STUDIO

Abbiamo utilizzato i dati da GWSA, includendo 51.080 fumatori attuali e 190.178 non fumatori (87% europei) per identificare nuovi loci che influenzano BMI e obesità addominale centrale, misurato come la circonferenza della vita adattata per BMI (WCadjBMI), e il rapporto tra la vita e il fianco adattato per BMI (WHRadjBMI).

### MATERIALI E METODI

Abbiamo applicato quattro approcci per identificare i loci genetici che influenzano i tratti dell'obesità, tenendo conto dello status degli attuali fumatori di tabacco.

Abbiamo definito fumatori quelli che hanno risposto di essere fumatori attuali: non sono stati considerati gli ex fumatori. Abbiamo valutato tre tratti: indice di massa corporea (BMI), circonferenza vita adattata al BMI (WcadjBMI), e il rapporto tra la vita e i fianchi adattato al BMI (WHRadjBMI). L'approccio 1 mirava a determinare le associazioni genetiche con i tratti fenotipici dell'obesità calibrandoli nella popolazione dei fumatori (SNPadjBMI). L'approccio 2 considerava l'impatto congiunto dei principali effetti nella popolazione dei fumatori e l'interazione di tali effetti come già descritto in nello studio di Ashard e altri "Genome-wide meta-analysis of joint tests for genetic and gene-environment



interaction effects. Hum Hered 70, 292-300 (2010)” (SNPjoint); mentre l’approccio 3 e 4 hanno mirato a determinare effetti d’interazione, l’approccio 3 (SNPint) ha considerato solo gli effetti di interazione, mentre l’approccio 4 (SNPscreen) ha seguito i loci identificati nell’approccio 1 per valutare la presenza di altri ulteriori effetti d’interazione.

### DESCRIZIONE DEL GRUPPO E DIMENSIONI DEI CAMPIONI

Il GIANT consortium è un’organizzazione formata da un gruppo internazionale di ricercatori interessati a capire l’architettura genetica dei tratti antropometrici. Nel totale, abbiamo incluso più di 79 studi, comprendendo più di 241.258 individui (incluso più di 51.080 fumatori e 190.178 non fumatori) con attribuzione dati tramite HapMap II (più di 2.8M di SNP nelle analisi associate), e/o con dati da genotyped MetaboChip (~195K SNPs nelle analisi associate).

### DESCRIZIONE DEI FENOTIPI

Il nostro studio evidenzia tre tratti di interesse: BMI, WCadjBMI, WHRadjBMI.

Per ciascun sesso, i tratti sono stati adattati utilizzando la regressione lineare per età<sup>2</sup> (come per BMI, così per WCadjBMI e WHRadjBMI).

Per lo studio della familiarità abbiamo utilizzato modelli di effetti lineari misti, tenendo conto delle relazioni familiari e condotto analisi per i maschi e le femmine combinati includendo il sesso nel modello.

I residui fenotipi sono stati ottenuti dall’adattamento dei modelli e sono stati trasformati successivamente per facilitare l’equiparazione attraverso gli studi e con le analisi pubblicate in precedenza. La trasformazione del tratto è stata condotta separatamente per i fumatori e i non fumatori per il modello stratificato per il fumo, e usando tutti gli individui per il modello adattato al fumo.

## DEFINIZIONE DEI FUMATORI

Gli studi coinvolti hanno mutevoli livelli di informazione sull'abitudine al fumo, alcuni con una sola semplice variabile e altre con dati più precisi. Il dato binario (attuali fumatori “SMK” versus non attuali fumatori) è stato usato per le analisi di molti studi, per cui si può facilmente ricavare questa variabile.

Non abbiamo usato una variabile di “Sempre-fumatore versus mai-fumatore” poiché poteva aumentare l'eterogeneità tra i vari studi, inoltre questa definizione avrebbe reso l'armonizzazione tra i diversi studi ancora più difficoltosa.

## IDENTIFICAZIONE E IMPUTAZIONE DEL GENOTIPO

Gli studi con vasta gamma di dati da GWAS o MetaboChip contribuiscono ai risultati. Ciascuno studio ha adoperato delle esclusioni studio-specifiche per gruppi di campioni, verifica del genere, campioni eterogenei e gruppi etnici diversi. Per molti studi (tranne quelli che impiegano direttamente i genotipi del MetaboChip), il dato del genoma wide chip era attribuito al HapMap II reference data via MACH<sup>102</sup>, IMPUTE<sup>103</sup>, BimBam<sup>104</sup> o Beagle<sup>105</sup>.

## STUDIO DEI LIVELLI DELLE ANALISI

Per ottenere la somma dei dati statistici studio-specifici utilizzati nelle metanalisi sequenziali, i seguenti modelli lineari (o modelli lineari per effetti misti negli studi con familiari/individui legati) sono stati tenuti separatamente per i maschi e per le femmine e separatamente per i casi e i controlli negli studi caso-controllo utilizzando i residuali fenotipi dai modelli descritti sopra. Gli studi con i dati di familiarità hanno anche condotto analisi con questi modelli per i maschi e le femmine combinati dopo aver tenuto conto della dipendenza fra i membri della famiglia come una funzione delle loro correlazioni parentali. Abbiamo supposto un modello genetico aggiuntivo.

Fumo(SMK)- adattato: caratteristica=  $\beta_0 + \beta_1 \text{SNP} + \beta_2 \text{SMK}$

Fumo(SMK)- stratificato: caratteristica=  $\beta_0 + \beta_1 \text{SNP}$  (tenuti separati fumatori/non fumatori.)

Le analisi sono state ottenute utilizzando vari software di GWAS, includendo MACH2QTL<sup>106</sup>, SNPTEST<sup>107</sup>, ProbABEL<sup>108</sup>, GenABEL<sup>109</sup>, Merlin<sup>110</sup>, PLINK<sup>111</sup>, o QUICKTEST<sup>112</sup>.

### IL CONTROLLO DELLA QUALITA' DELLO STUDIO SPECIFICO DEI DATI STATISTICI DEL RIEPILOGO

Il riepilogo dei dati statistici aggregati erano qualitativamente controllati in base a un protocollo standard<sup>113</sup>. Questo includeva dei controlli per i problemi con le trasformazioni dei tratti, le frequenze degli alleli. La bassa qualità dell'SNP in ciascuno studi era esclusa per i seguenti criteri:

- (i) SNPs con una ridotta somma di minor alleli ( $\text{MAC} \leq 5$ ,  $\text{MAC} = \text{MAF} * N$ ) e SNPs monomorfo,
- (ii) genotipando SNPs con un basso SNP call-rate (<95%) o un basso test di equilibrio di Hardy-Weinberg del valore P (< $10^{-6}$ ),
- (iii) imputato SNPs con ridotta qualità di imputazione (MACH-  $R_{sq}$  o  $\text{OEVAR} < 0.3$ , punteggio di informazione <0.4 per SNPTEST/IMPUTE/IMPUTE, o <0.8 per PLINK).

Per testare i problemi dei campioni non collegati o sovrapposti e per correggere la stratificazione potenziale della popolazione, gli errori standard dello studio specifico e le associazioni di valore P erano corretti dal controllo genomico (GC) utilizzando fattori  $\lambda$ <sup>114</sup>.

La correzione del GC per i dati del GWAS usava tutti gli SNPs, ma la correzione di GC per i dati di MetaboChip erano limitati nell'intervallo QTat the chip era arricchito per le associazioni dei tratti relativi all'obesità. Qualche studio di livello

del GWAS col fattore lambda >1.5 era stato rimosso da altre analisi. Mentre noi abbiamo stabilito questi criteri, nessun risultato è stato rimosso per questa ragione.

## LE METANALISI

Le metanalisi usavano i dati statistici del riepilogo dello studio specifico per le associazioni dei fenotipi per ciascuno dei modelli sopraindicati. Abbiamo utilizzato la variazione inverse degli effetti stabili ponderata al metodo per le analisi degli effetti principali di SNP. Tutte le analisi utilizzavano il METAL<sup>115</sup>. Come i risultati dello studio vengono suddivisi in due gruppi (Stage 1 e 2), le meta analisi dai due gruppi erano ulteriormente meta analizzati. La seconda correzione di GC era applicata a tutte le SNP quando le meta-analisi dello stage 1 e 2 si combinavano nelle meta analisi finali. Primo, attribuiva Hapmap e i dati di GWAS i quali venivano analizzati insieme, come degli studi sul MetaboChip. Questo step era seguito dalla combinazione dal GWAS + analisi del MetaboChip. Per le analisi primarie, abbiamo condotto delle analisi attraverso gli antenati e il sesso. Per le meta analisi, abbiamo condotto meta analisi solamente su studi dei discendenti europei, e meta analisi sul sesso specifico. C'erano due ragioni che conducevano alle meta analisi secondarie. Primo, entrambi WCadjBMI e WHRadjBMI hanno mostrato gli effetti genetici del sesso specifico<sup>7,30,116</sup>. Secondo, includendo la popolazione per molteplici antenati nelle nostre prime meta analisi, potevamo spiegare l'eterogeneità dovuta alle differenze delle taglie, frequenze degli alleli, campioni di LD sugli antenati, il potere decrescente per rilevare gli effetti genetici. Guardare la suppl. fig.1 per il riepilogo del disegno primario delle meta analisi. L'ottenuto SMK stratificato viene usato dopo per calcolare il riepilogo dei dati statistici di SNPjoint e SNPint usando Easystrata<sup>117</sup>.

Rapidamente, questo software esegue due campioni, un largo test di campioni dei parametri di eguale regressione tra fumatori e non fumatori come descritto da Randall et al<sup>30</sup> per SNPint e il test due gradi di libertà dei principali effetti di interazione per SNPjoint come descritto da Aschrad et al<sup>31</sup>.

## SELEZIONE DEL PRINCIPALE SNP

Prima di selezionare per ciascun loco un principale SNP, sono stati esclusi SNPs con alta eterogeneità  $I^2 > 0.75$  o la taglia del campione sotto il 50% del massimo N per ciascun strato.

I principali SNPs che presentano dei criteri significativi sono stati selezionati basandosi sulla distanza ( $\pm 500\text{Kb}$ ), e abbiamo definito l'SNP con il più basso valore di P come il vertice di SNP per il loco. Gli SNPs che raggiungono il significato del genome wide (GWS), ma non aveva altre SNPs entro i 500Kb con valore di  $P < 1E^{-5}$ , era escluso dall'SNP il processo di selezione. Due varianti erano escluse dall'approccio 2 basato su questi criteri, rs 2149656x WCadjBMI e rs236267 per WHRadjBMI.

## ZONE DI LOCUSZOOM

Le zone di associazione regionale erano prodotte per i nuovi loci utilizzando il programma Locuszoom<sup>118</sup>. Per ciascuna zona, LD era calcolato utilizzando un campione multietnico di 1000 gruppi genomici<sup>119,120</sup>, includendo europei, africani, asiatici e americani. Le precedenti associazioni dei tratti di SNP evidenziavano entro le zone incluse i tratti di interesse (cardiometabolici, comportamenti dipendenti, antropometrici) trovati nel catalogo NGHRI-EMI di GWAS e integrato con i recenti studi della letteratura<sup>6,7,116,121</sup>.

## ANALISI CONDIZIONALI

Per determinare se i molteplici segnali di associazione erano presenti entro un singolo loco, abbiamo utilizzato GCTA<sup>32</sup> per compiere delle analisi condizionali nei dati di SNPadjSMK e SMK stratificato. I criteri successivi per selezionare i loci candidati per le analisi condizionali sono stati la vicinanza di SNP ( $\pm 500\text{Kb}$ ) con  $R^2 > 0.4$  e un'associazione di  $P < 1E^{-5}$ . Il GCTA usa associazioni dalla nostra metanalisi e gli stimati LD dai gruppi dei dati di

riferimento contenendo dati genotipici individuali per compiere delle analisi condizionali. Per calcolare la struttura di LD, abbiamo utilizzato due gruppi americani, lo studio dell'AIRC ( rischio di aterosclerosi nelle comunità) comprendendo 9.713 individui di discendenti europei e 580 individui di discendenti afroamericani, e lo studio di FramHS (Framingham Heart Study) comprendendo 8.481 individui di antenati europei, entrambi gli studi attribuiti all'HapMap r22. Comunque, perché le nostre prime analisi fossero condotte su molteplici antenati, ciascuno studio integrava i dati genetici usando le popolazione di riferimento di HapMap perciò questo gruppo finale di riferimento era composto da quasi 1-3% di asiatici e 4-6% di africani per il totale di campioni di riferimento.

Abbiamo estratto una regione di 1MB intorno al nostro candidato di SNPs, abbiamo compiuto delle analisi condizionali approssimative, e poi abbiamo ripetuto questi steps per un appropriato l'approccio per identificare un segnale di associazione aggiuntivo.

Molti degli SNPs identificati nelle analisi attuali erano nelle vicinanze dei SNPs associati in precedenza con i tratti antropometrici collegati e l'obesità (altezza, tessuto adiposo viscerale). Per tutti i principali SNPs vicino a SNP associati in precedenza con queste caratteristiche, GCTA era anche usato per compiere delle analisi condizionali approssimative nei dati di SNPadjBMI e SMKstratificato in modo da determinare se i loci identificati qui sono indipendenti dalle associazioni di SNP-trait identificato in precedenza.

#### ANALISI DI FOLLOW-UP

Per identificare i geni che possono essere implicati nell'associazione tra i nostri principali SNPs e BMI, WHRadjBMI e WCadjBMI e per mettere in luce la complessa relazione tra le variazioni genetiche, SMK e obesità, abbiamo compiuto delle ricerche dettagliate nella letteratura nella vicinanza dei geni candidati. Snipper 1.2 era utilizzato per identificare alcuni geni e cis- o trans-eQTLs entro i 500Kb dei nostri principali SNPs. Tutti i geni identificati mediante

Snipper erano curati manualmente ed esaminati per dimostrare la relazione con il fumo o l'obesità. Per esplorare qualunque potere regolatore o ruolo di funzione delle regioni di associazione, i loci erano anche esaminati usando differenti dati bioinformatici, includendo HaploReg v4.1<sup>122</sup>, Browser del genoma UCSC<sup>123</sup>, GTeX Portal<sup>124</sup>, e RegulomeDB<sup>125</sup>.

### ANALISI eQTL (expression quantitative trait loci)

Abbiamo utilizzato due approcci per esplorare sistematicamente il ruolo dei nuovi loci nell'espressione genetica.

Primo, per acquisire una visuale generale del ruolo regolatore delle regioni di GWAS di recente identificate, abbiamo condotto un eQTL usando > 50 studi Eqtl<sup>126</sup>, con specifiche citazioni per >100 informazioni incluse nell'attuale quesito per le cellule sanguigne collegate agli studi eQTL e tessuto cellulare non sanguigno (adiposo e il tessuto del cervello). In aggiunta l'informazione eQTL era integrata da fonti online includendo ScanDB, il Broad Institute GTex Portal, e il Laboratorio Pritchard). Solo il significativo ci-eQTLs nell'alto LD con il nostro nuovo principale SNPs, proxy SNPs, erano tenuti in considerazione.

Secondo, dato che le banche dati con i dati di eQTL non hanno informazioni disponibili sullo status dei fumatori attuali, abbiamo anche condotto analisi sul cis-eQTL usando i risultati derivanti dal digiuno del sangue periferico usando l'Human Exon 1.0 (Affymetrix, inc,santa clara, CA). i dati grezzi erano normalizzati, trasformata log2, seguita dal riassunto usando Robust Multiarray Average<sup>127</sup> e ulteriormente adattati per le covarianti tecniche, includendo il principale componente del espressione del dato, le somme delle cellule sanguigne, il gruppo di appartenenza. Abbiamo valutato tutte le trascrizioni +/- 1b attorno ad ogni nuova variante nello studio del cuore di Framingham mentre tenendo conto dello status degli attuali fumatori, usando i seguenti quattro approcci simili a quelli usati nella nostra prima analisi delle nostre caratteristiche: 1)eQTLadattato per SMK, 2)eQTL stratificato da SMK, 3) interazione di eQTL x SMK, 4) joint main + interazione di eQTLxSMK). Il valore di livello era valutato da FDR<5%

per eQTL analisi e su tutti i loci identificati per quel modello nella meta analisi primaria.

### VALUTAZIONE DELLA VARIAZIONE ESPLICATA

Abbiamo valutato la variazione fenotipica nei fumatori e non fumatori esplicita mediante i segnali di associazione utilizzando una metodica descritta in precedenza con Kutalik et al<sup>128</sup>. Per ogni regione associata, abbiamo selezionato delle sottocategorie di SNPs entro i 500Kb dei nostri principali SNPs e basandoci sul mutevole valore soglia di P (oscillando da  $1 \times 10^{-8}$  a 0.1) dall'approccio 1 (modello di SNP<sub>adj</sub>SMK). Primo, ciascuna sottocategoria di SNPs era raggruppata dentro regioni indipendenti per identificare il principale SNP per ciascuna regione. La variazione esplicita da ciascuna sottocategoria di SNPs nel livello SMK e non SMK era stimata sommando la vibrazione esplicita da SNPs indipendenti.

### ESAMINARE IL COMPORTAMENTO DEL FUMATORE

Al fine di determinare se alcuni dei loci identificati nello studio attuale sono associati con il comportamento del fumo, siamo andati a cercare di tutti i principali SNPs dai nuovi loci e dall'approccio 3 nell'esistente GWAS del comportamento del fumatore<sup>9</sup>. L'analisi consiste nel frazionare i campioni dello studio specifico di GWAS contribuendo alle meta analisi del comportamento del fumo, imputazione, verificando l'associazione e le meta analisi. Per assicurarsi che tutti gli SNPs di interesse erano disponibili nel GWAS del fumatore, il programma SHAPEIT2<sup>129</sup> era usato per frazionare la regione di 500kb in qualunque parte di ciascun principale SNP, e l'imputazione era eseguita utilizzando IMPUTE2<sup>130</sup> con il set di dati di 1000 genomi come gruppo di riferimento.

Ogni regione era analizzata per tre fenotipi legati il fumo: (i)sempre fumatori o mai, (ii) attuali contro i non attuali fumatori, (iii) una misura categorica della



quantità del fumo<sup>91</sup>. I livelli della quantità del fumo erano 0 (definiti come 1-10 sigarette al giorno), 1 (da 11 a 20), 2 (da 21 a 30) e 3 (più di 31). Ciascun incremento rappresenta un aumento nella quantità del fumo per 10 sigarette al giorno. C' erano 10.058 che non hanno mai fumato, 13.418 che hanno sempre fumato, 11.796 fumatori non attuali, 6966 attuali fumatori e 11.436 campioni con i fenotipi SQ. SNPPMETA<sup>91</sup> era usato per compiere una meta analisi degli effetti invariabili ponderati della variazione inversa attraverso i gruppi di tutti gli SNPs di ciascuna regione, e includeva una sola correzione di GC. A ciascun SNP, solo quei gruppi che avevano una imputazione con punteggio > 0.5 che erano inclusi nella meta analisi.

## RISULTATI

Abbiamo utilizzato i risultati di studi specifici da un totale di 57 studi da GWAS e da 22 studi genotipati con Illumina MetaboChip, includendo più di 241,258 individui, e più di 51,080 fumatori e più di 190,178 non fumatori.

Abbiamo condotto delle meta analisi sulla variante inversa degli effetti ponderate dal GWAS per tre caratteristiche di adiposità (BMI, CVadj BMI, WHRadjBMI) e per due modelli di regressione per un duplice stato di fumatori (SMK: attuali fumatori contro i non fumatori attuali):

- 1) gli effetti del polimorfismo a singolo nucleotide(SNP) adattati per SMK (SNPadjSM)
- 2) gli effetti principali di SNP stratificati da SMK.

I risultati stratificati di SMK sono stati utilizzati dopo per calcolare l'interazione (SNPint)<sup>30</sup> e gli effetti combinati(principale + interazione) (SNP joint)<sup>31</sup>.

Per delle analisi primarie, abbiamo condotto delle meta-analisi attraverso gli antenati e il sesso. Per delle analisi secondarie, abbiamo condotto delle analisi sui discendenti europei, e analisi su uno specifico sesso.

Abbiamo considerato 4 approcci analitici:

Approccio 1 (SNPadjSMK) esamina le associazioni genetiche con i tratti dell'obesità dopo l'adattamento per SMK.

Approccio 2 (SNPjoint) considerato l'impatto combinato degli effetti principali adattati per SMK + effetti di interazione come descritto da Aschard et al<sup>31</sup>.

Approccio 3 (SNP int) e 4. (SNPscreen) entrambi mirano ad identificare gli effetti dell'interazione tra i due fattori, nell'approccio 3., abbiamo considerato solo gli effetti dell'interazione, nel 4, abbiamo fatto follow up ai loci identificati nell'approccio 1 per identificare gli effetti delle interazioni addizionali. (Snpint).

I risultati dall'approccio 1 e 3 sono considerati significativi se il valore di  $p < 5 \times 10^{-8}$ , mentre l'approccio 4 ha usato la correzione di Bonferroni dopo lo screening.

Se la variante principale era  $> 500\text{kb}$  rispetto alla precedente associazione con BMI, WCadjBMI, e WHRadjBMI, venivano considerate nuove associazioni.

In totale e attraverso le tre caratteristiche di distribuzione di grasso, abbiamo identificato 23 nuovi loci genetici associati (6 per BMI, 11 per WCadjBMI, 6 per WHRadjBMI) e 10 loci con significativi effetti di interazione GxSMK ( 2 per BMI, 2 per WCadjBMI, 5 per WHRadjBMI).

Siamo in grado di fornire un confronto globale con i loci precedentemente identificati<sup>6</sup> mediante i tratti nel materiale aggiuntivo.

#### BASANDOCI SULLO STATO DEGLI ATTUALI FUMATORI SI IDENTIFICANO NUOVE ASSOCIAZIONI

Per l'analisi primaria di BMI (antenati combinati e sessi combinati), per totale di 58 loci raggiunti per GWS nel approccio 1 (  $P_{\text{snpadjsmk}} < 5 \times 10^{-8}$ ), includendo due nuovi loci vicini SOX11, e SRRM1P2.

Altri tre loci di BMI sono stati identificati usando l'approccio 2 (SNpjoint), includendo un aggiuntivo nuovo loco vicino CCDC93.

Per identificare i loci che nascondono molti indipendenti SNPs associati con qualsiasi risultato, abbiamo usato GCTA<sup>32</sup> per compiere un'analisi approssimativa usando un testo statistico marginale dalla nostra prima analisi.

L'analisi condizionata di BMI non rivela i segnali secondari entro 500kb della nostra nuova principale SNPs.

In aggiunta, abbiamo svolto un'analisi condizionata per determinare se le nuove variazioni del BMI nell'attuale analisi erano indipendenti dagli precedenti loci identificati in GWAS associati con caratteristiche relativi di interesse.

L'SNP associato con BMI nel cromosoma 2, s10929925 (vicino SOX11), identificato in approccio 1, rimasto significativo ( $P_{\text{snpadsmkj}} = 1,4 \times 10^{-9}$ ) dopo il condizionamento nella vicinanza SNP rs10495537 precedentemente associato con tessuto adiposo viscerale (VAT)<sup>33</sup> indicando che questo loco è indipendente dal segnale di associazione del VAT.

Dalla primaria analisi di WCadjBMI, abbiamo identificato 62 GWS loci per l'approccio 1 (SNPadjSMK) e due loci aggiuntivi per l'approccio 2 (SNPjoint), includendo 8 nuovi loci vicini KIF1B, HDLBP, DOCK3, ADAMTS3, CDK6, GSDMC, TMEM38D, e ARFGEF2.

Le principale varianti al primo loco vicino a PSMB10 identificato nell'approccio 1 e 2 (rispettivamente rs14178 e s113090) sono > 500kb dalla precedente identificata variante associata WCadjBMI; comunque, dopo il condizionamento delle varianti conosciute, il nostro segnale è molto ridotto ( $P = 3,02 \times 10^{-2}$  e  $P = 5,22 \times 10^{-3}$ ), indicando che questi risultati non sono nuovi. non erano stati identificati i segnali secondari di GWS su nessun nuovo loco per WCadjBMi.

Sei nuovi loci per WCadjBMi erano vicini a delle precedenti associazioni con caratteristiche antropometriche relative. Di questi sei, alcuni segnali associati per s6743226 vicino a HDLBP, RS10269774 vicino CDK6, e rs6012558 vicino a ARFGEF2 erano ridotti ( $P > 1E^{-5}$  e beta ridotta della metà) dopo il condizionamento di almeno una della vicina altezza e circonferenza fianchi

adattato per BMI (HIPadjBMI) SNP, ma il segnale di associazione è rimasto indipendente dalle altre caratteristiche associate di SNP.

Per la prima analisi di WHRadjBMi, un totale di 32 loci erano identificati nell'approccio 1 (SNPadjSMK), includendo un nuovo loco vicino HLA-C, con nessun loco identificato tenendo conto dell'interazione con SMk nell'approccio 2. (SNPjoint).

Nessun segnale secondario di GWS è stato identificato in un nuovo loco per WHRadjBMI. Tuttavia, il condizionamento su altre quattro SNPs conosciute associate con l'altezza, HIPadjBMI o VAT, abbiamo trovato che il nostro segnale GWAS è attenuato dal condizionamento nelle due varianti conosciute (rs6457374 e s2247056), però rimane significativa per altre analisi condizionali. quindi, mentre questo loco rappresenta una nuova associazione per tratti relativi all'obesità, il nostro principale SNP (rs1049281) non è indipendente dai loci responsabili dell'altezza precedentemente identificati .

diversi loci aggiuntivi erano identificati per l'approccio 1 e 2 in almeno una della analisi secondarie che non erano identificati in nessuna delle analisi primarie. Per BMI, due nuovi loci erano identificati mediante l'approccio 1. un loco era identificato solamente nelle analisi sulle femmine e nelle analisi degli europei (vicino EPHA3) e un loco è stato identificato nelle analisi europee combinate al sesso ( vicino INADL). Per WCadjBMi, due nuovi loci sono stati identificati solo nelle analisi sugli europei, un loco per approccio 1 ( solo uomini, vicino RAI14), e un loco per approccio 2 (solo nelle donne, vicino PRNP). Per WHRadjBMI, 5 nuovi loci sono stati identificati nella seconda analisi. due loci, vicino BBX, TRBI1, sono stati identificati nelle femmine in tutti i discendenti gruppi utilizzando approccio 1. Tre loci sono stati identificati nella analisi del sesso combinato degli europei, includendo due per approccio 1 (vicino HMT2 e SMIM2) e uno per approccio 2 (vicino EYA4).

## PROVE DELLA MODIFICAZIONE DELLA PREDISPOSIZIONE GENETICA SULL'OBESITA' DAL FUMO

Utilizzando i risultati dell'analisi SMK stratificato, l'approccio 3 direttamente valuta l'effetto dell'interazione tra GxSMk.

Per l'analisi primaria per BMI, due loci raggiunti GWS per SNPint includendo l'interazione precedentemente identificata Gx SMK vicino CHRNA4<sup>9</sup>, e nuovi loci vicino INPP4B.

Entrambi i loci mostrano gli effetti GWS sul BMI nei fumatori e nessun effetto di BMI nei non fumatori. Per la variante vicino CHRNA4 ( recettore nicotinolinergico B4), l'allele minore (G) mostra un effetto ridotto sul BMI nei fumatori attuali (beta smk = - 0.047) però non ha effetti sui non fumatori ( beta non fumatori= 0.002).

Precedenti studi hanno identificato che in vicinanza dei polimorfismi a singoli nucleotidi in alto LD (linkage disequilibrium) , associato col fumo ( non sinonimo rs16969968 nel CHRNA5)<sup>9</sup> e calcificazione atriale (rs3825807, variante nel ADAMTS7)<sup>34</sup>. Il condizionamento su queste varianti ha attenuato gli effetti dell'interazione ma non le ha eliminate, suggerendo una relazione complessa tra il fumo, l'obesità, patologie cardiache, e le varianti genetiche in questa regione.

Significativamente, il gene CHRNA5- CHRNA3- CHRNA4 è stato associato con il ridotto BMI nei fumatori attuali<sup>9,35</sup>, però con alto BMI nei non fumatori<sup>9</sup>, il quale è supportato dalla mancanza dell'associazione per BMI nei non fumatori nel nostro studio così come la mancanza dei precedenti risultati GWAS su 15q25 per BMI<sup>6</sup>.

Il gene CHRNA5- CHRNA3- CHRNA4 codifica il recettore della subunità nicotina acetilcolina, alfa3, alfa 5 e beta4 che sono espressi nel sistema nervoso centrale<sup>36,37</sup>.

La nicotina ha differenti effetti sul corpo e sul cervello, sulle alterazioni del metabolismo, sulle calorie prese e sul comportamento alimentare<sup>38,40</sup>.

Questi risultati suggeriscono che l'esposizione al fumo può modificare gli effetti genetici su 15q24-25 per influenzare le malattie relative al fumo mediante diverse vie metaboliche, come l'obesità.

Nella primaria analisi di WCadjBMI, un nuovo loco (vicino GRIN2A) con effetto di direzione opposta tra fumatori e non fumatori è stato identificato per l'approccio 3 (SNPint). L'allele T di s4141488 mostra un effetto positivo su WCadjBMI nei fumatori attuali in confronto ai ridotti effetti nei non fumatori (beta SMK= 0.037, beta non fumatori= -0.015). Inoltre, nella nostra seconda analisi sulle sole donne europee abbiamo identificato un'interazione tra rs6076699, vicino PRNP, e SMK su WCadjBMI, un loco che era stato identificato nell'approccio 2 (SNPjoint) per donne europee. L'allele maggiore, A, ha un effetto positivo sui fumatori attuali in confronto a un effetto debole e negativo sulla WC nei non fumatori (beta fumatori= 0,169, beta non fumatori= -0,070). La direzione opposta degli effetti nei fumatori in confronto ai non fumatori per questi loci può spiegare il perché questa variante è rimasta inosservata in precedenza GWAS di WCadjBMI.

Abbiamo utilizzato un approccio aggiuntivo, approccio 4 (SNPscreen), (fig.1 online metodi) per valutare l'interazione GxSMK per cui abbiamo controllato i risultati di SNPadjSMK (da approccio1) e poi abbiamo esaminato l'interazione con la correzione di Bonferroni nella sottocategoria delle varianti con i principali effetti di GWS adattati.

Mentre nessun loco è stato identificato per WHRadjBMI utilizzando l'approccio 3 abbiamo identificato delle significative interazioni per WHRadjBMI con l'approccio 4.

Nella prima analisi, due SNPs, vicino LYPLA1 e RSPO3, abbiamo trovato delle interazioni significative; entrambi hanno precedentemente pubblicato effetti principali su tratti antropometrici<sup>1</sup>.

Questi loci mostrano effetti su WHRadjBMI nei non fumatori, però non nei fumatori. Nella nostra seconda analisi, abbiamo identificato tre conosciuti loci con

una significativa interazione di effetti tra GxSMK sul WHRadjBMI utilizzando l'approccio 4(SNP screen) vicino MAP3K1, HOXC4- HOXC6, e JUND.

Inoltre, la variante, rs1809420, vicino il conosciuto gene cluster CHRNA5-CHRNA3- CHRNA4 erano identificati utilizzando l'approccio 4 per BMI nei soli maschi, analisi dei combinati ascendenti.

Il calcolo della potenza dimostra che l'approccio 4 ha incrementato la forza per identificare SNPs che mostra (i) un effetto in uno strato (fumatori e non fumatori) e meno pronunciato ma effetto concordante negli altri strati, o (ii) un effetto più largo nello strato dei non fumatori e nessun effetto nei fumatori. In contrasto, approccio 3, ha aumentato la potenza dell'SNPs che mostra (i) un effetto più piccolo nello strato dei fumatori e nessun effetto nei non fumatori, oppure (ii) un effetto opposto tra fumatori e non fumatori.

I nostri risultati per entrambi gli approcci concordano con la previsione dalle analisi sulla potenza, supportando l'utilità del lavoro di entrambi gli approcci analitici per identificare le interazioni GxSMK.

#### ARRICCHIMENTO DEGLI EFFETTI GENETICI DEI FUMATORI IN RELAZIONE ALL'OBESITA'

Quando si esaminano gli effetti specifici del fumo per il BMI e i loci WCadjBMI attraverso la nostra analisi, non si nota alcun arricchimento significativo degli effetti genetici nei fumatori. Contrariamente al BMI e WCadjBMI, i nostri risultati per WHRadjBMI erano arricchiti per i loci con effetto più forte nei non fumatori in confronto ai fumatori, con 35 dei 45 loci visualizzati numericamente con un effetto maggiore nei non fumatori ( $P_{\text{binomial}} = 1,2 \times 10^{-4}$ ).

Nell'ulteriore indagine il risultato della contribuzione dei loci sull'adiposità è diverso da SMK, abbiamo calcolato la variazione esplicita dalle sottocategorie di SNPs selezionate in 15 soglie incrementate significativamente per l'approccio 1 da  $P_{\text{snpadjSMK}} = 1 \times 10^{-8}$  a  $P_{\text{snpadjsmk}} = 0$ .

Le differenze nella variazione è esplicita tra fumatori e non fumatori erano significative (utilizzando  $P < 0.0003 = 0.05/15$ , correzione di Bonferroni per 15 soglie) per BMI per ogni soglia, con più variazione esplicita nei fumatori rispetto ai non fumatori. Così, mentre non si trova un arricchimento per i loci con più forti effetti in uno strato in confronto all'altro nei nostri loci GWS, loci BMI esplicano più variazioni nei fumatori rispetto ai non fumatori e queste differenze aumentano con l'aumento della soglia.

Per WCadjBMI, la differenza era significativa per i siti di SNP iniziando con  $P_{\text{snpadjBmi}} < 3,16 \times 10^{-4}$ , e per WHRadjBMI a  $P_{\text{snpadjBmi}} < 1 \times 10^{-6}$ , però diversamente da BMI, SNPs dall'approccio 1 ha spiegato un maggiore rapporto della varianza nei non fumatori rispetto ai fumatori attuali.

Questi risultati suggeriscono che il tabacco può incrementare la suscettibilità genetica su tutto il tessuto adiposo, ma attenua gli effetti genetici sulla distribuzione del grasso. Questo risultato, che non stupisce totalmente, ha dato delle osservazioni fenotipiche più alte di obesità totale e più basso di obesità addominale nei fumatori<sup>11,12,14,16</sup>.

Inoltre, il fumo aumenta lo stress ossidativo e infiammazione generale del corpo<sup>41</sup> e recenti studi hanno scoperto che lo stress ossidativo può aumentare il peso, specialmente quando combinato con il consumo di alte calorie<sup>42,43</sup>. Molti geni implicati nel BMI sono coinvolti nella regolazione dell'appetito e nel comportamento alimentare<sup>6</sup>.

Comunque, per la circonferenza della vita, abbiamo adattato per il totale adiposo e sono probabili le evidenziazioni di altre vie metaboliche mediante le quali il fumo altera la suscettibilità genetica nella distribuzione del grasso del corpo.

Le differenze nella variazione esplicita erano più grandi per il BMI (le differenze oscillavano da 1,8%- 21% per i fumatori) e più basse per WHRadjBMI (oscillavano tra 0.3% a 8.8% per i non fumatori)<sup>19,21,44</sup>.



## RUOLO FUNZIONALE O BIOLOGICO DEI NUOVI LOCI RELATIVI ALL'OBESITA'

Affinché si possa ottenere la comprensione sulla rilevanza biologica dei nostri nuovi 23 loci identificati, abbiamo svolto mediante delle ricerche di letteratura e database bioinformatiche disponibili al pubblico per capire il ruolo funzionale di tutti i geni entro i 500kb dei nostri principali SNPs. Inoltre, abbiamo esplorato sistematicamente il ruolo potenziale dei nostri nuovi loci nell'effettuare l'espressione dei geni con e senza tenendo conto dell'influenza del comportamento del fumo.

## SUPPORTO ADDIZIONALE PER LE FUNZIONI BIOLOGICHE E LE VIE METABOLICHE COINVOLTE PRECEDENTEMENTE

Abbiamo scoperto che la maggioranza dei nuovi loci sono vicini ai forti geni candidati con funzioni biologiche simili ai loci relativi all'obesità identificati in precedenza, includendo la regolazione del grasso corporeo/peso, angiogenesi/adipogenesi, omeostasi di glucosio e lipidi, crescita generale e sviluppo, ecc.

Per esempio, abbiamo identificato rs17396340 associato con WCadjBMI per l'approccio 1 e 2, un introne variabile nel gene KIF1B. Questa variante è associata con l'espressione di KIF1B in tutto il sangue tenendo conto e non per i SMK<sup>45</sup> (GTEx). Come altri geni associati all'obesità, KIF1B è assai espresso nel cervello<sup>46</sup>. Le forme mutate di KIF1B nel topo sono risultate molte anomalie cerebrali, includendo la morfologia del ippocampo<sup>47</sup>, una regione coinvolta nella memoria e nella cognizione, includendo la memoria alimentare<sup>48,49</sup>. In aggiunta, l'apice del SNP, rs17396340, è associata con l'espressione dei livelli di ARSA (arylsulfatase A) nel tessuto LCL (legamento laterale collaterale).

L'adiposo umano esprime una funzionale ARSA, la quale converte il solfato dopamina in dopamina attiva. La dopamina è coinvolta nella regolazione dell'appetito a lungo termine attraverso la regolazione della leptina e dei livelli

dell'adiponectina, suggerendo un ruolo regolatore per ARSA nella regolamentazione dell'appetito<sup>50</sup>.

In aggiunta, il gene vicino al nostro principale SNP per WHRadjBMI per l'approccio 1 per le sole donne, rs670752, può essere d'interesse per gli studi del follow up. CD47 (molecola CD47) codifica un antigene superficiale della cellula coinvolto nella risposta immunitaria per i batteri, adesione cellulare, risposta infiammatoria, e il segnale da una cellula all'altra<sup>51,52</sup>. L'espressione CD47 è ridotta significativamente negli individui obesi e correlata negativamente con il BMI, CV, e la circonferenza fianchi<sup>53</sup>. Al contrario, nei topi, la carenza di CD47 mostra un ridotto acquisto del peso con una dieta ad alto contenuto lipidico, aumento della spesa energetica, migliore profilo del glucosio, e riduzione dell'infiammazione.

#### RECENTI FUNZIONI BIOLOGICHE E VIE METABOLICHE COINVOLTE

Diversi nuovi loci coinvolti nelle uniche funzioni biologiche e le vie metaboliche includono quei comportamenti dipendenti e la risposta allo stress ossidativo, evidenziando la rilevanza potenziale tenendo conto del fumo nelle analisi del GWAS. Questi potenziali geni candidati vicino ai nostri segnali associati sono altamente espressi nei tessuti rilevanti per la regolazione dell'adiposo e nei comportamenti del fumo (esempio: cervello, tessuto adiposo, fegato, polmone, muscolo)(suppl. nota 2, suppl. tab.19).

Per esempio, il cluster CHRNA5- CHRNA3- CHRNA4 dei geni sono coinvolti nell'enzima ossido nitrico sintasi segnalando la via metabolica che è la chiave per neutralizzare le specie reattive dell'ossigeno introdotto dal tabacco e dall'obesità<sup>55,56</sup>. La distribuzione di questa via metabolica è stata associata con la disregolazione di adiponectina nell'adiposo dei topi obesi, implicando questa via metabolica verso l'effetto della regolazione del peso<sup>56,57</sup>.

Mentre non c'è nessuna relazione diretta tra INPP4B e l'obesità o lo stress ossidativo, recenti studi hanno coinvolto l'INPP4B nella regolazione del

PI3K/Akt segnalando la via metabolica<sup>58,59</sup>. Questa via metabolica è importante per un numero di processi biologici impegnato nella crescita e la proliferazione cellulare, però anche nel segnale del'enzima ossido sintasi eNOS, metabolismo dei carboidrati, e l'angiogenesi<sup>60</sup>.

GRIN2A, vicino a rs4141488 associato con WCadjBMI nell'approccio3 (SNPint), aiuta nel controllare la memoria a lungo termine e la conoscenza attraverso la regolazione e l'efficienza della trasmissione sinaptica<sup>61</sup> e è stato associato con la dipendenza da eroina e le malattie ossessive- compulsive<sup>62-68</sup>. In aggiunta, la nicotina aumenta l'espressione di GRIN2A nella corteccia prefrontale nei modelli dei topi<sup>69</sup>. Non c'è nessuna relazione stabilita tra GRIN2A e i fenotipi relativi all'obesità nella letteratura, ancora la memantina e ketamina, gli antagonisti farmacologici dell'attività di GRIN2A<sup>70-74</sup>, sono implicati nel trattamento dell'obesità associata alle malattie, includendo il disturbo da alimentazione incontrollata e l'obesità (clinical trails identifica: NCT0030655, NCT02334059, NCT01997515, NCT01724983).

Inoltre, la memantina è oggetto di investigazione clinica per il trattamento della dipendenza da nicotina (clinical trails identifica: NCT01535040, NCT00136786, NCT00136747).

Mentre il nostro principale SNP non è entro il gene descritto, rs4141488 e le varianti nell'alto LD ( $r > 0.7$ ) con il nostro principale SNP sono entro delle regioni attive di enhancer per diversi tessuti, includendo il fegato, il muscolo della gamba del feto, il tessuto liscio dello stomaco e il muscolatura intestinale, la corteccia, e diversi tipi di cellule embrionali e pluripotenti, e di conseguenza può rappresentare un'importante regione regolatoria per i geni vicini come GRIN2A.

Nella nostra seconda analisi relativa alle sole donne europee, abbiamo identificato una significativa interazione del GxSMK per rs6076699 su WCadjBMI. Questo SNP è 100Kb controcorrente al gene PRNP (proteina prione), il quale è un segnale transdotto implicato in molteplici processi biologici relativi al sistema nervoso, sistema immunitario, e molte funzioni generale delle cellule<sup>75</sup>.

Curiosamente, le forme alterne dei oligomeri hanno mostrato la risposta allo stress ossidativo causato dall'esposizione al rame<sup>76</sup>.

Il rame è presente nelle sigarette ed è elevato nel siero dei fumatori, però non è considerato fuori dal sicuro rang secondo i centri per il controllo e la prevenzione delle malattie, il centro nazionale per la prevenzione delle malattie croniche e la promozione della salute, e l'Office on smoking and health<sup>77,78</sup>. Un altro gene vicino a rs6076699, SLC23A2 (solut carrier family23) è essenziale per l'assorbimento e il trasporto della vitamina C, un importante nutriente per il DNA e la riparazione cellulare in risposta allo stress ossidativo entrambi direttamente e attraverso il supporto della riparazione della vitamina E dopo l'esposizione agli agenti ossidativi<sup>79-81</sup>. SLC23A2 è presente nel surrene e nei modelli dei topi (murinae) che gioca un ruolo importante nella regolazione del livello della dopamina<sup>82</sup>, regolatore dell'appetito<sup>50</sup>. Per di più, questa regione è associata con successo nella cessazione del fumo ed è implicata nei comportamenti di dipendenza in generale<sup>83,84</sup>.

Il nostro tag SNP è localizzato entro una regione attiva di enhancer (segnato da segnali aperti di cromatina, DNase ipersensitive, e la trascrizione Motifmap\* factor binding motifs\* ); questa attività regolatoria appare tessuto specifico (tessuti specifici del sesso e dei polmoni).

Inoltre, la nicotamina mononecluata adenotransferasi (NMNAT1), la quale è a monte alla WCadjBMI variante rs17396340, è responsabile per la sintesi del NAD (Nicotinammide adenina di nucleotide) dall'ATP e NMN<sup>85-87</sup>. Il NAD è necessario per la riparazione delle cellule a seguito dello stress ossidativo. La sovraregolazione di NMNAT1 ha dimostrato di proteggere contro i danni causati dalle specie dell'ossigeno reattivo nel cervello, e specialmente nell'ippocampo<sup>88,89</sup>. Anche per le variante di WCadjBMI, entrambi CDK6, vicino al nostro nuovo SNP rs10269774, e FAM49B, vicino il nostro nuovo SNP rs6470765, sono i bersagli del fattore di trascrizione BACH1, il quale è coinvolto nella risposta cellulare allo stress ossidativo e gestisce il ciclo cellulare<sup>90</sup>.

## L'INFLUENZA DEI NUOVI LOCI SULLE RELATIVE CARATTERISTICHE

Al fine di determinare se alcuni dei nuovi 26 SNPs identificati sono associati con il comportamento del fumo, abbiamo osservato il comportamento del fumo sull'esistente GWAS<sup>91</sup>. La tabella supplementare 17 mostra l'analisi e il valore di P che valuta il nostro nuovo SNPs per sempre/mai, attuali/non attuali e la quantità del fumo (SQ). Otto SNPs attraverso i tre caratteristiche dell'adiposità mostrano la prova dell'associazione al valore nominale per almeno uno dei tratti del fumo. Tuttavia, dopo diversi test di correzione ( $P < 0.05/26 = 0.0019$ ), solo un SNP rimane significativo: rs12902602, identificato per l'approccio 2(SNPjoint) e 3(SNPint) per BMI, mostrano l'associazione con SQ ( $P = 1.45 \times 10^{-9}$ ), e nell'LD (Linkage disequilibrium) con una regione conosciuta attorno al recettore della nicotina.

Abbiamo condotto una ricerca nel catalogo NHGRI-EBI GWAS<sup>92,93</sup> al fine di determinare se alcuni dei nostri loci identificati di recente hanno un alto LD con le variante associate con le caratteristiche comportamentali e cardiometabolici o le malattie. Da sette nuovi SNPs per BMI, solo rs12902602 aveva un alto LD (r elevato  $> 0.7$ ) con SNPs precedentemente associato con i tratti del fumo (dipendenza da nicotina), cancro al polmone, e malattie cardiovascolari (malattie delle coronarie) (suppl. tab.22). Dai dodici nuovi WCadjBMI di SNPs, cinque avevano un alto LD con le variazioni precedentemente sostenute di GWAS per il volume medio della piastrina, l'altezza, la lunghezza del bambino, e il melanoma. Da sei nuovi SNPs WHRadjBMI, tre erano parecchio vicino alle variante precedentemente associate, includendo le caratteristiche cardiometabolici (colesterolo LDL, trigliceridi, e la misura della funzionale renale).

In conclusione, come il fumo ha un effetto negativo sul BMI (riduzione del peso), è probabile che il fumo associato alle variante genetiche abbia un effetto sul BMI negli attuali fumatori. Di conseguenza, ci si aspetta che il fumo associato all'SNPs manifesti alcune interazioni con il fumo sul BMI. Per esaminare queste ulteriori, abbiamo guardato a tutte gli SNPs ( $P < 5 \times 10^{-8}$ ) che sono identificati nella GWAS di un fumatore che sono presenti nel catalogo NHGRI-EBI GWAS, includendo 10 variazioni nei sei loci. Due di questi loci hanno raggiunto il valore nominale ( $P < 0.05$ ) per l'interazione GxSMK sul BMI, ma solo un loco ha raggiunto il

valore corretto di Bonferroni ( $P < 0.005$ ). Nessun SNP associato col fumo mostra un'interazione col GxSMK. Di conseguenza, non abbiamo visto un forte arricchimento per la bassa interazione del valore P tra i loci precedentemente identificati col fumo.

## LE SFIDE TENENDO CONTO DELL'ESPOSIZIONE AMBIENTALE NEI GWAS

Una possibile limitazione può essere la definizione e l'armonizzazione dello status del fumo. Abbiamo scelto di stratificare basandoci solo sullo stato dei fumatori attuali senza alcuna considerazione del tipo di fumo (sigarette, pipa) per due ragioni.

La prima, focalizzandosi solo sul peso, gli ex fumatori tendono a ritornare al loro peso previsto molto velocemente a seguito della cessazione del fumo<sup>16,23,25</sup>.

La seconda, questa definizione ci permette di massimizzare una taglia campione, come altri studi coinvolti hanno solo lo status del fumo disponibili al momento della raccolta dei dati.

Abbiamo chiesto degli studi che hanno delle informazioni disponibili per escludere dei partecipanti che solo recentemente hanno smesso di fumare nel tentativo di ridurre i tratti eterogenei.

Comunque, il CV e il WHR non possono comportarsi nella stessa maniera come il peso e il BMI con gli ex fumatori immagazzinando eccessivo grasso intorno alla vita<sup>24</sup>. Considerata la differenza tra il cambiamento della circonferenza della vita e il peso per i fumatori, è possibile che i risultati possono differire leggermente se potessimo armonizzare la variabile in maniera diversa.

Un'altra limitazione può essere il potenziale per errore sistematico nei nostri effetti stimati quando adattando per un covariate nei nostri modelli di associazioni (esempio: collider bias)<sup>94,95</sup>.

Questo fenomeno è di un particolare interessamento quando la correlazione tra il risultato e la covariate è alta e quando le significative associazioni genetiche si verificano con entrambi i tratti nelle direzioni opposte. In altre parole, quando la correlazione tra il risultato e la covariate è positiva, ma le direzioni dell'effetto genetico osservato sono opposte. In questo lavoro, abbiamo adattato per il BMI e per entrambi i tratti di WCadjBMI e WHRadjBMI. WHR ha una correlazione di 0.49 con BMI, mentre CV ha una correlazione di 0.85<sup>1,95</sup>. Usando i risultati pubblicati precedentemente per BMI, WCadjBMI e WHRadjBMI, abbiamo trovato tre dei nostri nuovi loci per WCadjBMI (vicino DOCK3, ARFGEF2, TMEM38B) e due dei nostri nuovi loci per WHRadjBMI (vicino EHMT<sup>2</sup>, HLA-C) con significative associazioni con BMI e con direzione opposta dell'effetto. A questi loci, l'effetto genetico stimato dovrebbe essere interpretato con prudenza. In aggiunta, abbiamo adattato per SMK in l'approccio 1 (SNPadjSMK). Comunque binari fumo status che è stato utilizzato nelle attuali analisi ha una bassa correlazione da BMI, WC e WHR, come stimato negli studi di ARIC con i discendenti europei come partecipanti (rispettivamente - 0.13, 0.08, e 0.12) e in Framingham Heart Study (-0.05, 0.08, 0.16).

Inoltre, non ci sono loci identificati nell'approccio1 che sono associati con alcuni comportamenti del fumo e che esibisce una direzione opposte di effetto da quello identificato nei nostri tratti dell'obesità. Perciò preclude l'errore sistematico e acquisendo una giusta ipotesi forzando i loci presenti nel adattamento di SMK.

Al fine di stimare quanta informazione aggiuntiva è fornita tenendo conto di SMK e GxSMK nel GWAS per i tratti dell'obesità, abbiamo calcolato e confrontato i risultati del rischio genetico (GRSs) basati su varie sottocategorie dei principali genotipi di SNP e basati sui vari modelli di regressione. Mentre alcuni GRS erano associati con la caratteristica obesità ( $P < 1.6 \times 10^{-7}$ ), aggiungendo i termini SMK e GxSMK al modello di regressione includendo insieme alle nuove varianti nel calcolo del GRSs notevolmente aumentata la variazione esplicata. Per esempio, la variazione esplicata aumentata del 38% per BMI (da 1.53% a 2.11%,  $P = 4.3 \times 10^{-5}$ ), da 27% per WCadjBMI (da 2.59% a 3.29%,  $P = 3.9 \times 10^{-6}$ ) e da 168% per WHRadjBMI (da 0.82% a 2.20%,  $P = 3.2 \times 10^{-11}$ ). Pertanto, nonostante le limitazioni

potenziali, c'è molto da acquisire tenendo conto dell'esposizione ambientale negli studi nel GWAS.

## DISCUSSIONE

Per aiutarci a capire gli effetti del fumo sulla suscettibilità genetica all'obesità, valutata da BMI, WCadjBMI, e WHRadjBMI, abbiamo condotto una serie di analisi per scoprire le variazioni genetiche che possono essere mascherate quando l'influenza ambientale del fumo non viene considerata, e per scoprire dei loci genetici che interagiscono con il fumo per influenzare i tratti relativi all'obesità. Abbiamo identificato 161 loci in totale, includendo 23 nuovi loci che influenzano i tratti dell'obesità (6 per BMI, 11 per WCadjBMI, e 6 per WHRadjBMI). Mentre molti dei nostri loci identificati di recente supportano l'ipotesi che il fumo possa influenzare la fluttuazione del peso attraverso la regolazione dell'appetito, questi nuovi loci altresì hanno evidenziato nuovi processi biologici e vie metaboliche implicate nella patogenesi dell'obesità.

Soprattutto, abbiamo identificato nuovi dieci loci con la prova plausibile dell'interazione del GxSMK sui tratti relativi all'obesità. Siamo in grado di replicare la precedente interazione di GxSMK con i risultati di BMI entro il cluster genico CHRNA5-CHRNA3-CHRNA4.

Ulteriormente, un nuovo loco associato con BMI vicino INPP4B e due nuovi loci associati con WCadjBMI vicino a GRIN2A e PRNP mostrano una interazione significativa di GxSMK. Mediante l'uso dei metodi fortemente analitici, siamo in grado di identificare l'interazione significativa GxSMK per un conosciuto loco associato al BMI vicino a ADAMTS7 in aggiunta a cinque conosciuti loci associati a WHRadjBMI vicino a LYPLAL1, RSPO3, MAP3K1, HOXC4-HOXC6 e JUND. I nostri risultati mostrano che la maggioranza di questi loci nascondono candidati geni con la prova per influenzare adiposità totale, e anche la prova che indica un possibile ruolo per la modulazione degli effetti attraverso l'uso del tabacco.



Abbiamo identificato 18 nuovi loci nell'approccio 1 mediante l'adattamento dello stato degli attuali fumatori. Le nostre analisi non ci hanno permesso di determinare se queste scoperte sono dovute alle differenti sottocategorie dei soggetti inclusi nell'analisi comparata agli studi precedenti<sup>6,7</sup> o dovute all'adattamento per gli attuali fumatori.

L'adattamento per gli attuali fumatori su questi loci delle nostre analisi, però, rivelano le associazioni che sono state identificate precedentemente. Specificamente dopo aver tenuto conto dei fumatori nelle nostre analisi, tutti i nuovi loci di BMI mostrano i valori P hanno almeno un ordine di grandezza in meno rispetto alla precedente indagine di GIANT (Genome-scale Integrated Analysis of gene Networks in Tissues), nonostante la dimensione comune è più piccola nelle analisi attuali<sup>7</sup>.

Mentre le dimensioni dei campioni per entrambi WCadjBMI e WHRadjBMI sono più paragonabili con la precedente indagine di GIANT, i nostri valori P per la variante identificato dell'approccio 1 hanno almeno due ordini di grandezza meno rispetto ai risultati precedenti. Perciò, l'adattamento per i fumatori può avere rivelare nuovi loci. Ulteriormente, i loci identificati nell'approccio 2, includendo nove nuovi loci, suggeriscono che tenendo conto dell'interazione migliora la nostra abilità di localizzare questi loci anche in presenza di una sola prova dell'interazione di GxSMK.

Abbiamo esaminato se ci fossero degli effetti più forti nei fumatori attuali rispetto ai non fumatori. Mentre abbiamo trovato che gli effetti non erano significativamente più forti negli attuali fumatori rispetto ai non fumatori per BMI, c'è un rapporto maggiore della variazione del BMI spiegato dalle varianti che sono significative per l'approccio 1, il quale può essere atteso dato che ci sono un più alto numero di varianti con più alto effetto stimato negli attuali fumatori. Per WCadjBMI, non c'era arricchimento per effetti più forti in uno strato rispetto agli altri anche per i nostri loci significativi; comunque, c'era una grande proporzione della variazione nel WHRadjBMI per i loci identificati nell'approccio 1 nei non fumatori rispetto ai fumatori. Per WHRadjBMI, c'erano più loci che mostravano un effetto maggiore nei non fumatori rispetto ai fumatori, e questo

schema era rispecchiato nella variazione dell'analisi esplicita. La grande differenza tra l'effetto nei fumatori e nei non fumatori spiega presumibilmente il sottolivello di GWS dei nostri loci delle indagini precedenti del GIANT<sup>7</sup>. Per esempio, l'allele T di rs7697556, 81kb dal gene di ADAMTS3, era associato con l'aumento di WCadjBMI e mostra un effetto sei volte maggiore nei non fumatori rispetto ai fumatori, benché l'effetto di interazione era solo nominale; nelle indagini precedenti di GWAS da Locke et al. Questa variante era vicina a GWS. Queste differenze nell'effetto stimato tra i fumatori e i non fumatori attraverso i loci può aiutare a spiegare le ricerche discordanti nelle analisi precedenti che mostrano l'aumento dell'adiposità addominale con l'aumento del fumo, ma è associato con la diminuzione del peso e di BMI<sup>13,20,21,96</sup>.

I nostri risultati supportano le ricerche precedenti che riguardano i geni coinvolti nella trascrizione e nell'espressione del gene, la regolazione dell'appetito, il metabolismo macronutriente, e l'omeostasi del glucosio.

Diversi dei nostri nuovi loci hanno candidato geni entro i 500kb delle nostre varianti che sono altamente espresse e/o attive nel tessuto cerebrale (BBX, KIF1B, SOX11, e EPHA3) e, come altri geni associati all'obesità, possono essere coinvolti nelle precedenti identificate vie metaboliche connesse alla regolazione neuronale dell'appetito. (KIF1B, GRIN2A, e SLC23A2), adiposo/angiogenesi (ANGPYL3 e TNF) e dell'omeostasi del glucosio, dei lipidi e dell'energia (CD47, STK25, STK19, RAGE, AIF1, LYPLAL1, HDLBP, ANGPTL3, DOCK7, KIF1B, PREX1, e RPS12).

Comunque, molti dei nostri loci identificati di recente evidenziano nuove funzioni biologiche e vie metaboliche dove la disregolazione può aumentare la suscettibilità dell'obesità, includendo le risposte allo stress ossidativo, al comportamento dipendente, alle funzioni regolatorie identificate di recente. C'è un numero crescente di prove che supportano l'idea che l'esposizione allo stress ossidativo porta all'aumento dell'adiposo, rischio di obesità, e ridotto risultato metabolico<sup>56,87,100</sup>. I nostri risultati per BMI e WCadjBMI, specificamente le associazioni vicine a CHRNA5- CHRNA3- CHRNB4, PRNP, SLC23A2, BACH1 e NMNAT1, evidenziano le nuove vie metaboliche biologiche e i processi per le

future esaminazioni e possono guidare verso una buona comprensione su come lo stress ossidativo porta a delle alterazioni nei fenotipi dell'obesità e i rischi cardiometabolici.

Per l'incorporazione dei fumatori attuali nelle nostre analisi, siamo in grado di identificare sei nuovi loci per BMI, 11 per WCadjBMI, e 6 per WHRadjBMI, ed evidenzia nuovi processi biologici e funzioni regolatorie per i geni implicati nel rischio dell'aumento dell'obesità.

Molti dei loci precedentemente identificati erano confermati nelle nostre analisi con l'adattamento per lo stato degli attuali fumatori con la più piccola dimensione dei campioni che erano necessari nelle analisi precedentemente scoperte.

Un approccio tipico in larga scala delle analisi del GWAS non si modifica per la covariate come negli attuali fumatori; comunque, i nostri risultati evidenziano l'importanza di tener conto dell'esposizione ambientale nelle analisi genetiche.

## CONCLUSIONI

Lo studio ha evidenziato la presenza di molteplici loci che codificano una suscettibilità genetica per l'insorgenza dell'obesità nella sottopopolazione dei fumatori, inserendo un ulteriore tassello del mosaico della stratificazione del rischio cardiovascolare e aprendo la strada ad indagini future.

## BIBLIOGRAFIA

- I. **B. Caballero**, *The global epidemic of obesity* in *Epidemiol Rev.*, vol 29, pagg. 1-5, 2007.
- II. *Epidemiologia dell'obesità- Il quadro internazionale ed europeo*, 2013
- III. **G. Faglia, P. Beck- Peccoz, A. Spada**, *Malattie del sistema endocrino e del metabolismo*, pag. 602, Milano, 2013.
- IV. **G. Faglia, P. Beck- Peccoz, A. Spada**, op. cit., pag.605
- V. **G. Faglia, P. Beck- Peccoz, A. Spada**, op. cit, pag. 603.
- VI. **J. Sobal, A.J. Stunkard**, *Socioeconomic status and obesity: A review of the literature*, in *Psychol Bull*, vol.105, n.2, pagg. 260- 75, 1989
- VII. **National Institutes of Health**, *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults*, Public Health Service, 1998.
- VIII. **G. Faglia, P. Beck- Peccoz, A. Spada**, op.cit. , pag.597.
- IX. **M- Vanoli, P. Riboldi**, *Principi di medicina interna (Harrison)*, pag. 911, Milano, 2009.
- X. **G. Faglia, P. Beck- Peccoz, A. Spada**, op.cit., pag.607.
- XI. **DW. Haslam e WP. James**, cit.
- XII. **I e D. Hoffman**, *The chancing cigarette: Chemical Studies and bioassays in Risks Associated with Smoking Cigarette*, pagg. 159-191, Bethesda, 2001.
- XIII. **Indagine Doxa-Iss**, 2014.
- XIV. **Yusuf S., Hawken S.**, *Effect of potentially modifiable risk factors associated with myocardial infarction*, pag. 364, 2004.
- XV. **E. Staffiere**, *Il fumo come fattore di rischio cardiovascolare e nuove strategie di trattamento* in *Giornale di Cardiologia*, pag. 15s, 2010.
- XVI. **Health and Social Care Information Centre (HSCIC)**, *Lifestyles Statistics. Statistics on Smoking*, England, 2012.
- XVII. **A. E. Locke**, et al. *Genetic studies of body mass index yield new insights for obesity biology*, *Nature* 518, 197-206, 2015.

XVIII. **C. Clair** et al. *Dose-dependent positive association between cigarette smoking, abdominal obesity and body fat: cross-sectional data from a population based survey*, BMC Public Health 11, 23,2011.

XIX. **B. J. Nicklas, N. Tomoyasu, J. Muir & A.P. Goldberg**, *Effects of cigarette smoking and its cessation on body weight and plasma leptin levels*, Metabolism 48, 804-8, 1999.

XX. **B. J. Nicklas, N. Tomoyasu, J. Muir & A.P. Goldberg**, cit.

XXI. **C. K. Weeckley, 3rd, R.C. Klesges & G. Reylea**, *Smoking as a weight control strategy and its relationship to smoking status*, Addict Behav 17, 259- 71, 1992.

1. Department of Epidemiology, University of North Carolina, Chapel Hill, NC 27599.
2. Department of Genetic Epidemiology, Institute of Epidemiology and preventive Medicine, University of Regensburg, D-93053 Regensburg, Germany.
3. Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine; St. LOUIS, MO, 63108 USA.
4. Department of biostatistics, Boston University School of public Health, Boston, MA 02118.
5. Population health research institute, st. Georg's, University of London,London, SW17 ORE, UK.
6. TransMed Systems, inc, Cupertino, CA 95014.
7. Department of Biostatistics ,Johns Hopkins Bloomberg school of public health, Baltimore MD.
8. The Novo Nordisk Foundation Center of Basic Metabolic Research, section of metabolic genetics, Faculty of Health and medical Sciences, University of Copenhagen, Copenhagen, Denmark.
9. NHLBI Framingham Heart Study, Framingham, MA,01702 USA.
10. Division of Preventive Medicine,Brigham and womens hospital and Harvard medical school, Boston,MA USA
11. Department of Neurology, Boston University School of Medicine, Boston, MA 02118, USA.

12. Population Sciences Branch, National Heart, Lung, and blood institute, National Institutes of Health, The Framingham Heart Study , Framingham, MA,USA.
13. Institute of Social and Preventive Medicine, Center Hospitalier University Vaudois, Lausanne,Switzerland.
14. Department of medical genetics, University of Lausanne, Lausanne, Switzerland.
15. Swiss institute of bioinformatics.
16. Wellcome Trust Center for Human Genetics, University of Oxford, Oxford, OX37BN, UK.
17. Department of Biobank Research, Umea University, Umea, Sweden.
18. .Department of clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University, skan University Hospital Malmo,SE 205-02 MALMO, Sweden.
19. Department of Epidemiology and population health, Albert Einstein Colleg of medicine, Bronx, NY, USA.
20. The Charles Bronfman institute for Personalized Medicine, Icahn School of medicine at Mount Sinai, New York, NY, USA.
21. The genetics of Obesity and Related Metabolic Traits Program, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
22. Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, 3015GE, The Netherlands.
23. Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA.
24. University of Lille, CNRS, Institut Pasteur of Lille, UMR 8199 - EGID, Lille, France.
25. Internal Medicine - Nephrology, University of Michigan, Ann Arbor, Michigan, USA.
26. Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 48109, USA.
27. Centre for Genetic Origins of Health and Disease, University of Western Australia, Crawley, Australia.

28. Dept. Health Sciences, University of Milan, Via A. Di Rudiní, 8 20142, Milano, Italy.
29. Department of Biostatistics, University of Washington, Seattle, WA 98195.
30. . Department of Epidemiology, Erasmus Medical Center, Rotterdam, Netherlands.
31. Department of Psychiatry, Dokuz Eylul University, Izmir, Turkey.
32. Centre for Bone and Arthritis Research, Department of Internal Medicine and Clinical Nutrition, University of Gothenburg, Sweden.
33. Estonian Genome Center, University of Tartu, Tartu 51010, Estonia.
34. Department of Nephrology, University Hospital Regensburg, Regensburg, Germany.
35. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, OX3 7LJ, UK.
36. Epidemiology Domain, Saw Swee Hock School of Public Health National University of Singapore, Singapore 117549.
37. Department of Nutrition, Harvard School of Public Health, Boston, MA 02115, USA.
38. MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland.
39. . William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK.
40. Department of Health, National Institute for Health and Welfare, Helsinki, FI-00271 Finland.
41. Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany.
42. Kuopio Research Institute of Exercise Medicine, Kuopio, Finland.
43. ISER, University of Essex, Colchester, Essex, UK CO43SQ.
44. Department of Epidemiology and Public Health, UCL, London, UK. WC1E 6BT.
45. Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI 96813, USA.

46. . MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge, CB2 0QQ, UK.
47. Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland.
48. Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere 33014, Finland.
49. Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.
50. NIHR Biomedical Research Centre at Guy's and St. Thomas' Foundation Trust, London, UK.
51. Center for Public Health Genomics and Biostatistics Section, Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia 22903.
52. Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
53. Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, D-85764 Neuherberg, Germany.
54. Department of Medicine I, Ludwig-Maximilians-Universität, D-81377 Munich, Germany.
55. DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany.
56. Department of Kinesiology, Faculty of Medicine, Université Laval, Québec, Canada.
57. Institute of Nutrition and Functional Foods, Université Laval, Québec, Canada.
58. Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Tartu 51010, Estonia.
59. Department of Social and Health Care, City of Helsinki, Helsinki, Finland.
60. Icelandic Heart Association, Kopavogur, Iceland.
61. Faculty of Medicine, University of Iceland, Reykjavik, Iceland.
62. Department of Medicine, Institute of Clinical Medicine, University of Eastern Finland, 70210 Kuopio, Finland.



63. Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden.
64. Center for Molecular Medicine, Karolinska University Hospital Solna, Stockholm, Sweden.
65. Division of Biostatistics, Washington University School of Medicine, St Louis, MO.
66. Translational Gerontology Branch, National Institute on Aging, Baltimore MD, USA.
67. Institute for Community Medicine, University Medicine Greifswald, Germany.
68. Department of Cardiology, Leiden University Medical Center, The Netherlands.
69. Department of Gerontology and Geriatrics, Leiden University Medical Center, The Netherlands.
70. Laboratory of Experimental Cardiology, Department of Cardiology, Division Heart & Lungs, UMC Utrecht, the Netherlands.
71. Department of Epidemiology, University of Groningen, University Medical Center Groningen, The Netherlands.
72. Department of Endocrinology, University of Groningen, University Medical Center Groningen, the Netherlands.
73. Divisions of Endocrinology and Genetics and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston MA 02115 USA.
74. Broad Institute of Harvard and MIT, Cambridge, MA 02142 USA.
75. Department of Cardiology, University Medical Center Groningen, University of Groningen, the Netherlands.
76. Department of Biological Psychology, Vrije Universiteit, Amsterdam, the Netherlands.
77. NCA, Vrije Universiteit & Vrije Universiteit Medical Center, Amsterdam, the Netherlands.
78. Department of Genetics, University of North Carolina, Chapel Hill, NC 27599 USA.

79. Dept Epidemiology and Biostatistics, School of Public Health, Imperial College London, UK.
80. Cardiology, Ealing Hospital NHS Trust, Middlesex, UK.
81. . Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle WA USA.
82. Department of Nutrition, University of North Carolina, Chapel Hill, NC 27599.
83. Department of Medicine, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands.
84. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA 98101.
85. Busselton Population Medical Research Institute, Nedlands, WA 6009, Australia.
86. PathWest Laboratory Medicine of WA, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
87. School of Pathology and Laboratory Medicine, The University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia.
88. . Diabetes and Obesity Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA.
89. Clinic for Prosthetic Dentistry, Gerostomatology and Material Science, University Medicine Greifswald, Germany
90. South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley, Brownsville, TX.
91. Human Genetics Center, The University of Texas Health Science Center, PO Box 20186, Houston, TX 77225.
92. Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA.
93. Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
94. . Department of Pharmacology and Therapeutics, University College Cork, Ireland.
95. Department of Genetics, Rutgers University, Piscataway, NJ 08854, USA.

96. Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ 08854, USA.
97. Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Scotland, UK.
98. Imperial College Healthcare NHS Trust, London, UK.
99. Department of Vascular Surgery, Division of Surgical Specialties, UMC Utrecht, the Netherlands.
100. EMGO+ Institute Vrije Universiteit & Vrije Universiteit Medical Center.
101. Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece.
102. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA.
103. Robertson Center for Biostatistics, University of Glasgow, United Kingdom.
104. Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Mona, JMAAW15 Jamaica.
105. Unit of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.
106. Hypertension and Related Disease Centre, AOU-University of Sassari.
107. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany.
108. Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, D-85764 Neuherberg, Germany.
109. German Center for Diabetes Research, D-85764 Neuherberg, Germany.
110. Department of Public Health and Clinical Medicine, Section for Nutritional Research, Umeå University, Umeå, Sweden.
111. Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, National Institutes of Health, Bethesda, MD.
112. Department of Psychiatry, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands.
113. Department of Psychiatry, Washington University School of Medicine, St. Louis, MO.

114. Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA.
115. Division of Genomic Medicine, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892.
116. Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen, UK, AB25 2ZD.
117. Generation Scotland, Centre for Genomic and Experimental Medicine, University of Edinburgh, Edinburgh, Scotland.
118. St. Olav Hospital, Trondheim University Hospital, Trondheim, Norway.
119. Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Germany.
120. Department of Human Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK.
121. School of Medicine and Pharmacology, The University of Western Australia, 25 Stirling Hwy, Crawley, WA 6009, Australia.
122. Department of Cardiovascular Medicine, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
123. Department of Pediatrics, Tampere University Hospital, Tampere 33521, Finland.
124. Department of Pediatrics, University of Tampere School of Medicine, Tampere 33014, Finland.
125. Department of Medical Sciences, Molecular Epidemiology, Uppsala University, Uppsala, 751 85, Sweden.
126. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA.
127. Science for Life Laboratory, Uppsala University, Uppsala, 750 85, Sweden.
128. Department of Pulmonary Physiology and Sleep Medicine, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
129. Department of Physiology, Institute of Neuroscience and Physiology, the Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden.

130. Department of Epidemiology and Biostatistics, MRC–PHE Centre for Environment & Health, School of Public Health, Imperial College London, UK.
131. Center for Life Course Epidemiology, Faculty of Medicine, P.O.Box 5000, FI-90014 University of Oulu, Finland.
132. Biocenter Oulu, University of Oulu, Finland.
133. Unit of Primary Care, Oulu University Hospital, Kajaanintie 50, P.O.Box 20, FI-90220 Oulu, 90029 OYS, Finland.
134. Steno Diabetes Center, Gentofte, Denmark.
135. Department of Medicine, University of Turku, Turku 20520 Finland.
136. Division of Medicine, Turku University Hospital, Turku 20521, Finland.
137. Department of Clinical Physiology, Tampere University Hospital, Tampere 33521, Finland.
138. Department of Clinical Physiology, University of Tampere School of Medicine, Tampere 33014, Finland.