



UNIVERSITÀ DEGLI STUDI DI CAGLIARI
FACOLTÀ DI MEDICINA E CHIRURGIA
DIPARTIMENTO DI SANITA' PUBBLICA, MEDICINA CLINICA E MOLECOLARE

Ciclo XXV_ DOTTORATO DI RICERCA IN

“Terapia Pediatrica e Farmacologia dello Sviluppo”

*Studio dei tratti immunofenotipici e valutazione del
loro contributo nell'insorgenza di patologie*

Coordinatore Scientifico:

Prof. Paolo Moi

Tutori:

Dott.ssa Valeria Orrù

Dott. Edoardo Fiorillo

Dottoranda:

Dott.ssa Francesca Viridis

Relatore: **Prof. Francesco Cucca**

Anno accademico 2011–2012

Settore scientifico disciplinare di afferenza MED/06

INDICE

Introduzione	3
Disegno sperimentale dello studio	5
Sezione 1	6
1.1 Le Popolazioni isolate	6
1.2 Il Progetto ProgeNIA	7
Sezione 2	8
2.1 Immunità innata e Immunità adattativa	8
2.2 Origine dei componenti del sistema immunitario	9
2.3 Tolleranza Immunologica e Autoimmunità	10
2.4 Popolazioni cellulari oggetto di studio	11
a) Cellule T, B e natural killer	11
b) Cellule T regolatorie	12
c) Maturazione T cellulare	13
d) Cellule dendritiche	14
Sezione 3	15
3.1 Principi di citometria a flusso	15
a) Sistema fluidico	16
b) Sistema ottico: ottica di eccitazione, emissione e rilevazione	17
c) Sistema elettronico	17

Sezione 4	18
Materiali e metodi	18
4.1 Misurazioni citofluorimetriche	18
a) Pannello T-B-NK	18
b) Pannello Treg	19
c) Pannello della Maturazione T cellulare	20
d) Pannello DC	20
4.2 Settaggio strumentale dei citofluorimetri	21
4.3 Riproducibilità dati FACS	22
4.4 Ereditabilità	23
4.5 Genotipizzazione con Gene chip arrays	23
4.6 Sequenziamento dei campioni	24
4.7 Imputazione genotipica	24
4.8 Analisi statistiche e bioinformatiche	24
4.9 Ricerca delle associazioni coincidenti con i disturbi complessi	26
Sezione 5	26
Risultati	26
5.1 Stime Ereditabilità	26
5.2 Associazioni genetiche	27
5.3 Associazioni coincidenti tra tratti immunofenotipici e malattie	29
Conclusioni	32
Tabelle e Figure	33
Referenze Bibliografiche	65
Ringraziamenti	71

Introduzione

Il sistema immunitario è costituito da un insieme specializzato di organi, tessuti, cellule e molecole, la cui funzione principale è proteggere l'organismo dall'attacco di agenti patogeni. Tale sistema è caratterizzato dalla capacità di distinguere le strutture appartenenti all'organismo (*self*) da quelle estranee ad esso (*non self*) in modo che la risposta immunitaria venga innescata unicamente nei confronti di queste ultime. Tuttavia, l'alterazione dei normali meccanismi di tolleranza immunologica attuati verso le strutture *self* può provocare una reazione diretta contro le cellule ed i tessuti dello stesso organismo; questo fenomeno è chiamato autoimmunità. Le malattie autoimmuni rappresentano un serio problema sanitario spesso privo di cure risolutive; esistono infatti circa 90 diversi tipi di patologie autoimmuni, che nel loro insieme colpiscono il 5% della popolazione europea e, malgrado questo dato allarmante, solo il 15-20% dei lavori scientifici pubblicati in riviste mediche di rilevanza internazionale ha come interesse specifico l'immunologia umana [Davis M.M. 2008]. Numerosi dati sperimentali hanno evidenziato una forte componente genetica alla base dell'insorgenza delle patologie autoimmuni spesso correlate a variazioni del numero o della funzione di specifici sottoinsiemi di cellule linfocitarie. Ne è un esempio un recente studio effettuato dal gruppo del Prof. Todd che ha dimostrato che un polimorfismo genetico associato al diabete di tipo 1 (T1D), localizzato a nel gene che codifica per la catena α del recettore dell'interleuchina 2 (*IL-2Ra*, CD25), altera l'espressione di tale recettore in specifiche popolazioni leucocitarie, modificando in tal modo l'omeostasi del sistema immunitario. Un ulteriore prova del ruolo della genetica sulla regolazione del sistema immunitario è dimostrato da un polimorfismo codificante localizzato nel gene *PTPN22* che, oltre ad essere associato a numerose patologie autoimmuni, rappresentare un fattore prognostico per la progressione della leucemia linfatica cronica [Negro R. *et al.*, 2012]. Negli ultimi anni è stato fatto uno sforzo a livello internazionale, per colmare le lacune riguardo l'immunologia umana, infatti nel 2010 è stato fondato il Human Immunology Consortium Project che ha il fine di caratterizzare i meccanismi che regolano il sistema immunitario sia in condizioni di salute che patologiche.

Tuttavia, lo studio delle popolazioni cellulari, quando effettuato in modo multicentrico, richiede dei processi di standardizzazione complessi che limitano i fenotipi cellulari studiati alle popolazioni immunologiche maggiormente rappresentate e più facilmente catalogabili, senza la possibilità di analizzare popolazioni specifiche e meno frequenti che sono spesso maggiormente implicate nella patogenesi delle malattie autoimmuni [Maecker H.T. *et al.*, 2012].

Da questo contesto nasce l'esigenza di concentrare i nostri sforzi sullo studio delle cellule del sistema immunitario ed in particolare sulla definizione della componente genetica in grado di regolare i livelli di tali cellule. Lo studio delle componenti genetiche delle malattie complesse è stato rivoluzionato nel corso degli ultimi anni grazie all'avvento degli studi di associazione sull'intero genoma (GWAS). Questo metodo di analisi ad alta risoluzione, combinato con nuovi approcci quali il sequenziamento dell'intero genoma e con adeguati studi funzionali, sta determinando un rapido avanzamento delle conoscenze sui meccanismi e sulle vie patogenetiche coinvolte in numerose patologie comuni. In virtù di questo approccio genetico, il nostro studio prevede l'incrocio dei dati genetici di individui afferenti alla coorte ProgeNIA con il loro profilo immunologico (o tratti immunofenotipici) ottenuto mediante l'analisi quantitativa di diverse popolazioni cellulari appartenenti ai leucociti circolanti nel sangue periferico. Attraverso uno studio di associazione sull'intero genoma, per ogni tratto immunologico considerato verranno evidenziate le varianti genetiche che ne regolano i livelli fenotipici. Tali varianti saranno successivamente ricercate in database pubblici per verificare la loro coincidenza con varianti associate a patologie già descritte in letteratura. L'associazione coincidente fra varianti associate a malattia e varianti associate ad uno o più tratti immunofenotipici suggerirà che quel tratto è implicato nel meccanismo patogenetico. La comprensione dei fattori eziopatogenetici alla base della suscettibilità nei confronti delle malattie offre un'opportunità senza precedenti per individuare potenziali bersagli terapeutici e generare nuovi farmaci. Il nostro studio è quindi distinto da numerosi altri studi, in quanto il disegno sperimentale non prevede il confronto dell'assetto immunologico tra pazienti e controlli [Marwaha AK. *et al.*, 2010], bensì considera un campione di individui appartenenti alla popolazione generale. Il fine di questo approccio consiste nell'evitare gli effetti confondenti dovuti alla patologia o al suo trattamento farmacologico che potrebbero alterare quantitativamente i tratti immunofenotipici esaminati.

Disegno sperimentale dello studio

Il disegno sperimentale del presente studio prevede la caratterizzazione sia immunofenotipica, mediante citometria a flusso, che genetica, mediante chip e sequenziamento, di individui sardi appartenenti alla coorte ProgeNIA. L'obiettivo principale è quello di identificare le varianti genetiche che regolano i livelli di numerose popolazioni cellulari circolanti appartenenti all'immunità innata ed adattativa per un totale di 272 tratti immunofenotipici che derivano dalla stima delle conte cellulari assolute espresse in numero di cellule per microlitro e dalla percentuale di una popolazione nei confronti della linea cellulare gerarchicamente superiore sia di una generazione (popolazione parentale) che di due generazioni (popolazione granparentale) (Figura 2). Per chiarezza consideriamo l'esempio delle cellule T citotossiche, che sono state valutate come numero di cellule per microlitro, come percentuale rispetto alle cellule T (popolazione parentale) e rispetto al totale dei linfociti (popolazione granparentale). Tale quantificazione delle popolazioni leucocitarie rispetto alle popolazioni progenitrici è stata effettuata allo scopo di fornire una misurazione più accurata e robusta dei livelli cellulari esaminati diminuendo, per contro, la variabilità nella misurazione dovuta a fattori operatore-dipendente e fattori influenzati dal microambiente circostante. Inoltre, poiché vi è un diretto equilibrio omeostatico tra cellule T e B e tra cellule T helper e cellule T citotossiche, sono stati considerati come dato immunofenotipico anche tali rapporti. Parallelamente, gli individui immunofenotipizzati sono stati caratterizzati genotipicamente mediante l'utilizzo di due saggi custom Illumina, MetaboChip ed ImmunoChip, [Cortes A. and Brown M.A., 2011; Voight B.F. *et al.*, 2012] e attraverso il sequenziamento dell'intero genoma di 1146 individui sardi, grazie al quale è stato possibile testare un totale di ~8.2M varianti genetiche.

La prima analisi genetica effettuata sui tratti immunofenotipici è rappresentata dalla stima dell'ereditabilità (h^2) che necessita della valutazione del tratto sulle famiglie verticali presenti nella coorte ProgeNIA. Tale analisi ha permesso di comprendere quanto della variabilità fenotipica osservata fosse geneticamente determinata. Successivamente, per ciascun tratto immunofenotipico quantificato è stato effettuato uno studio di associazione sull'intero genoma che ha permesso di identificare le varianti genetiche associate alle variazioni dei tratti analizzati. In dettaglio, sono state selezionate le varianti indipendenti per ciascun tratto ($r^2 < 0.1$)

che hanno mostrato un p-value minore di 5.26×10^{-10} . Tale valore corrisponde alla significatività statistica universalmente accettata per gli studi GWAS, che è pari a 5×10^{-8} , corretta con il metodo di Bonferroni per 95 test indipendenti (che corrispondono al numero di popolazioni considerate come conta assoluta). Questi primi risultati di associazione, effettuati su 1629 individui, sono stati poi confermati sfruttando un campione esteso che comprendeva 1241 individui addizionali per un totale di 2870 campioni.

Poiché spesso i segnali di associazione non cadono all'interno di geni, è stato necessario applicare una strategia per individuare i geni, prossimali alle varianti associate, biologicamente rilevanti sulla base sia della caratteristica del tratto immunofenotipico in oggetto che della funzione biologica delle varianti (o dei geni) associate ai tratti. Infine, abbiamo ricercato quali varianti genetiche associate ai tratti immunologici coincidessero o fossero in forte LD ($r^2 > 0.8$) con le varianti associate a patologie descritte in database pubblici, come il GWAS catalogue [Hindorff L.A.: www.genome.gov/gwastudies] ed ImmunoBase [<http://www.immunobase.org/>].

Sezione 1

1.1 Le popolazioni isolate

Le popolazioni isolate, come i volontari appartenenti alla coorte ProgeNIA, sono rappresentate da un gruppo di individui che hanno avuto origine da un ridotto numero di soggetti fondatori che per un certo numero di generazioni hanno vissuto in una condizione di isolamento geografico o culturale. Tale condizione ha portato ad una diminuzione dei flussi immigratori, ad un'elevata percentuale di matrimoni fra consanguinei con una conseguente riduzione del contributo genetico apportato dall'esterno. Inoltre, l'isolamento genetico può derivare da ulteriori fattori, il più importante dei quali sono le barriere geografiche che tendono a ridurre la possibilità di spostamento degli individui e quindi a creare le condizioni ideali per l'isolamento. Le popolazioni isolate costituiscono uno strumento valido per lo studio delle basi genetiche delle malattie complesse, indispensabile per la comprensione dei loro meccanismi patogenetici. La Sardegna rappresenta un esempio di popolazione isolata: essa è caratterizzata da condizioni ambientali, quali il clima e le abitudini alimentari, molto simili in tutto il territorio, è una popolazione omogenea che ha vissuto per un lungo periodo in una condizione di

isolamento genetico che ha favorito la formazione di un differente ed unico pool genetico rispetto a quello presente in altre popolazioni che vivono in zone circostanti. L'unicità del patrimonio genetico della popolazione Sarda, e in particolare di alcune aree della provincia dell'Ogliastra, risiede nella presenza di alcune varianti genetiche originali, in quanto già presenti negli individui che popolavano la Sardegna dal paleolitico superiore in poi (14.000 anni fa) ed al contempo molto più rare o assenti nelle altre popolazioni. In conclusione, i geni o le varianti di suscettibilità per malattie complesse comuni sono più facilmente identificabili se si sceglie di studiare "le popolazioni geneticamente isolate" [Kristiansson K. *et al.*, 2008] caratterizzate da un'alta incidenza di malattie comuni e portatrici di varianti predisponenti per tali patologie.

1.2 Il progetto ProgeNIA

ProgeNIA è uno studio longitudinale che si propone di studiare, su una coorte di individui di origine sarda, numerosi parametri biochimici ed antropometrici di interesse biomedico al fine di identificare le varianti genetiche in grado di influenzarne i relativi valori [Pilia G. *et al.*, 2006]. La coorte reclutata nel progetto comprende circa 7000 volontari, di età compresa tra i 14 ed i 102 anni, che vivono in 4 paesi della provincia dell'Ogliastra (Lanusei, Arzana, Ilbono ed Elini). Dal 2001 ogni individuo si sottopone volontariamente ad accurate visite mediche che includono l'analisi dell'apparato cardiovascolare e muscolo-scheletrico oltre che l'esame emocromocitometrico. Pur nascendo con l'obiettivo di ricercare i geni coinvolti nei processi dell'invecchiamento, dall'anno 2010, l'analisi dei tratti fenotipici ematici e clinici della coorte ProgeNIA è stata implementata con la valutazione di oltre 100 popolazioni cellulari immunologiche circolanti nel sangue periferico e rilevanti nell'insorgenza di patologie a carico del sistema immunitario come il diabete di tipo 1 e la sclerosi multipla, malattie particolarmente diffuse in Sardegna. Oltre a possedere i dati fenotipici di ciascun volontario, il centro ProgeNIA dispone anche dei loro profili genetici ottenuti mediante l'utilizzo di diverse piattaforme di tipizzazione. Tali piattaforme permettono di esaminare simultaneamente fino a diverse centinaia di migliaia di polimorfismi a singolo nucleotide (SNPs) distribuiti lungo tutto il genoma di un individuo. Numerosi studi di associazione condotti su questa coorte hanno portato all'identificazione di

diverse varianti genetiche implicate nella regolazione della maggior parte dei tratti esaminati nello studio. Quest'ultimo dato conferma che gli studi genetici effettuati su una popolazione isolata sono ideali per individuare effetti genetici che si trasmettono attraverso le generazioni.

Sezione 2

2.1 Immunità innata e Immunità adattativa

Il sistema immunitario è un complesso di organi, tessuti, cellule e molecole, il cui compito consiste nel proteggere l'organismo dall'attacco di patogeni, quali batteri, virus, elminti, o molecole tossiche. Tale sistema ha la capacità di riconoscere le molecole appartenenti all'organismo da quelle estranee ad esso, e di innescare una risposta immune che coinvolge una rete articolata di popolazioni cellulari che interagiscono tra loro in modo sinergico e il cui comune obiettivo è eliminare le sostanze estranee che infettano il nostro organismo. Ciascuna popolazione cellulare richiede il differenziamento e l'attivazione di cellule effettrici, essenziali per l'individuazione e l'eliminazione di patogeni, e di cellule regolatorie, necessarie per il mantenimento della tolleranza al *self* ed il contenimento della risposta immune. Quest'ultima è resa efficace dalla coesistenza di due tipi di reazioni, chiamate immunità innata ed immunità adattativa.

L'immunità innata è una risposta precoce ed aspecifica messa in atto essenzialmente da barriere epiteliali, monociti e macrofagi, cellule dendritiche (DC) e cellule natural killer (NK). Essa non ha bisogno di una pre-esposizione all'antigene per essere innescata e costituisce la risposta immediata dell'organismo all'attacco di un patogeno senza conferire una protezione immunitaria a lungo termine. Per contro, l'immunità adattativa (o specifica), coinvolge cellule altamente specializzate, i linfociti, e viene attivata in maniera specifica da differenti antigeni microbici e non microbici. Le sue risposte sono potenziate dalla ripetuta esposizione ad uno stesso antigene. Esistono due tipi di risposta immunitaria specifica, denominati immunità umorale e immunità cellulare. L'immunità umorale costituisce il principale meccanismo di difesa contro i microbi extracellulari e contro le tossine da essi prodotti. Essa è mediata dagli anticorpi presenti sulla superficie dei linfociti B che, dopo essersi attivati e differenziati in plasmacellule, producono anticorpi in grado di legare i batteri successivamente

eliminati dai fagociti. Al contrario, l'immunità cellulare (o cellulo-mediata) è attuata dai linfociti T. Questi ultimi sono in grado di riconoscere differenti peptidi antigenici mediante recettori presenti sulla superficie cellulare chiamati T-cell Receptor (o TCR) in grado di riconoscere in modo specifico solo i peptidi antigenici legati a proteine codificate da una famiglia di geni polimorfici nota come "complesso maggiore di istocompatibilità" (Major Histocompatibility Complex, MHC) ed espresse sulla membrana delle cellule che presentano l'antigene (Antigen-Presenting Cells, APC), quali DC, cellule B e macrofagi [Lanzavecchia A. *et al.*, 1985; Caux C. *et al.*, 1994; Sebastien A. *et al.*, 2009]. Grazie a quest'ultimo legame, le risposte funzionali dei linfociti T innescano la trasmissione di segnali di attivazione all'interno delle cellule. Oltre ad esprimere il TCR, i linfociti T esprimono anche molecole accessorie importanti per l'attivazione cellulare, alcune delle quali interagiscono direttamente con i ligandi espressi sulle APC o sulle cellule bersaglio stabilizzandone il legame, mentre altre trasmettono all'interno dei linfociti T i segnali di attivazione. Globalmente, il sistema immunitario protegge il nostro organismo dall'infinita varietà di agenti patogeni con i quali giornalmente viene a contatto, tuttavia l'immunità innata e adattativa, seppur didatticamente indicate come distinte, sono in realtà due tipi di risposte complementari e difficilmente separabili.

2.2 Origine dei componenti del sistema immunitario

Il midollo osseo è la sede di produzione delle cellule del sistema immunitario che derivano da un unico progenitore staminale ematopoietico pluripotente. Da quest'ultimo originano due linee cellulari: la mieloide, da cui prendono origine la maggior parte delle cellule dell'immunità innata, e la linfoide, da cui si differenziano le cellule natural killer (NK) e i linfociti T e B. La maturazione dei linfociti B ed NK avviene nel midollo osseo, mentre i linfociti T si sviluppano nel timo. Poiché il presente studio si incentra sulla dissezione delle cellule T, verrà descritto in dettaglio il loro sviluppo. Durante la maturazione dei linfociti T si ha la formazione del loro repertorio immunologico attraverso il riarrangiamento dei geni che codificano per le catene (α - β o γ - δ) che compongono il recettore delle cellule T (TCR), assicurando in tal modo la specificità del riconoscimento dell'antigene. Il successivo processo di "educazione" a cui vengono sottoposti i linfociti T, e che si realizza mediante processi di selezione positiva e negativa, garantisce la

sopravvivenza delle cellule il cui TCR è funzionale e l'eliminazione delle cellule reattive contro gli antigeni *self*. Inoltre, l'esposizione del sistema immunitario ad un antigene estraneo aumenta la sua capacità di rispondere a quell'antigene. Più precisamente, quando il nostro organismo subisce per la prima volta l'attacco di un antigene, una parte delle cellule immunitarie specifiche per tale antigene subiscono un'espansione clonale e saranno caratterizzate dalla capacità di circolare nei vasi ematici e linfatici per numerosi anni. Questo fenomeno è chiamato "memoria immunologica" e darà luogo, in seguito ad una successiva esposizione, ad una risposta immunitaria più veloce e potente rispetto alla prima.

2.3 Tolleranza immunologica ed autoimmunità

La protezione del nostro organismo dall'attacco di agenti patogeni (antigeni) avviene mediante la sintesi di strutture altamente specializzate, gli anticorpi, e la generazione di elementi cellulari, quali linfociti e fagociti, che eliminano gli antigeni. In condizioni fisiologiche, gli individui sono "tolleranti" verso le molecole *self* poiché, durante il processo di sviluppo del sistema immunitario, i linfociti che le riconoscono in modo specifico vengono eliminati o inattivati, o addirittura indotti a cambiare la loro specificità. Tuttavia, i meccanismi molecolari da cui originano i geni che codificano i recettori per i diversi antigeni sono casuali e non sono condizionati da ciò che è estraneo o autologo all'organismo. Da quest'ultimo concetto risulta chiaro che i linfociti, che riconoscono molecole autologhe, possono generare una risposta immunitaria causando un danno tissutale. Per evitare che si verificano reazioni di questo tipo il sistema immunitario ha sviluppato dei meccanismi di tolleranza immunologica. Con il termine "tolleranza immunologica" si intende precisamente "la perdita della capacità del sistema immunitario di rispondere ad un determinato antigene indotta da un precedente incontro con lo stesso antigene". Una delle peculiarità della tolleranza immunologica è quella che si verifica nei confronti degli antigeni *self*. I meccanismi della tolleranza si distinguono in meccanismi centrali e meccanismi periferici che hanno luogo rispettivamente a livello del timo e degli organi linfoidi secondari. I meccanismi centrali si verificano per mezzo della selezione negativa che permette l'eliminazione dei linfociti T con un'elevata affinità per gli antigeni autologhi. I meccanismi periferici regolano le risposte dei linfociti T maturi sfuggiti alla selezione negativa che risultano essere

quindi autoreattivi. I principali meccanismi alla base della tolleranza linfocitaria sono: 1) la morte cellulare per apoptosi, detta “delezione clonale”, causata dalla ripetuta stimolazione dei linfociti attivati da parte di antigeni autologhi persistenti; 2) l’inattivazione funzionale della cellula, chiamata “anergia clonale”, la quale si verifica quando un peptide antigenico è presentato in assenza di molecole costimolatorie o quando i linfociti T specifici per un determinato antigene vengono in contatto con una forma mutata dell’antigene; 3) la soppressione dell’attivazione linfocitaria e delle funzioni effettrici ad opera dei linfociti ad azione regolatoria [Bluestone J.A., 2011]. Quest’ultimo meccanismo coinvolge una classe di linfociti T, detti appunto regolatori, che svolge un ruolo fisiologico importante per il mantenimento della non responsività verso gli antigeni *self* e più in generale per limitare l’entità delle risposte immuni specifiche verso gli antigeni estranei. Sulla base delle considerazioni appena fatte, è evidente che la deficienza o l’alterazione dei meccanismi normalmente responsabili del mantenimento della tolleranza immunologica possono determinare una risposta immune nei confronti del *self* tale da provocare l’insorgenza di fenomeni autoimmuni. La comprensione dei meccanismi responsabili dell’induzione della tolleranza è importante per definire con chiarezza specifici target molecolari per lo sviluppo di nuovi composti immunomodulatori in grado sia di inibire la risposta effettrice sia di aumentare quella regolatoria per non incorrere in tali fenomeni a carattere autoimmune.

2.4 Popolazioni cellulari oggetto di studio

a) *Cellule T, B e natural killer*

I linfociti, possono essere suddivisi in diverse sottopopolazioni, ciascuna con un compito preciso nelle diverse fasi della risposta immunitaria. I linfociti costituiscono il 20-40% delle cellule bianche totali del sangue, o leucociti, che riconoscono e rispondono in modo specifico agli antigeni estranei che entrano in contatto con il nostro organismo. I linfociti si distinguono in tre popolazioni principali: linfociti T, B e NK. I linfociti T riconoscono gli antigeni derivanti da agenti patogeni intracellulari ed agiscono uccidendo tali patogeni o le cellule infettate da essi. I linfociti T sono costituiti da popolazioni funzionalmente distinte: i linfociti T helper ed i linfociti T citotossici. In risposta alla stimolazione antigenica, i linfociti T helper secernono proteine denominate citochine, che

stimolano la proliferazione e la differenziazione dei linfociti T stessi. Nel dettaglio, i linfociti T helper legano i frammenti peptidici derivati dal processamento degli antigeni e presentati dalle molecole MHC di classe II presenti sulla superficie di membrana delle cellule che presentano l'antigene (Antigen Presenting Cell, APC). Invece, i linfociti T citotossici riconoscono e uccidono direttamente le cellule infettate da virus e altri microrganismi interagendo con le molecole MHC di classe I presenti sulla superficie di tutte le cellule.

L'altra classe di linfociti è rappresentata dai linfociti B, che costituiscono il 15% delle cellule bianche del sangue. L'attivazione dei linfociti B è innescata dal riconoscimento di antigeni extracellulari da parte degli anticorpi (o immunoglobuline, Ig) presenti sulla superficie della loro membrana. I linfociti B, una volta attivati, si differenziano a loro volta in cellule secernenti anticorpi, o plasmacellule. Gli anticorpi secreti dalle plasmacellule sono specifici per ogni dato antigene. Inoltre, i linfociti B, una volta legato l'antigene, presentano i peptidi antigenici ai linfociti T attivandoli.

Una terza classe di linfociti, le cellule natural killer, è coinvolta nella risposta immunitaria innata contro i virus e altri patogeni intracellulari e possiede una funzione citotossica simile a quella dei linfociti T CD8 positivi. Il loro nome deriva dal fatto che non hanno bisogno di essere attivate, mediante riconoscimento delle molecole MHC, per assolvere la loro funzione citolitica. Recentemente è emerso che le cellule NK abbiano caratteristiche che le accomunano alle cellule della immunità adattativa essendo sottoposte durante lo sviluppo ad una "educazione", esprimendo recettori antigenici specifici e generando cellule con memoria immunologica [Sun J.C. & Lanier L.L. 2011].

b) *Cellule T regolatorie*

Le cellule T regolatorie (o Treg) sono un insieme eterogeneo di sottopopolazioni specializzate di linfociti T che possiedono la capacità di sopprimere l'attivazione del sistema immunitario, limitando l'entità delle risposte immuni dirette verso gli antigeni estranei, e di mantenere la non responsività verso gli antigeni *self*. Nel processo di soppressione mediato dalle cellule Treg, un ruolo cruciale viene svolto da un fattore trascrizionale, denominato FoxP3, che è il principale regolatore dello sviluppo e della funzione delle cellule Treg [Hori S. *et al.*, 2003; Fontenot J. D. *et al.*, 2003]. Le Treg possono essere identificate in base alla mancata espressione

della catena alfa del recettore per l'interleuchina 7 (CD127) e alla forte espressione del fattore trascrizionale FoxP3 e della catena alfa del recettore per l'interleuchina 2 (CD25). Sino a poco tempo fa si pensava che le cellule Treg che esprimevano il recettore CD4 (Treg CD4+) fossero una popolazione omogenea e funzionalmente matura, infatti sino ad allora non erano emerse evidenze riguardanti sottoclassi di cellule Treg [Dendrou C.A. *et al.*, 2009]. Recentemente, queste cellule sono state divise in tre sottoclassi: quiescenti, attivate e secretorie. Le prime due classi hanno attività soppressoria in vitro secernendo citochine ad attività anti-infiammatoria quali TGF- β ed IL-10, mentre il terzo gruppo esplica un'attività secretoria proinfiammatoria attraverso la produzione di citochine quali l'IL-17, IL-2, INF- γ [Miyara M. *et al.*, 2009; Battaglia M. & Roncarolo MG., 2009]. Le cellule Treg hanno origine nel timo e vengono rilasciate poco dopo la nascita come cellule Treg naturali caratterizzate dal fenotipo rTreg (resting, CD25⁺⁺/CD45RA⁺) che esprimono l'antigene CD45RA, una quantità intermedia di CD25 e una bassa quantità di FoxP3. Dopo aver incontrato l'antigene in periferia, le cellule Treg proliferano ed, in parallelo, si ha una riduzione dell'espressione dell'antigene CD45RA ed un aumento dell'espressione degli antigeni CD25 e FoxP3, in tal modo si attivano trasformandosi in cellule aTreg (attivate, CD25⁺⁺⁺/CD45RA⁻) con attività soppressoria sulla proliferazione e sulle funzioni delle cellule T. L'alta attività proliferativa delle cellule aTreg sopprime anche la proliferazione delle cellule rTreg generando un controllo a feedback negativo che blocca l'ulteriore produzione di cellule Treg. La terza sottopopolazione, detta non Treg-cells (CD25⁺⁺, CD45RA⁻), esprime il CD25 solo dopo attivazione ed una bassa quantità di FoxP3; queste cellule non possiedono attività soppressoria, bensì proinfiammatoria. Sinora le basi genetiche della variabilità delle cellule Treg non sono state ancora ben chiarite, sebbene numerose evidenze sperimentali suggeriscano il ruolo delle loro alterazioni quantitative nell'influenzare il rischio di sviluppare una malattia autoimmune.

c] *Maturazione delle cellule T*

Il processo di maturazione dei linfociti T naive, che avviene in contemporanea alla loro migrazione dal sangue ai linfonodi, è costituito da tappe sequenziali. L'homing dei linfociti T naive ai tessuti linfoidei delle mucose avviene a livello delle venule postcapillari specializzate e coinvolge numerose molecole di adesione diverse e chemochine. In generale, il processo infiammatorio di un tessuto periferico

provoca un forte aumento dell'afflusso dei linfociti T naive ai linfonodi che drenano il sito di infiammazione. Gli antigeni si concentrano negli organi linfoidi secondari (linfonodi, i tessuti linfoidi associati alle mucose, milza) dove vengono presentati dalle cellule dendritiche mature ai linfociti T naive che riconoscono i loro antigeni specifici e si attivano, proliferando e differenziandosi in cellule effettrici che migrano nei tessuti infiammati. Una parte dei linfociti T attivati, dopo aver incontrato l'antigene nei linfonodi, rimane in circolo per lunghi periodi di tempo come "cellule di memoria". La presenza di cellule della memoria può essere l'unico indice di una risposta immunitaria precedente e rappresenta l'elemento distintivo dell'immunità specifica. Al contrario, i linfociti T naive che non riconoscono l'antigene all'interno dei linfonodi non si attivano e ritornano in circolo, attraverso i vasi linfatici efferenti, per ricercare in altri linfonodi l'antigene. I marcatori di superficie che abbiamo selezionato nel nostro studio per la discriminazione dei differenti stadi della maturazione delle cellule T, sono il CD45RA ed il CCR7 [Sallusto F. *et al.*, 1999]. Il CD45RA è una isoforma dell'antigene CD45, marcatore specifico dei leucociti, e rappresenta l'antigene distintivo dei linfociti T di tipo "naive" [Pilling D. *et al.*, 1996], mentre il recettore per la CC chemochina recettore di tipo 7 (CCR7), dopo aver interagito con i suoi ligandi (CCL19, CCL20 e SLC "secondary lymphoid tissue chemokine"), è in grado di promuovere una ferma adesione e transmigrazione delle cellule T naive nel linfonodo. Nel dettaglio, questi due marcatori ci permettono di discriminare quattro sottotipi di cellule T: le cellule naive (CD45RA+CCR7+), che sono linfociti maturi non sono ancora entrati in contatto con l'antigene; tre sottotipi funzionalmente distinti di cellule della memoria: le cellule T_{CM} o "central memory" (CD45RA-/CCR7+), che hanno bisogno di incontrare nuovamente l'antigene prima di diventare effettrici; le cellule T_{EM} o "effector memory" (CD45RA-/CCR7-) che sono pronte per esplicare la loro funzione effettrice nel sito di infiammazione; le cellule T_{TD} o "terminally differentiated" (CD45RA+/CCR7-), che ormai trascorso un intervallo di tempo troppo lungo dall'ultima stimolazione antigenica vanno incontro ad apoptosi [Carrasco J. *et al.*, 2006].

d] *Cellule dendritiche*

Le cellule dendritiche (DC) sono localizzate principalmente nei tessuti a contatto con l'ambiente esterno, quali la cute e le mucose, e rappresentano le più potenti cellule APC (Antigen Presenting Cell) in grado di catturare, processare e presentare

l'antigene ai linfociti T. Esse sono in grado di attivare i linfociti T naive sia a livello timico per indurre il loro differenziamento, consentendo l'instaurarsi della memoria immunologica [Banchereau J. *et al.*, 2000], che in periferia per iniziare la risposta immunitaria [Bonasio R. *et al.*, 2006]. I precursori delle cellule dendritiche sono i monociti che possono differenziarsi in una cellula dendritica o in un macrofago a livello tissutale. Le DC si possono classificare in due famiglie: le cellule mieloidi (CD11c positive) e le cellule plasmacitoidi (CD123 positive) [Yang G.X. *et al.*, 2005; Charles J. *et al.*, 2010] e vanno incontro a due stadi differenziativi: cellule dendritiche immature e cellule dendritiche mature. Le prime sono presenti nei tessuti periferici dove catturano, internalizzano e processano gli antigeni in peptidi che vengono esposti sulla superficie cellulare mediante le molecole MHC di classe II (HLA-DR, -DQ, -DP) per la loro presentazione ai linfociti T naive. Tuttavia, le cellule DC per poter migrare negli organi linfoidi secondari, dove avviene la presentazione antigenica, devono essere sottoposte a stimoli infiammatori caratteristici quali l'esposizione a citochine quali TNF- α e IL-1 β , prodotti batterici (LPS) e particelle virali. Durante la migrazione verso gli organi linfoidi, le DCs subiscono cambiamenti morfologici, fenotipici e funzionali che incrementano la loro capacità di attivare i linfociti T naive, trasformandosi in cellule mature. In considerazione del potenziale immunostimolatorio delle DC è plausibile ipotizzare che un'alterazione delle loro funzioni possa costituire una delle cause scatenanti l'insorgenza di patologie autoimmuni. Quest'ultima osservazione è in linea con alcune evidenze che descrivono le DC mieloidi e plasmacitoidi come possibili fattori coinvolti nell'insorgenza di patologie autoimmuni quali lupus eritematoso, psoriasi e diabete di tipo I [Banchereau J, *et al.*, 2006; Albanesi C. *et al.*, 2010; Chen X. *et al.*, 2007].

Sezione 3

3.1 Principi di citometria a flusso

L'avvento della citometria a flusso o citofluorimetria (CFM) risale agli anni 70 e, nell'arco di quattro decenni, è diventata una delle tecnologie più avanzate per la caratterizzazione, fenotipica e quantitativa, delle cellule in un modo veloce ed automatico. L'analisi dell'assetto immunofenotipico linfocitario rappresenta una delle sue principali applicazioni cliniche e sperimentali. Per mezzo di questa

moderna tecnica è possibile analizzare l'insieme delle differenti classi linfocitarie grazie all'utilizzo di anticorpi monoclonali coniugati con sostanze fluorescenti che riconoscono i loro specifici antigeni permettendone la fine discriminazione, numerazione e caratterizzazione e consentendo allo stesso tempo la quantificazione dei loro livelli di espressione in superficie. Il citometro a flusso è uno strumento che può rilevare la fluorescenza emessa da una singola cellula in sospensione e determinare, all'interno di un campione, il numero di cellule che esprimono la molecola a cui l'anticorpo fluorescente si lega. La capacità di discriminare le molteplici classi cellulari che costituiscono il sistema immunitario è un elemento essenziale per arricchire le nostre conoscenze riguardo l'immunità cellulare e la patogenesi delle malattie, e risulta ancor più rilevante nelle malattie che coinvolgono direttamente il sistema immunitario, come leucemie o malattie autoimmunitarie ed infezioni da virus dell'immunodeficienza, caratterizzate da una alterazione dei rapporti tra le sottopopolazioni linfocitarie [Ferreira M.A., 2010]. Un citometro a flusso consiste di tre componenti principali: a) un sistema fluidico che controlla la captazione ed il flusso cellulare; b) un sistema ottico in cui le cellule vengono attraversate da un raggio laser; c) un sistema elettronico che raccoglie, rappresenta ed analizza i dati (Figura 1).

a) Sistema fluidico

Un campione in soluzione per essere analizzato da un citometro deve essere ordinato in un flusso di particelle singole o fase monodispersa. Questo processo è gestito dal sistema fluidico che è costituito da un canale di conta centrale attraverso il quale il campione, una volta iniettato, viene racchiuso da una guaina esterna (o fluido di trascinamento), che abbraccia le particelle creando intorno ad esse una spirale, da destra verso sinistra, tale da farle muovere dal basso verso l'alto vincolandole a disporsi in fila indiana. Il fluido di trascinamento permette alle particelle di fluire più velocemente grazie al principio della "focalizzazione idrodinamica" secondo cui le cellule fluiscono una dietro l'altra grazie alla differente pressione esistente tra il campione ed il fluido di trascinamento (flusso laminare). In questo modo, il flusso di particelle presenti all'interno del capillare è in grado di passare, alla stessa distanza ed alla stessa velocità, davanti ad un raggio laser di forma

ellittica focalizzato in modo da colpire le cellule al centro del canale di conta; il punto in cui il laser incontra le particelle si chiama “interrogation point”.

b) Sistema ottico: Ottica di eccitazione, emissione e rilevazione

Dopo la focalizzazione idrodinamica, ogni cellula viene colpita da uno o più fasci laser. La misura delle fluorescenze, fornita dalle differenti lunghezze d'onda a cui emettono i diversi fluorocromi, forniscono dati qualitativi e quantitativi riguardo i recettori di superficie cellulare a cui si legano attraverso gli anticorpi. Questo sistema permette, inoltre, di misurare le proprietà fisiche delle cellule grazie alla capacità delle cellule di disperdere la luce anteriormente (FSC_forward scatter) e lateralmente (SSC_side scatter); questi ultimi sono parametri fisici che riflettono rispettivamente le dimensioni cellulari e la struttura interna, intesa come rugosità di membrana, granulosità esterna e rapporto tra nucleo e citoplasma. Il FSC ed il SSC sono specifici per ogni tipo cellulare e la loro combinazione è utilizzata per differenziare un campione eterogeneo di cellule. L'insieme di questi segnali vengono raccolti da un sistema di lenti, specchi e filtri ottici, posizionati dietro la camera di conta, ed inviati ai rispettivi sensori di rilevazione rappresentati da fotodiodi o fotomoltiplicatori di silicio (PMTs) che, dopo aver trasformato il segnale ottico in segnale elettrico, lo amplificano ed infine ne misurano la specificità e l'intensità. Le diverse emissioni di colore, dato l'utilizzo di più di una sostanza fluorescente, vengono separate attraverso un sistema di filtri ottici per poter analizzare separatamente le diverse fluorescenze.

c) Sistema elettronico

I fotomoltiplicatori convertono i segnali luminosi in impulsi elettrici direttamente proporzionali alla luce raccolta. I segnali elettrici provenienti da ogni sensore, opportunamente amplificati e digitalizzati, sono infine inviati ad un analizzatore di dati che provvede alla loro visualizzazione su monitor, rappresentazione grafica e definizione statistica. I segnali digitali raccolti vengono raffigurati su uno schermo sotto forma di puntini: ogni puntino rappresenta “un evento” definito dalle proprietà della dispersione di luce

possedute da ogni singola cellula. Le cellule possono essere visualizzate mediante “citogramma o dot plot”, un diagramma bidimensionale che mette in correlazione due differenti parametri rilevati, o mediante un istogramma che definisce il diagramma delle conte cellulari rispetto a ciascun canale.

Sezione 4

Materiali e Metodi

4.1 Misurazioni citofluorimetriche

La caratterizzazione immunofenotipica è stata effettuata su 2870 individui afferenti al progetto ProgeNIA. Per i dettagli relativi alla coorte si veda il paragrafo “Progetto ProgeNIA”. I campioni di sangue periferico prelevati da ciascun volontario sono stati raccolti in provette contenenti eparina come anticoagulante, processati e colorati con anticorpi fluorescenti specifici ed acquisiti al citometro a flusso (FACSCanto II, BD Biosciences). Per evitare artefatti causati da variabili tempo-dipendenti, la fenotipizzazione leucocitaria dei campioni di sangue fresco è stata effettuata entro due ore dalla loro raccolta. L’analisi immunofenotipica delle cellule bianche del sangue periferico di ciascun volontario è stata effettuata grazie all’utilizzo di quattro differenti pannelli di anticorpi monoclonali. Ogni pannello è costituito dall’insieme di anticorpi diretti verso gli antigeni di superficie che caratterizzano le popolazioni leucocitarie selezionate per lo studio. I cloni e gli anticorpi utilizzati nel presente studio sono ampiamente sfruttati dalla comunità scientifica e sono in linea con quelli utilizzati dallo Human Immunology Project Consortium [Maecker H.T. *et al.*, 2012]. I quattro pannelli, che sono stati denominati in modo da indicare in sintesi i tipi cellulari o le caratteristiche principali delle cellule analizzate, sono i seguenti: pannello T-B-NK, pannello T regolatorie, pannello Maturazione T cellulare e pannello cellule dendritiche circolanti.

Nel dettaglio:

a) Il pannello T-B-NK permette di misurare 10 parametri con 9 anticorpi. Il sangue è stato processato mediante l’utilizzo di tubi BD TruCount™ al cui interno sono presenti le biglie necessarie per calcolare la conta assoluta (o numero

cellule/ μ l; Bead Count BD, cat. 340334). Le biglie di conta vengono risospese in 50 μ l di sangue periferico fresco; in seguito, i campioni di sangue vengono colorati con gli anticorpi fluorescenti riportati nella Tabella 1. Dopo 15 minuti dalla colorazione i campioni sono stati sottoposti a lisi cellulare per 30 minuti con 1 ml di FACS Lysing Solution 1x (BD, cat. 349202) e, dopo 3 ore di incubazione (+4°C), sono stati acquisiti al citometro a flusso. Tale protocollo è indicato come “Lyse-No-Wash”. Lo schema di analisi cellulare e la creazione della gerarchia delle popolazioni studiate è il seguente: i leucociti sono stati suddivisi in base ai livelli di positività per l’antigene CD45 e ai parametri morfologici in granulociti, monociti e linfociti. I linfociti sono stati differenziati in base all’espressione dell’antigene di superficie CD3. Le cellule CD3+, o linfociti T, sono state ulteriormente suddivise in 6 sottoclassi in base al grado di espressione degli antigeni CD4 e CD8 (CD4-CD8-, CD4- CD8dim, CD4- CD8bright , CD4+ CD8+ e CD4+ CD8dim, CD4+ CD8-), mentre le cellule CD3- sono state suddivise in cellule B (CD3- CD19+) e cellule natural killer (CD3- CD16+ e/o CD3- CD56+). Per rilevare il grado di attivazione cellulare dei linfociti T e delle cellule NK, è stato utilizzato l’anticorpo diretto contro l’antigene HLA-DR, il quale è costitutivamente espresso nelle cellule che presentano l’antigene quali le cellule B e le cellule dendritiche. I monociti sono stati caratterizzati anche mediante la positività per il marcatore CD14. La strategia di analisi utilizzata per il pannello delle T-B-NK è mostrata nella Figura 3, mentre la descrizione statistica degli immunofenotipi è riportata nella Tabella 2.

b) Il pannello delle cellule T regolatorie è caratterizzato da 10 parametri ed 8 anticorpi. Il sangue è stato processato mediante la tecnologia BD Lyotube™, (BD, cat. 624644) che utilizza un cocktail liofilizzato pre-titolato di anticorpi. Il protocollo per il processamento del sangue con i tubi Lyotube™ è chiamato “Lyse-Wash” in virtù del fatto che le cellule vengano lavate con opportune soluzioni prima della loro acquisizione. Gli anticorpi fluorescenti utilizzati per la misurazione delle cellule Treg sono riportati nella Tabella 1. In dettaglio, il cocktail di anticorpi liofilizzati viene eluito per 5 minuti in 100 μ l di una soluzione di lavaggio (FACS FLOW, BD cat. 342003), sono stati quindi aggiunti 100 μ l di sangue fresco per effettuare la colorazione della durata di 15 minuti. Successivamente, gli eritrociti sono stati lisati in 3 ml di FACS Lysing Solution 1x e le cellule incubate per 30 minuti a temperatura ambiente e centrifugate a 1500 rpm per 5 minuti. Infine, le cellule sono state risospese in 500 μ l di soluzione FACS FLOW ed analizzate al citofluorimetro. La conta assoluta delle cellule Treg è stata calcolata tenendo conto

del numero assoluto di linfociti CD3+ misurati nel pannello T-B-NK. Lo schema di analisi cellulare e la creazione della gerarchia delle popolazioni studiate è il seguente: l'anticorpo contro l'antigene CD3 ha permesso di discriminare i linfociti T dagli altri leucociti presenti nel sangue. I linfociti T sono stati suddivisi in CD4+ e CD8+. Le cellule Treg CD4+ sono state discriminate in base all'alto grado di espressione dell'antigene CD25 e alla bassa espressione dell'antigene CD127 (CD25high CD127low). Le cellule Treg sono state ulteriormente suddivise in attivate (CD25+++ CD45RA-), quiescenti (CD25++, CD45RA+) e secretorie non soppressorie (CD25++ CD45RA-). Inoltre, sono state analizzate le cellule T non regolatorie CD4+ CD25high (CD25high CD4+ not Treg) valutando le cellule T CD4+ esprimenti un'elevata quantità dell'antigene CD25 a cui sono state sottratte le cellule Treg, la popolazione risultante è stata ulteriormente suddivisa in base alla positività per l'antigene CD45RA. In questo pannello sono state anche analizzate le cellule T CD8+, divise in base all'espressione della molecola co-stimolatoria CD28 e dell'antigene CD45RA. Infine, in specifiche popolazioni è stato misurato il marcatore di attivazione ectonucleotidasi trifosfato difosfoidrolasi 1 (ETNPD1), anche conosciuto come antigene CD39. La strategia di analisi utilizzata per il pannello delle cellule T regolatorie è mostrata nella Figura 4, mentre la descrizione statistica degli immunofenotipi è riportata nella Tabella 2.

c) Il pannello della maturazione delle cellule T è caratterizzato da 7 parametri e 5 anticorpi. La colorazione è stata effettuata utilizzando il protocollo Lyse-Wash e la tecnologia Lyotube BD™ (vedi Pannello Treg). Gli anticorpi fluorescenti utilizzati per la colorazione delle cellule bianche del sangue sono riportati nella Tabella 1. Il pattern di maturazione delle 6 sottoclassi di cellule T descritte nel pannello T-B-NK (CD4-CD8-, CD4-CD8dim, CD4- CD8bright, CD4+CD8+ e CD4+CD8dim, CD4+CD8-) è definito da 2 marcatori: CCR7 e CD45RA. La caratterizzazione delle fasi della maturazione delle cellule T prevede l'identificazione delle cellule naive (CD45RA+CCR7+), cellule T_{CM} (CCR7+CD45RA-), cellule T_{EM} (CD45RA-CCR7-) e cellule T_{TD} (CCR7-CD45RA+). La conta assoluta delle cellule T della maturazione è stata calcolata tenendo conto del numero assoluto di linfociti CD3+ misurati nel pannello T-B-NK. La strategia di analisi utilizzata per il pannello degli stadi della maturazione T cellulare è mostrata nella Figura 5, mentre la descrizione statistica degli immunofenotipi è riportata nella Tabella 2.

d) Il pannello relativo alle cellule dendritiche circolanti (cDC) consente la discriminazione di 10 parametri con l'utilizzo di 13 anticorpi. La colorazione è stata effettuata utilizzando il protocollo Lyse-Wash e la tecnologia Lyotube BD™ (vedi Pannello Treg). Il numero assoluto di cellule DCs è stato calcolato mediante l'utilizzo di biglie di conta (o "Liquid BD Counting Beads", BD cat. 335925). Gli anticorpi fluorescenti utilizzati per la misurazione delle cellule cDCs sono riportati nella Tabella 1. Per poter discriminare le cDCs abbiamo considerato la positività di queste cellule per l'antigene HLA-DR e la loro negatività per cocktail di anticorpi (detto Lin1) composto da anticorpi contro gli antigeni CD3, CD14, CD19, CD20, CD56 e CD16. Nello specifico, l'antigene CD16 è espresso da uno specifico sottotipo di cDC mieloidi [Chen L. *et al.*, 2007], che non è stato considerato nel nostro studio. Le cellule DC sono state quindi suddivise in base alla loro positività per le molecole di superficie CD11c e CD123, rispettivamente, in mieloidi e plasmacitoidi. Inoltre, lo stato di maturazione delle cDC è stato determinato in base all'espressione della molecola di adesione CD62L, indispensabile per l'ingresso della cellula dendritica nel linfonodo, e del recettore per le chemochine CCR2 che media l'arrivo delle DC al sito di infiammazione. Infine il loro grado di attivazione è stato valutato grazie ad due marcatori diretti verso gli antigeni CD86 e CD80, ligandi della molecola costimolatoria CD28 presente sulla superficie del recettore delle cellule T. La strategia di analisi utilizzata per il pannello delle cellule dendritiche circolanti è mostrata nella Figura 6, mentre la descrizione statistica degli immunofenotipi è riportata nella Tabella 2.

I dati citofluorimetrici sono stati analizzati mediante il software BD FACSDiva™ (BD Biosciences).

4.2 Settaggio strumentale dei citofluorimetri

La caratterizzazione immunofenotipica è stata effettuata mediante l'utilizzo di due citofluorimetri FACSCanto™ II (BD, Biosciences) a 3 laser (blu, emissione 488 nm; rosso, emissione 633 nm; viola, emissione 407 nm) che permettono la rilevazione di 8 fluorescenze e dieci parametri. La taratura degli strumenti è stata finalizzata a garantire la massima resa e sensibilità su entrambi gli strumenti. Tale sincronizzazione dei citofluorimetri ha l'obiettivo di renderli interscambiabili per

l'acquisizione del campione. In dettaglio, è stato valutato lo SDEN (Standard deviation of Electronic background Noise) per tutti i parametri di fluorescenza su entrambi gli strumenti, mediante l'utilizzo di apposite biglie di calibrazione BD CS&T Beads (BD, cat. 642412) e mediante uno specifico software BD FACSDiva CS&T Module. Successivamente è stata generata una matrice dello SDEN sulla base del valore più elevato per ogni parametro di fluorescenza; tale matrice è stata utilizzata come referenza per entrambi i citofluorimetri per definire la potenza dei fotomoltiplicatori in modo che la varianza (SDEN²) del segnale di fondo aspecifico, influenzasse la varianza dei linfociti non marcati in misura del 10-20%. Il settaggio dei fotomoltiplicatori è stato attuato in uno strumento mediante la creazione di valori di riferimento che costituiscono la "baseline" e che sono stati applicati inizialmente alle cellule e successivamente alle biglie CS&T. Questi valori sono stati trasferiti indirettamente sul secondo strumento mediante le biglie CS&T che hanno permesso di generare in esso un settaggio strumentale identico al primo. Per mantenere una corretta taratura strumentale nel tempo è stato utilizzato il software del controllo di qualità CS&T che si basa sull'acquisizione di biglie CS&T caratterizzate da tre diversi tipi di fluorescenza (bassa, media, alta) su tutti i parametri misurati da una macchina e da differenti coefficienti di variazione (CV). Il CV è una misura che fornisce un'indicazione della variabilità delle misurazioni effettuate e risulterà basso per le biglie brillanti che portano un segnale distinto dal rumore di fondo, pertanto la loro misurazione sarà sempre molto precisa, mentre l'opposto accadrà per le biglie a bassa fluorescenza. Infine, la compensazione, ovvero il processo che permette di correggere il fenomeno per cui un fluorocromo viene letto da un fotomoltiplicatore non deputato alla lettura di quel fluorocromo (spillover), è stata calcolata utilizzando le CompBeads BD (BD Biosciences) e validata su entrambi gli strumenti. La compensazione sulla seconda macchina è risultata essere molto simile, se non identica, a quella calcolata sulla prima macchina.

4.3 Riproducibilità delle misurazioni citofluorimetriche

Per assicurarci che i campioni di sangue venissero acquisiti in modo costante per tutta la durata dello studio sono stati effettuati giornalmente dei controlli di qualità. In dettaglio, abbiamo acquisito le biglie CS&T che ci hanno fornito un controllo elettronico giornaliero delle macchine. Successivamente, per monitorare e

mantenere le stesse condizioni di acquisizione e lettura nel tempo, sono state utilizzate le biglie ad 8 picchi Rainbow (Sphero™, cat. 559123). Queste ultime hanno fluorescenza standardizzata al fine di controllare e sostenere nel tempo le stesse condizioni di acquisizione e, in caso di necessità, tali particelle con fluorescenza nota, ci permettono di normalizzare le intensità di fluorescenza media ottenute nei nostri campioni. Settimanalmente le misurazioni citofluorimetriche sono state confermate grazie all'acquisizione del Multicheck™ Quality Control, ovvero campioni di sangue stabilizzati e per i quali i valori del saggio ed i range di misurazioni previsti per la conta assoluta delle popolazioni leucocitarie sono stati forniti dalla ditta produttrice. Infine, per valutare nel tempo la riproducibilità delle misurazioni dei tratti immunofenotipici, 13 volontari sono stati caratterizzati fenotipicamente due volte con un intervallo di tempo di almeno 3 mesi fra la prima e la seconda misurazione. Grazie a questi controlli mirati siamo stati in grado di monitorare i parametri di acquisizione nelle diverse fasi del progetto.

4.4 Ereditabilità

Le stime dell'ereditabilità sono state effettuate sui primi 1629 campioni immunofenotipizzati comprendenti 211 individui non imparentati, 1,418 soggetti raggruppati in 249 famiglie, principalmente costituite da 567 coppie di fratelli, 30 coppie di fratellastri, 248 coppie di cugini, 609 coppie di genitore-figlio, 32 coppie di nonno-nipote e 561 coppie di zii. Le stime dell'ereditabilità sono state eseguite, attraverso l'utilizzo del software "*poly*". L'età ed il genere sono stati presi in esame come covariate, cioè come variabili che possono essere utili per spiegare la variabilità dei dati ottenuti.

4.5 Genotipizzazione con GeneChip arrays

La coorte ProgeNIA è stata caratterizzata con due array custom (Illumina): il Cardio-MetaboChip e l'ImmunoChip. Queste piattaforme di genotipizzazione sono state progettate da consorzi internazionali per identificare regioni d'interesse correlate a tratti e/o malattie legate, rispettivamente, al sistema metabolico ed al sistema immunitario [Cortes A. and Brown M.A., 2011; Voight B.F. *et al.*, 2012]. Il

Cardio-MetaboChip e l'ImmunoChip analizzano varianti genetiche sia comuni che rare alternando regioni genetiche ad alta copertura con regioni a bassa copertura. I risultanti 283,955 SNPs autosomali che hanno superato i controlli di qualità sono stati utilizzati come genotipi di riferimento per imputare altre varianti attraverso l'utilizzo di un pannello di sequenze sarde.

4.6 Sequenziamento dei campioni

In questo studio abbiamo utilizzato per la prima volta un pannello di riferimento ottenuto dal sequenziamento a bassa risoluzione (4x) sull'intero genoma di 1146 individui sardi, di cui 481 individui sono afferenti al progetto ProgeNIA, mentre i restanti 665 hanno preso parte a studi relativi a Sclerosi Multipla e Diabete di tipo 1 [Sanna S. *et al.*, 2010]. Le librerie genomiche sono state generate a partire da 3-5 µg di DNA genomico utilizzando due kits differenti: "Paired-End Genomic Sample Prep Kit" (Illumina e NewEnglandBiolabs). Il sequenziamento è stato effettuato in due centri di ricerca localizzati presso il centro CRS4 (Centro di Ricerca, Sviluppo e Studi Superiori in Sardegna, Italia) che dispone di 2 analizzatori "Genome Analyzer Iix" (Illumina) e 3 strumenti "Hiseq2000" (Illumina) e presso il centro "DNA Sequencing Core" dell'università del Michigan, U.S.A., il quale possiede 2 analizzatori "Genome Analyzer Iix" e 2 strumenti "Hiseq2000".

4.7 Imputazione dei genotipi

L'imputazione o inferenza statistica è un modello di calcolo probabilistico attraverso il quale si cerca di ricostruire delle intere sequenze di DNA a partire da dati genetici preesistenti e parziali derivanti da genotipizzazione mediante piattaforme. Questo strumento di analisi è rilevante per lo studio del genoma poiché permette di stimare con un certo grado di accuratezza varianti genetiche che non sono state direttamente genotipizzate sfruttando un pannello di riferimento costituito da sequenze dell'intero genoma. La qualità dell'imputazione, calcolata come probabilità media che il genotipo imputato sia corretto, è definita da un parametro chiamato RSQR. Questo parametro varia da 0 a 1 e la soglia che rimuove il 90% degli SNPs inferiti in maniera non corretta è ≤ 0.3 [Li Y. *et al.*, 2009]. In questo studio i dati di genotipizzazione derivanti dai ChipArrays sono stati integrati con i dati generati dal sequenziamento di 1146 individui, mediante la

costruzione di un pannello di riferimento contenente i profili genetici degli individui risultati non imparentati tra tutti quelli sequenziali. L'imputazione degli aplotipi è stata effettuata utilizzando il "MiniMac" [Howie B. et al., 2012]. Sono stati genotipizzati o imputati (con una qualità di imputazione >0.3) un totale di 8,213,003 marcatori genetici.

4.8 Analisi statistiche e bioinformatiche

Per ciascun tratto immunofenotipico misurato è stato effettuato un GWAS utilizzando circa 8.2 milioni di varianti genetiche derivanti dall'integrazione, mediante imputazione [Li Y. et al., 2009], dei dati ottenuti dai due saggi Illumina, Immunochip e CardioMetabochip, sfruttando quelli derivanti dal sequenziamento a bassa risoluzione (4x) di 1146 individui sardi. L'analisi di associazione è stata effettuata usando un'analisi di regressione basata su un modello a componenti di varianza, che sfrutta le strutture familiari qui analizzate ed include l'età ed il sesso come covariate [Chen W.M and Abecasis G.R., 2007]. Sono state quindi selezionate le varianti genetiche indipendenti ($r^2 < 0.1$) associate a ciascun tratto, che hanno mostrato un p-value minore di 5.26×10^{-10} . Successivamente, sono state rimosse: i) le varianti con una bassa qualità di imputazione, ii) le associazioni tratto/variante ridondanti dando priorità alle varianti più associate a ciascun locus e iii) le associazioni in linkage disequilibrium (LD) con le varianti primariamente associate. Abbiamo inoltre incluso due associazioni suggestive ($p < 5 \times 10^{-8}$) localizzate nelle regioni geniche di *SH2B3/ATXN2* e *SLFN13* descritte precedentemente dal gruppo di Ferreira (Ferreira et al., 2010). Per validare i risultati ottenuti, sono state condotte le analisi di associazione tra variante associata e tratto immunofenotipico su un campione esteso di 2870 individui che comprende gli iniziali 1629 volontari e ulteriori 1241 individui afferenti alla stessa coorte ProgeNIA. Inoltre, sono state genotipizzate le varianti imputate che hanno dato segnali di associazione e che non sono supportate da variante proxy ($r^2 > 0.85$) direttamente genotipizzate. Le varianti che mostravano un eccesso di discordanza tra genotipi inferiti e genotipizzati o un p-value meno significativo, dopo l'aggiunta dei 1241 individui, sono state escluse dalle analisi successive.

Infine, per misurare la percentuale di varianza fenotipica spiegata dai fattori genetici, per ciascun tratto è stato creato un modello lineare che include età, sesso e tutti gli SNPs indipendenti associati a un tratto specifico (full model), ed un

modello lineare che include solo età e sesso (basic model). La varianza spiegata è stata calcolata come la differenza tra la varianza spiegata da due modelli full e basic.

Per identificare i geni candidati a ciascun locus, abbiamo preso in considerazione: a) geni la cui espressione fosse regolata dagli SNP associati agli immunofenotipi (eQTLs), b) i geni in cui erano localizzate varianti codificanti c) i geni immunocorrelati prossimali alle varianti associate come riportato in OMIM (Online Mendelian Inheritance in Man).

Infine le analisi bioinformatiche sono state effettuate utilizzando il sito web UCSC Genome Browser per verificare se le varianti associate fossero localizzate in regioni regolatrici e i database pubblici clinicaltrials.gov, FDA.gov e USPTO.gov per verificare se le stesse varianti fossero di interesse farmacologico.

4.9 Ricerca delle associazioni coincidenti tra tratti immunologici e malattie

Per individuare le associazioni coincidenti tra tratti immunologici e malattie sono state utilizzate le informazioni genetiche disponibili nei seguenti siti web: GWAS Catalog [Hindorff L.A.] ed ImmunoBase [<http://www.immunobase.org/>]. In questi database pubblici è possibile trovare le varianti genetiche risultate associate a patologie, autoimmuni e non, grazie ai risultati emersi da recenti studi GWAS. Nello specifico, è stata effettuata la ricerca degli SNPs in linkage disequilibrium ($r^2 \geq 0,8$) con gli SNPs trovati associati ai tratti immunologici misurati nel nostro studio.

Sezione 5

Risultati

5.1 Stime dell'ereditabilità

L'ampia eterogeneità di popolazioni di cellule immunitarie ed i loro ruoli nel mediare e nel controllare le reazioni immuni è stata ampiamente descritta nel corso degli ultimi decenni. Al contrario, la misura in cui la variabilità quantitativa dei livelli di cellule immunitarie circolanti è ereditariamente trasmessa rimane tuttora ancora da definire con chiarezza, fatta eccezione per poche classi di linfociti [Ferreira M.A. *et al.*, 2010; Okada Y. *et al.*, 2011; Nalls M.A. *et al.*, 2011]. Per verificare in che misura i geni influenzano i livelli quantitativi dei tratti

immunofenotipici analizzati in questo studio è stata eseguita la stima dell'ereditabilità esaminando anche l'influenza esercitata dalle covariate età e sesso. I risultati emersi dal nostro lavoro dimostrano che una porzione rilevante della variazione dei livelli quantitativi delle popolazioni leucocitarie analizzate è "geneticamente programmata". Infatti, i valori di ereditabilità sono compresi tra il 3% e l'87% con un valore medio pari al 41%. Nello specifico, i tratti che hanno mostrato una più alta ereditabilità sono le sottopopolazioni di cellule T regolatorie ($h^2=55\%$) (Figure 7- 8), che rappresentano una classe cellulare la cui filogenesi è avvenuta in maniera tardiva rispetto ad altri tipi cellulari. La caratteristica comune della maggior parte delle popolazioni cellulari con ereditabilità eccezionalmente alta (> 60%) è legata alla loro positività per il marcatore CD39, il quale sembra avere pertanto un ruolo predominante nel determinare una porzione apprezzabile della componente genetica per i tipi cellulari in cui esso è espresso. Per quanto concerne il sesso, esso esercita un debole effetto sui tratti immunofenotipici analizzati, mentre l'età ha un'influenza più marcata in particolare sulla popolazione delle cellule T CD8+ naive in linea con risultati pubblicati precedentemente che evidenziano come la riduzione di questa subset cellulare sia responsabile del ridotto successo delle vaccinazioni negli anziani [Sansonì P. *et al.*, 2008; Buchholz V.R. *et al.*, 2011].

5.2 Associazioni genetiche

Successivamente, considerata la notevole ereditabilità media dei tratti, è stato effettuato uno studio di associazione sull'intero genoma per ogni tratto immunofenotipico analizzato. I risultati delle analisi dei GWAS hanno portato all'identificazione di 23 varianti genetiche in 13 loci, ciascuna con un livello di significatività (p-value) inferiore a 5.26×10^{-10} (Figura 9). In generale, la maggior parte delle associazioni identificate nel nostro studio sono nuove, sebbene alcune di esse siano in linea con altri segnali di associazione più deboli rilevati in precedenza da altri gruppi di ricerca, come l'associazione tra i livelli di cellule NK e varianti localizzate nel cluster genico della famiglia *Schlafen* e quella tra i livelli di cellule T CD4+ e varianti genetiche localizzate nella regione genica *ATXN2/SH2B3* [Ferreira M.A. *et al.*, 2010]. Tra le 23 varianti genetiche identificate nel nostro studio, 7 di esse cadono in elementi con chiara funzione regolatoria, quali elementi promotori e repressori, enhancers ed ancora siti di legame per fattori di

trascrizione. La variante genetica con maggior effetto fenotipico è costituita da un polimorfismo intronico localizzato nel gene *ENTPDI* (EctoNucleoside Triphosphate diPhosphohydrolyase 1) che codifica per l'antigene CD39 (Figura 10a). Il CD39 è un ectoenzima che idrolizza l'ATP e l'ADP presenti a livello extracellulare in AMP. Nei leucociti, il CD39 è espresso su monociti, neutrofili, e sulle cellule B, ed in misura minore sulle cellule T e sulle cellule NK [Pulte E.D. *et al.*, 2007]. In particolare, tra le cellule T, il CD39 è espresso principalmente dalle cellule Treg attivate CD4+, dove svolge una funzione antinfiammatoria, attraverso la riduzione della molecola ATP extracellulare pro-infiammatoria [Borsellino G. *et al.*, 2007]. Quest'associazione rappresenta un chiaro meccanismo biologico in cui una variante *in cis* regola i livelli di espressione della proteina codificata dal gene in cui la variante è localizzata facendo anche aumentare il numero di cellule che esprimono tale marcatore. Un altro chiaro candidato biologico emerso grazie alle associazioni identificate in questo lavoro include una variante genetica localizzata vicino al gene che codifica per il class II transactivator gene, o *CIITA*, forte attivatore della trascrizione dei geni *HLA di classe II*. La variante è infatti associata alla regolazione dei livelli di linfociti T attivati che sono stati caratterizzati mediante l'espressione della molecola HLA-DR (Figura 10b).

Il nostro studio ha rivelato anche la presenza di più varianti alleliche indipendenti presenti all'interno della regione che comprende i geni *GALM* (GALactose Mutarotase) e *HNRPLL* (Heterogeneous Nuclear Ribonucleoprotein L-like) che modulano il numero di cellule appartenenti alle sottopopolazioni di linfociti T naive e terminally differentiated, contraddistinti dalla positività per l'antigene CD45RA, con una concomitante ed opposta regolazione dei livelli di popolazioni di cellule T della memoria (CD45RA-). L'associazione con il gene *HNRPLL* è in linea con il suo ruolo di principale regolatore dello splicing dell'antigene CD45 che rappresenta l'antigene distintivo della maturazione delle cellule T [Wu Z. *et al.*, 2010]. Al contrario, i meccanismi biologici alla base delle associazioni identificate tra le varianti localizzate nel gene *GALM* sono meno chiari, infatti *GALM* potrebbe modulare la glicosilazione dell'antigene CD45, alterandone l'espressione sulla superficie cellulare. In alternativa, questa variante genetica potrebbe avere un effetto regolatore a distanza sul gene *HNRPLL* poiché questa variante è localizzata in regioni che potrebbero interagire con il suo promotore [Li G. *et al.*, 2012]. Altri chiari candidati biologici includono una variante localizzata nella regione genica *CD8A-CD8B* che codificano per la glicoproteina di membrana CD8, associata alla modulazione dei livelli della popolazione CD4+CD8dim, o una variante localizzata

nel gene *TNFSF13B* (Tumor necrosis factor (ligand) superfamily, member 13) che è associata ai livelli delle cellule B. Quest'ultimo gene codifica per la proteina BAFF (B Activating Factor of the TNF Family) la cui inattivazione è associata in modo specifico alla riduzione della maturazione delle cellule B circolanti [Kreuzaler M. *et al.*, 2012; Schiemann B. *et al.*, 2001].

Tra i segnali multipli indipendenti presenti nella stessa regione genetica e associati con più tratti immunofenotipici possiamo menzionare quelli localizzati nella regione del gene *ENTDPI* e nella regione *HLA* per il quale dei segnali di associazione multilocus e multiallelici sono già stati descritti essere legati a patologie complesse [Marrosu M.G. *et al.*, 2001]. Le varianti localizzate nei geni *HLA* mediano la funzione del sistema immune influenzando i livelli di specifici tipi cellulari sia in condizioni fisiologiche che patologiche, infatti alcune di esse sono associate con la modulazione di numerose sottopopolazioni T cellulari CD8+. Quest'ultimo dato è consistente con un recente lavoro secondo cui le molecole MHC di classe I favoriscono la sopravvivenza delle cellule T CD8+ [Takada K. and Jameson S.C., 2009].

Alcune associazioni identificate nel presente studio sono localizzate in geni apparentemente non correlati con il sistema immunitario. Ad esempio, un'associazione suggestiva ($p < 5 \times 10^{-8}$) è stata identificata tra una comune variante missenso, codificante e conservativa, presente all'interno dell'esone 13 del gene *LCT* (lattasi) e la variazione dei livelli di cellule pDC suggerendo per questo gene un ruolo ancora sconosciuto nella regolazione della risposta immunitaria innata. Tuttavia, un eventuale ruolo funzionale esercitato dal prodotto proteico del gene *LCT* sarebbe coerente con la posizione dello SNP associato che risulta localizzato in una regione con funzione regolatrice. A supporto di questa ipotesi, è stata recentemente identificata l'associazione tra due varianti missenso indipendenti del gene *LCT* con la conta leucocitaria in individui Afro-Americani [Auer P.L. *et al.*, 2012].

5.3 Associazioni coincidenti tra tratti immunofenotipici e malattie

Uno degli obiettivi principali del presente studio consiste nell'identificare dei punti di sovrapposizione tra tratti immunologici e malattie, indicativi di meccanismi patogenetici causali. Infatti, se la variazione quantitativa di una popolazione leucocitaria è associata a una variante genetica che è anche associata a una malattia,

allora la stessa popolazione leucocitaria è verosimilmente coinvolta nella patogenesi di tale malattia. Pertanto, grazie all'osservazione di queste associazioni coincidenti siamo riusciti ad identificare categorie di cellule del sistema immunitario potenzialmente coinvolte nella suscettibilità verso una malattia.

Il nostro lavoro ci ha permesso di identificare 4 associazioni genetiche coincidenti localizzate nei loci genici *HLA*, *IL2RA*, *CIITA* e *ATXN2/SH2B3* (Figure 10b-c-d e Figura 11).

In dettaglio, una variante genetica localizzata a monte del gene *HLA-DRA* è associata alla modulazione dei livelli di cellule T della memoria positive per l'antigene CD8 e negative per la molecola co-stimolatoria CD28 (cellule CD45RA-CD28-CD8+) (p-value= 4.3×10^{-13}) e correla con alcune varianti alleliche descritte in precedenti pubblicazioni come associate negativamente alla colite ulcerosa (UC), alla sclerosi sistemica (SSc), al morbo di Parkinson ed al linfoma di Hodgkin [Barrett J.C. *et al.*, 2009; Enciso-Mora V. *et al.*, 2010; Gorlova O. *et al.*, 2011; Hamza *et al.*, 2011]. La seconda associazione coincidente è stata identificata tra una variante genetica localizzata gene *IL2RA*, che codifica per l'antigene CD25, e le cellule T della memoria che esprimono alti livelli dello stesso CD25 (cellule CD45RA-CD25hi CD4+ non-Treg) (p-value= 5.65×10^{-29}). Questa variante è stata descritta recentemente come la variante più fortemente associata a un ridotto rischio di sviluppare diabete di tipo 1 (T1D) in questa regione genica [Huang J. *et al.*, 2012; Lowe C.E. *et al.*, 2007]. Inoltre, la nostra variante associata è in forte linkage disequilibrium [LD pari a 0.77] con una variante associata sia al T1D che allo stesso tratto immunofenotipico [Dendrou C.A., 2009]. L'allele responsabile dell'aumento delle cellule CD45RA-CD25hi CD4+ non-Treg riduce il rischio per il T1D. L'insieme di queste osservazioni indica che questo specifico tipo cellulare è coinvolto nella protezione contro il T1D e che il top SNP associato probabilmente rappresenta il polimorfismo genetico causale in questo locus (Figura 9). Questo dato ci permette di avanzare l'ipotesi secondo cui terapie immunosoppressive che utilizzano anticorpi anti-CD25 avrebbero come target cellulare una popolazione cellulare protettiva la cui deplezione potrebbe aumentare il rischio per il T1D. Questa ipotesi trova riscontro con i risultati ottenuti in alcuni trials clinici che hanno evidenziato un aumento del rischio di T1D in pazienti trapiantati trattati con anticorpi anti-CD25 [Bayes B. *et al.*, 2007; Vendrame F. *et al.*, 2010].

Un'altra associazione coincidente è stata identificata tra una variante (p-value= 9.29×10^{-30}) localizzata nel gene *CIITA* che è associata alla modulazione dei

livelli di linfociti T HLA-DR+ e risulta essere in forte LD con un allele di rischio per la colite ulcerosa (UC) [McGovern D.P. *et al.*, 2010].

Un ulteriore segnale coincidente è stato individuato per una variante genetica localizzata nel gene *ATXN2* che è associata sia ad un aumento dei livelli di linfociti T e delle cellule T CD4+ (p-value=1.87x10⁻⁰⁹) che ad alcuni disturbi autoimmuni, quali diabete di tipo primo, sclerosi multipla, artrite reumatoide e celiachia, oltre che all'ipertensione e patologie ad essa correlate (disturbi coronarici e disturbi cronici ai reni) [Hindorff L.A.]. Questa variante risulta in forte LD con una variante missenso (R262W) presente nel gene *SH2B3* che costituisce un forte candidato biologico. Infatti, il gene codifica per la proteina adattatrice LNK che svolge il ruolo di regolatore negativo dell'ematopoiesi, delle vie di segnalazione cellulare e dell'infiammazione [Devalliere J. and Charreau B., 2011]. L'allele non ancestrale del gene *SH2B3*, che codifica per W262, è associato con l'aumento dei linfociti CD3 e CD4 facendo ipotizzare una perdita di funzione della proteina. Questo dato è consistente con risultati ottenuti in modelli murini e umani i quali indicano una buona efficacia nell'utilizzo degli anticorpi monoclonali diretti contro l'antigene CD3 nel trattamento del T1D e di altri disturbi autoimmuni [Chatenoud L., 2010]. La variante nel gene *SH2B3* si trova in un aplotipo esteso di circa 200kb, che definisce una regione sottoposta a selezione positiva [Zhernakova A., 2010] risultando essere un esempio suggestivo di "evolutionary tempo mismatch". Infatti, la variante che causa linfocitosi può essere stata utile nel corso degli ultimi 10 mila anni di vita sedentaria per la resistenza ai patogeni, ma in un ambiente sempre più igienicamente controllato, come quello attuale, essa costituisce un fattore di predisposizione per l'autoimmunità e per altri disturbi analoghi.

In conclusione, le associazioni coincidenti tra malattie e tratti immunofenotipici assumono particolare importanza per l'identificazione di nuovi bersagli di interesse terapeutico, infatti alcune delle associazioni coincidenti qui identificate rivestono già un interesse farmacologico per la selezione di particolari targets terapeutici. Allo stesso tempo, è importante sottolineare che il nostro lavoro in alcuni casi non supporta associazioni genetiche identificate in precedenza sulla base di prove funzionali che coinvolgono livelli specifici di tipi cellulari in specifiche malattie. Per esempio, in diversi lavori pubblicati in precedenza è emerso un ruolo protettivo delle cellule Tregs attivate CD39+ CD4+ in diverse malattie autoimmuni [Chalmin F. *et al.*, 2012; Fletcher J.M. *et al.*, 2009], tuttavia nel nostro lavoro non è stata identificata alcuna associazione coincidente tra malattie e varianti genetiche in

grado di influenzare la regolazione quantitativa di questa sottopopolazione cellulare.

CONCLUSIONI

L'approccio metodologico utilizzato nel nostro studio ci ha permesso di fornire la prima chiara evidenza che la variazione dei livelli quantitativi delle principali popolazioni leucocitarie potenzialmente coinvolte nell'insorgenza di malattie è sotto stretto controllo genetico. Ci sono stati diversi fattori che hanno influito positivamente sull'alto grado di ereditabilità mostrato in particolare per alcuni sottotipi cellulari come la scelta di esaminare specifiche sottoclassi cellulari, piuttosto che ampie popolazioni di cellule in cui l'effetto genetico sarebbe diluito. Non meno importante è stata la strategia di analisi genetica utilizzata che si basa sul sequenziamento di oltre 1000 individui sardi e successiva imputazione delle varianti non direttamente genotipizzate, approccio chiamato "sequence-based GWAS", che ci ha permesso di testare mediante GWAS oltre 8.2 milioni di varianti per individuo. A conferma della rilevanza degli studi GWAS basati sul sequenziamento, il nostro progetto ha portato all'identificazione di alcuni segnali di associazione localizzati nei geni *NCAM*, *CD4*, *HLA-E* e *SORBI* e rilevati solo grazie al pannello di riferimento costruito sulla base dei dati di sequenziamento di individui sardi. Infatti, solo 3 delle varianti genetiche genotipizzate nei due array Illumina raggiungeva il livello di significatività utilizzato nel presente studio. Le rimanenti venti sono state identificate mediante inferenza statistica e due di esse non sono presenti sia nel dataset HapMap che nei dati relativi al progetto 1000 Genomi. Da questo si può chiaramente dedurre che queste varianti genetiche non potrebbero essere individuabili con il metodo dell'imputazione sfruttando database pubblici di sequenze non sardo-specifiche. Nonostante questi risultati, sarà comunque necessario eseguire indagini più approfondite in altre popolazioni per determinare se le varianti genetiche da noi identificate siano specifiche degli individui sardi. In conclusione, il nostro studio ha confermato l'importanza di studiare i tratti quantitativi connessi a malattie, permettendo la rilevazione di nuove varianti genetiche e geni correlati alle stesse ed offrendo l'opportunità di mettere in luce una moltitudine di aspetti chiave alla base della fisiologia del sistema immunitario e allo stesso tempo di capire i meccanismi responsabili delle sue alterazioni in caso di malattia.

Pannello	Anticorpo	Fluorocromo	Clone
T-B-NK	anti-CD45	PerCP-Cy5.5	2D1 (HLe-1)
	anti-CD14	V450	MφP9
	anti-CD3	FITC	SK7
	anti-CD4	PE-Cy7	SK3
	anti-CD8	APC-Cy7	SK1
	anti-CD19	APC	SJ25C1
	anti-CD16+CD56	PE	B73.1+NCAM16.2
	anti-HLA DR	V500	G46-6
Tregs	anti-CD3	FITC	SK7
	anti-CD4	APC-H7	SK3
	anti-CD8	V500	RPA-T8
	anti-CD25	PE-Cy7	2A3
	anti-CD45RA	V450	HI100
	anti-CD39	APC	TU66
	anti-CD28	PE	L293
	anti-CD127	PerCP-Cy5.5	hIL-7R-M21
T cell maturati on stages	anti-CD3	FITC	SK7
	anti-CD4	APC	SK3
	anti-CD8	APC-H7	SK1
	anti-CD45RA	V450	HI100
	anti-CCR7	PerCP-Cy5.5	150503

cDCs	anti-Lin cocktail (CD3+CD14+CD16+CD19+CD20+CD56)	PE	SK7+M5E2+B73.1+SJ25C1+L27+NCA M16.2
	anti-HLA DR	V500	G46-6
	anti-CD123	PE-Cy7	7G3
	anti-CD11c	V450	B-Ly6
	anti-CD86	PerCP-Cy5.5	FUN-1
	anti-CD80	APC-H7	L307.4
	anti-CD62L	FITC	SK11
	anti-CCR2	Alexa Fluor 647	48607

Tabella 1. Caratterizzazione degli immunofenotipi.

Lista degli anticorpi, fluorocromi e cloni utilizzati in ciascun pannello.

<i>Categoria</i>	<i>Pannello</i>	<i>Tratto</i>	<i>Parentale</i>	<i>Granparentale</i>	<i>N misurazioni</i>	<i>Minimo</i>	<i>Media</i>	<i>Mediana</i>	<i>Massimo</i>
General leukocytes	TBNK	Granulocyte AC	Leukocytes	no	1612	742.57	3812.21	3586.90	16285.00
General leukocytes	TBNK	Granulocyte %P			1627	23.90	56.15	56.10	85.20
General leukocytes	TBNK	Monocyte AC	Leukocytes	no	1612	0.00	335.92	315.70	1298.23
General leukocytes	TBNK	Monocyte %P			1627	0.00	5.06	4.90	15.50
General leukocytes	TBNK	Lymphocyte AC	Leukocytes	no	1612	485.38	2433.48	2344.30	7650.93
General leukocytes	TBNK	Lymphocyte %P			1627	9.50	36.79	36.90	71.40
General leukocytes	TBNK	T lymphocyte AC	Lymphocytes	Leukocytes	1612	307.40	1695.86	1630.50	5661.44
General leukocytes	TBNK	T lymphocyte %P			1627	21.90	69.61	70.60	90.80
General leukocytes	TBNK	T lymphocyte %GP			1627	6.40	25.67	25.40	57.30
General leukocytes	TBNK	CD4+ AC	T lymphocytes	Lymphocytes	1612	177.98	1062.16	1009.40	3739.70
General leukocytes	TBNK	CD4+ %P			1627	21.60	63.13	63.80	88.10
General leukocytes	TBNK	CD4+ %GP			1627	15.30	43.79	44.00	69.70
General leukocytes	Treg	CD25hi CD4+ AC	CD4+	T lymphocytes	1611	15.44	183.31	167.98	859.51
General leukocytes	Treg	CD25hi CD4+ %P			1628	5.10	17.84	17.20	46.30
General leukocytes	Treg	CD25hi CD4+ %GP			1628	1.30	11.20	10.60	37.40
General leukocytes	Treg	CD39+ CD4+ AC	CD4+	T lymphocytes	1611	3.22	53.99	51.70	413.89
General leukocytes	Treg	CD39+ CD4+ %P			1628	0.30	5.44	5.30	48.20
General leukocytes	Treg	CD39+ CD4+ %GP			1628	0.20	3.41	3.20	31.70
General leukocytes	Treg	CD39+ CD25hi CD4+ AC	CD4+	T lymphocytes	1611	1.93	27.73	26.32	153.26
General leukocytes	Treg	CD39+ CD25hi CD4+ %P			1628	0.20	2.77	2.80	11.00
General	Treg	CD39+ CD25hi			1628	0.10	1.73	1.70	7.20

leukocytes		CD4+ %GP							
General leukocytes	TBNK	CD8br AC	T lymphocytes	Lymphocytes	1612	53.19	479.50	428.48	3255.22
General leukocytes	TBNK	CD8br %P			1627	8.00	27.87	27.10	74.30
General leukocytes	TBNK	CD8br %GP			1627	3.20	19.51	18.80	59.90
General leukocytes	Treg	CD28- CD8br AC	CD8br	T lymphocytes	1611	5.12	194.77	139.15	2405.92
General leukocytes	Treg	CD28- CD8br %P			1628	2.70	37.10	35.75	97.10
General leukocytes	Treg	CD28- CD8br %GP			1628	0.50	11.13	9.00	74.60
General leukocytes	Treg	CD28+ CD8br AC	CD8br	T lymphocytes	1611	27.12	279.33	251.54	1119.31
General leukocytes	Treg	CD28+ CD8br %P			1628	2.70	61.92	63.10	97.20
General leukocytes	Treg	CD28+ CD8br %GP			1628	2.10	16.44	15.90	38.40
General leukocytes	TBNK	CD8dim AC	T lymphocytes	Lymphocytes	1612	2.38	45.47	34.84	615.90
General leukocytes	TBNK	CD8dim %P			1627	0.20	2.68	2.20	24.10
General leukocytes	TBNK	CD8dim %GP			1627	0.20	1.87	1.50	20.70
General leukocytes	Treg	CD28- CD8dim AC	CD8dim	T lymphocytes	1611	0.29	27.07	15.48	580.63
General leukocytes	Treg	CD28- CD8dim %P			1628	1.50	46.71	44.90	97.70
General leukocytes	Treg	CD28- CD8dim %GP			1628	0.00	1.56	0.90	22.70
General leukocytes	Treg	CD28+ CD8dim AC	CD8dim	T lymphocytes	1611	0.82	21.43	17.37	248.71
General leukocytes	Treg	CD28+ CD8dim %P			1628	1.30	51.79	53.10	98.50
General leukocytes	Treg	CD28+ CD8dim %GP			1628	0.10	1.30	1.10	14.40
General leukocytes	TBNK	CD8+ AC	T lymphocytes	Lymphocytes	1612	55.89	523.68	469.20	3535.04
General leukocytes	TBNK	CD8+ %P			1627	8.40	30.47	29.90	75.30

General leukocytes	TBNK	CD8+ %GP			1627	3.50	21.33	20.70	63.80
General leukocytes	Treg	CD28- CD8+ AC	CD8+	T lymphocytes	1611	5.56	221.12	163.80	2829.62
General leukocytes	Treg	CD28- CD8+ %P			1611	2.45	38.75	38.22	96.90
General leukocytes	Treg	CD28- CD8+ %GP			1628	0.60	12.64	10.60	74.90
General leukocytes	Treg	CD39+ CD8+ AC	CD8+	T lymphocytes	1611	0.00	9.45	6.84	164.69
General leukocytes	Treg	CD39+ CD8+ %P			1628	0.00	2.10	1.40	40.30
General leukocytes	Treg	CD39+ CD8+ %GP			1628	0.00	0.60	0.40	15.80
General leukocytes	TBNK	CD4+ CD8br AC	T lymphocytes	Lymphocytes	1612	0.00	7.19	4.78	115.52
General leukocytes	TBNK	CD4+ CD8br %P			1627	0.00	0.43	0.30	8.50
General leukocytes	TBNK	CD4+ CD8br %GP			1627	0.00	0.30	0.20	5.30
General leukocytes	TBNK	CD4+ CD8dim AC	T lymphocytes	Lymphocytes	1612	0.00	19.11	9.35	1104.94
General leukocytes	TBNK	CD4+ CD8dim %P			1627	0.00	1.09	0.60	35.00
General leukocytes	TBNK	CD4+ CD8dim %GP			1627	0.00	0.76	0.40	24.80
General leukocytes	TBNK	CD4+ CD8+ AC	T lymphocytes	Lymphocytes	1612	0.00	26.30	15.90	1106.56
General leukocytes	TBNK	CD4+ CD8+ %P			1627	0.00	1.52	1.00	35.50
General leukocytes	TBNK	CD4+ CD8+ %GP			1627	0.00	1.06	0.70	25.10
General leukocytes	TBNK	CD4- CD8- AC	T lymphocytes	Lymphocytes	1612	5.81	84.09	65.24	1295.67
General leukocytes	TBNK	CD4- CD8- %P			1627	0.30	4.91	4.00	36.80
General leukocytes	TBNK	CD4- CD8- %GP			1627	0.30	3.44	2.80	27.30
General leukocytes	TBNK	HLA DR+ T lymphocyte AC	T lymphocytes	Lymphocytes	1550	0.00	67.07	55.00	881.43
General	TBNK	HLA DR+ T			1565	0.00	4.04	3.40	18.60

leukocytes		lymphocyte %P							
General leukocytes	TBNK	HLA DR+ T lymphocyte %GP			1565	0.00	2.80	2.30	16.40
General leukocytes	TBNK	NKT AC	T lymphocytes	Lymphocytes	1612	4.86	174.51	125.06	3991.06
General leukocytes	TBNK	NKT %P			1627	0.30	10.03	8.00	81.20
General leukocytes	TBNK	NKT %GP			1627	0.20	7.04	5.50	63.30
General leukocytes	TBNK	CD3- lymphocyte AC	Lymphocytes	Leukocytes	1612	120.70	737.61	683.90	5975.80
General leukocytes	TBNK	CD3- lymphocyte %P			1627	9.20	30.39	29.40	78.10
General leukocytes	TBNK	CD3- lymphocyte %GP			1627	2.40	11.11	10.50	49.70
General leukocytes	TBNK	B cell AC	CD3- lymphocytes	Lymphocytes	1612	16.97	318.75	273.73	5327.86
General leukocytes	TBNK	B cell %P			1627	4.40	43.56	43.00	91.80
General leukocytes	TBNK	B cell %GP			1627	0.80	12.89	12.20	69.60
General leukocytes	TBNK	NK AC	CD3- lymphocytes	Lymphocytes	1612	25.26	382.93	337.60	3864.66
General leukocytes	TBNK	NK %P			1627	6.80	51.20	51.20	89.00
General leukocytes	TBNK	NK %GP			1627	1.70	15.96	14.60	65.80
General leukocytes	TBNK	HLA DR+ NK AC	NK cells	CD3- lymphocytes	1550	0.00	34.94	28.39	290.10
General leukocytes	TBNK	HLA DR+ NK %P			1565	0.00	9.82	8.20	65.60
General leukocytes	TBNK	HLA DR+ NK %GP			1565	0.00	4.88	4.00	26.20
General leukocytes	TBNK	Lympho-sum AC	Lymphocytes	Leukocytes	1612	451.59	2394.98	2312.10	7634.04
General leukocytes	TBNK	Lympho-sum %P			1627	83.50	98.36	98.70	100.00
General leukocytes	TBNK	Lympho-sum %GP			1627	9.30	36.20	36.40	71.10
General leukocytes	TBNK	CD4+/CD8br	no	no	1612	0.29	2.60	2.32	10.79

General leukocytes	TBNK	CD4+/CD8+	no	no	1612	0.29	2.36	2.13	10.35
General leukocytes	TBNK	T/B cells	no	no	1612	0.31	6.61	5.77	97.07
Tregs	Treg	CD4+ Treg AC	CD4+	T lymphocytes	1611	9.45	80.86	74.91	390.32
Tregs	Treg	CD4+ Treg %P			1628	1.30	7.88	7.60	22.20
Tregs	Treg	CD4+ Treg %GP			1628	0.60	4.94	4.80	15.60
Tregs	Treg	resting CD4+ Treg AC	CD4+ Tregs	CD4+	1611	0.02	16.07	13.14	186.96
Tregs	Treg	resting CD4+ Treg %P			1628	0.20	19.16	17.70	83.60
Tregs	Treg	resting CD4+ Treg %GP			1628	0.00	1.50	1.30	9.20
Tregs	Treg	CD39+ resting CD4+ Treg AC	resting CD4+ Tregs	CD4+ Tregs	1611	0.00	0.56	0.31	14.36
Tregs	Treg	CD39+ resting CD4+ Treg %P			1628	0.00	3.68	2.70	50.00
Tregs	Treg	CD39+ resting CD4+ Treg %GP			1628	0.00	0.69	0.40	18.20
Tregs	Treg	activated CD4+ Treg AC	CD4+ Tregs	CD4+	1611	0.82	14.61	13.02	71.48
Tregs	Treg	activated CD4+ Treg %P			1628	4.00	18.22	17.70	42.50
Tregs	Treg	activated CD4+ Treg %GP			1628	0.10	1.46	1.30	7.20
Tregs	Treg	CD39+ activated CD4+ Treg AC	activated CD4+ Tregs	CD4+ Tregs	1611	0.40	8.45	7.69	53.79
Tregs	Treg	CD39+ activated CD4+ Treg %P			1628	4.00	58.24	71.30	99.00
Tregs	Treg	CD39+ activated CD4+ Treg %GP			1628	0.40	10.65	11.00	35.50
Tregs	Treg	secreting CD4+ Treg AC	CD4+ Tregs	CD4+	1611	2.44	51.78	48.21	190.02
Tregs	Treg	secreting CD4+ Treg %P			1628	5.60	64.60	65.30	92.20
Tregs	Treg	secreting CD4+ Treg %GP			1628	0.50	5.08	4.90	12.50
Tregs	Treg	CD39+ secreting CD4+ Treg AC	secreting CD4+ Tregs	CD4+ Tregs	1611	0.50	12.05	11.16	74.79

Tregs	Treg	CD39+ secreting CD4+ Treg %P			1628	0.90	23.99	25.25	71.20
Tregs	Treg	CD39+ secreting CD4+ Treg %GP			1628	0.60	15.44	16.10	58.30
Tregs	Treg	CD39+ CD4+ Treg AC	CD4+ Tregs	CD4+	1611	1.58	21.65	20.48	122.86
Tregs	Treg	CD39+ CD4+ Treg %P			1628	2.30	27.52	30.25	74.80
Tregs	Treg	CD39+ CD4+ Treg %GP			1628	0.10	2.16	2.10	9.00
Tregs	Treg	secreting & activated CD4+ Treg AC	CD4+ Tregs	CD4+	1611	4.74	66.40	62.11	221.69
Tregs	Treg	secreting & activated CD4+ Treg %P			1624	10.90	82.78	84.30	99.60
Tregs	Treg	secreting & activated CD4+ Treg %GP			1628	0.90	6.54	6.30	19.70
Tregs	Treg	CD39+ (secreting & activated) CD4+ Treg AC	CD4+ Tregs	CD4+	1611	1.24	20.49	19.42	114.44
Tregs	Treg	CD39+ (secreting & activated) CD4+ Treg %P			1611	1.90	26.03	28.47	77.09
Tregs	Treg	CD39+ (secreting & activated) CD4+ Treg %GP			1611	0.07	2.04	2.02	7.66
Tregs	Treg	resting & activated CD4+ Treg AC	CD4+ Tregs	CD4+	1611	1.86	30.68	27.48	258.39
Tregs	Treg	resting & activated CD4+ Treg %P			1628	9.10	37.38	36.60	87.90
Tregs	Treg	resting & activated CD4+ Treg %GP			1628	0.30	2.96	2.80	12.70
Tregs	Treg	CD39+ (resting & activated) CD4+ Treg AC	CD4+ Tregs	CD4+	1611	0.47	9.00	8.16	54.61
Tregs	Treg	CD39+ (resting & activated) CD4+ Treg %P			1611	0.54	11.30	11.83	36.14
General leukocytes	Treg	CD4+ not Treg AC	CD4+	T lymphocytes	1611	172.22	970.08	921.80	3626.51

General leukocytes	Treg	CD4+ not Treg %P			1628	77.80	92.12	92.40	98.70
General leukocytes	Treg	CD4+ not Treg %GP			1628	18.50	57.61	58.10	85.90
T cell maturation stages	Maturation stages of T cell	naive CD4+ AC	CD4+	T lymphocytes	1601	0.46	367.51	323.82	1847.16
T cell maturation stages	Maturation stages of T cell	naive CD4+ %P			1618	0.25	33.29	33.25	72.01
T cell maturation stages	Maturation stages of T cell	naive CD4+ %GP			1618	0.10	21.22	20.80	55.40
T cell maturation stages	Maturation stages of T cell	CM CD4+ AC	CD4+	T lymphocytes	1601	1.93	318.50	298.38	1339.57
T cell maturation stages	Maturation stages of T cell	CM CD4+ %P			1618	0.19	30.57	30.11	64.48
T cell maturation stages	Maturation stages of T cell	CM CD4+ %GP			1618	0.20	19.30	18.60	45.70
T cell maturation stages	Maturation stages of T cell	EM CD4+ AC	CD4+	T lymphocytes	1601	40.37	330.70	312.49	1499.00
T cell maturation stages	Maturation stages of T cell	EM CD4+ %P			1618	4.84	32.51	31.05	86.89
T cell maturation stages	Maturation stages of T cell	EM CD4+ %GP			1618	3.00	20.11	19.10	58.90
T cell maturation stages	Maturation stages of T cell	TD CD4+ AC	CD4+	T lymphocytes	1601	0.49	39.59	23.37	3068.41
T cell maturation stages	Maturation stages of T cell	TD CD4+ %P			1618	0.06	3.64	2.37	75.58
T cell maturation stages	Maturation stages of T cell	TD CD4+ %GP			1618	0.00	2.25	1.50	54.20
T cell maturation stages	Maturation stages of T cell	CD45RA- CD4+ AC	CD4+	T lymphocytes	1601	52.41	649.20	621.47	2023.87
T cell maturation stages	Maturation stages of T cell	CD45RA- CD4+ %P			1618	9.11	63.08	63.17	97.53
T cell maturation stages	Maturation stages of T cell	CD45RA- CD4+			1618	4.30	39.41	38.50	75.30

stages	cell	%GP							
T cell maturation stages	Maturation stages of T cell	CD45RA+ CD4+ AC	CD4+	T lymphocytes	1611	7.09	402.67	355.69	2711.61
T cell maturation stages	Maturation stages of T cell	CD45RA+ CD4+ %P			1628	2.80	36.71	36.75	92.40
T cell maturation stages	Maturation stages of T cell	CD45RA+ CD4+ %GP			1628	1.20	23.23	22.70	61.40
T cell maturation stages	Treg	CD45RA- CD4+ not Treg AC	CD4+	T lymphocytes	1611	32.63	581.75	553.64	1870.05
T cell maturation stages	Treg	CD45RA- CD4+ not Treg %P			1628	6.30	56.75	56.60	91.00
T cell maturation stages	Treg	CD45RA- CD4+ not Treg %GP			1628	2.60	35.23	34.55	71.00
T cell maturation stages	Treg	CD45RA+ CD4+ not Treg AC	CD4+	T lymphocytes	1611	7.09	388.31	340.96	2707.94
T cell maturation stages	Treg	CD45RA+ CD4+ not Treg %P			1628	2.60	35.37	35.40	85.30
T cell maturation stages	Treg	CD45RA+ CD4+ not Treg %GP			1628	1.20	22.38	21.85	57.50
T cell maturation stages	Treg	CD45RA- CD25hi CD4+ not Treg AC	CD4+	T lymphocytes	1611	5.81	93.08	83.40	429.57
T cell maturation stages	Treg	CD45RA- CD25hi CD4+ not Treg %P			1628	1.20	9.08	8.40	32.80
T cell maturation stages	Treg	CD45RA- CD25hi CD4+ not Treg %GP			1628	0.50	5.69	5.20	24.30
T cell maturation stages	Treg	CD45RA+ CD25hi CD4+ not Treg AC	CD4+	T lymphocytes	1611	0.00	9.54	6.46	200.22
T cell maturation stages	Treg	CD45RA+ CD25hi CD4+ not Treg %P			1628	0.00	0.89	0.60	14.70
T cell maturation stages	Treg	CD45RA+ CD25hi CD4+ not Treg %GP			1628	0.00	0.58	0.40	8.60
T cell maturation stages	Maturation stages of T cell	naive CD8br AC	CD8br	T lymphocytes	1601	0.00	120.53	92.77	887.70
T cell maturation stages	Maturation stages of T	naive CD8br %P			1618	0.00	25.64	22.71	81.94

	cell								
T cell maturation stages	Maturation stages of T cell	naive CD8br %GP			1618	0.00	6.85	5.95	29.10
T cell maturation stages	Maturation stages of T cell	CM CD8br AC	CD8br	T lymphocytes	1601	0.00	31.48	25.91	180.69
T cell maturation stages	Maturation stages of T cell	CM CD8br %P			1618	0.00	7.15	6.05	52.96
T cell maturation stages	Maturation stages of T cell	CM CD8br %GP			1618	0.00	1.91	1.60	15.30
T cell maturation stages	Maturation stages of T cell	EM CD8br AC	CD8br	T lymphocytes	1601	1.13	206.90	174.39	1463.56
T cell maturation stages	Maturation stages of T cell	EM CD8br %P			1618	0.39	42.67	41.56	88.11
T cell maturation stages	Maturation stages of T cell	EM CD8br %GP			1618	0.10	12.18	11.00	51.40
T cell maturation stages	Maturation stages of T cell	TD CD8br AC	CD8br	T lymphocytes	1601	1.42	127.53	89.74	3275.97
T cell maturation stages	Maturation stages of T cell	TD CD8br %P			1618	0.49	24.54	21.75	91.38
T cell maturation stages	Maturation stages of T cell	TD CD8br %GP			1618	0.10	7.36	5.70	66.60
T cell maturation stages	Maturation stages of T cell	CD45RA- CD8br AC	CD8br	T lymphocytes	1601	9.26	238.38	203.18	1546.17
T cell maturation stages	Maturation stages of T cell	CD45RA- CD8br %P			1618	1.84	49.83	49.38	94.06
T cell maturation stages	Maturation stages of T cell	CD45RA- CD8br %GP			1618	0.60	14.08	12.90	54.60
T cell maturation stages	Treg	CD45RA- CD28- CD8br AC	CD28- CD8br	CD8br	1610	1.19	98.87	67.00	1697.27
T cell maturation stages	Treg	CD45RA- CD28- CD8br %P			1610	0.30	51.14	52.04	97.49
T cell maturation stages	Treg	CD45RA- CD28- CD8br %GP			1610	0.24	18.82	15.89	77.81

T cell maturation stages	Treg	CD45RA+ CD28-CD8br AC	CD28-CD8br	CD8br	1611	1.39	95.97	62.11	2253.85
T cell maturation stages	Treg	CD45RA+ CD28-CD8br %P			1610	2.51	48.86	47.97	99.70
T cell maturation stages	Treg	CD45RA+ CD28-CD8br %GP			1611	0.76	18.32	15.21	78.96
T cell maturation stages	Maturation stages of T cell	naive CD8dim AC	CD8dim	T lymphocytes	1601	0.00	1.67	1.12	15.91
T cell maturation stages	Maturation stages of T cell	naive CD8dim %P			1618	0.00	4.94	3.21	52.78
T cell maturation stages	Maturation stages of T cell	naive CD8dim %GP			1618	0.00	0.10	0.10	0.90
T cell maturation stages	Maturation stages of T cell	CM CD8dim AC	CD8dim	T lymphocytes	1601	0.00	0.71	0.44	9.86
T cell maturation stages	Maturation stages of T cell	CM CD8dim %P			1618	0.00	2.19	1.21	31.25
T cell maturation stages	Maturation stages of T cell	CM CD8dim %GP			1618	0.00	0.04	0.00	0.60
T cell maturation stages	Maturation stages of T cell	EM CD8dim AC	CD8dim	T lymphocytes	1601	0.37	26.47	18.54	442.42
T cell maturation stages	Maturation stages of T cell	EM CD8dim %P			1618	0.69	55.66	58.24	97.03
T cell maturation stages	Maturation stages of T cell	EM CD8dim %GP			1618	0.00	1.55	1.20	16.90
T cell maturation stages	Maturation stages of T cell	TD CD8dim AC	CD8dim	T lymphocytes	1601	0.00	18.88	11.18	448.68
T cell maturation stages	Maturation stages of T cell	TD CD8dim %P			1618	0.00	37.21	33.33	97.93
T cell maturation stages	Maturation stages of T cell	TD CD8dim %GP			1618	0.00	1.11	0.70	17.50
T cell maturation stages	Maturation stages of T cell	CD45RA-CD8dim AC	CD8dim	T lymphocytes	1601	0.37	27.17	19.25	442.91
T cell maturation stages	Maturation stages of T cell	CD45RA-CD8dim %P			1618	0.69	57.85	60.66	98.99

	cell								
T cell maturation stages	Maturation stages of T cell	CD45RA- CD8dim %GP			1618	0.00	1.59	1.20	16.90
T cell maturation stages	Treg	CD45RA- CD28- CD8dim AC	CD8dim		1581	0.01	12.34	4.90	442.92
T cell maturation stages	Treg	CD45RA- CD28- CD8dim %P			1581	0.03	19.33	14.42	91.76
T cell maturation stages	Treg	CD45RA+ CD28- CD8dim AC	CD8dim	T lymphocytes	1611	0.00	14.97	8.05	433.05
T cell maturation stages	Treg	CD45RA+ CD28- CD8dim %P			1611	0.00	27.84	23.08	94.20
T cell maturation stages	Treg	CD45RA+ CD28- CD8dim %GP			1628	0.00	0.87	0.50	16.80
T cell maturation stages	Maturation stages of T cell	naive CD8+ AC	CD8+	T lymphocytes	1601	0.00	122.01	93.65	893.55
T cell maturation stages	Maturation stages of T cell	naive CD8+ %P			1618	0.00	23.67	20.87	78.03
T cell maturation stages	Maturation stages of T cell	naive CD8+ %GP			1618	0.00	6.93	6.10	29.50
T cell maturation stages	Maturation stages of T cell	CM CD8+ AC	CD8+	T lymphocytes	1601	0.00	31.94	26.12	172.28
T cell maturation stages	Maturation stages of T cell	CM CD8+ %P			1618	0.00	6.64	5.56	50.26
T cell maturation stages	Maturation stages of T cell	CM CD8+ %GP			1618	0.00	1.93	1.60	15.30
T cell maturation stages	Maturation stages of T cell	EM CD8+ AC	CD8+	T lymphocytes	1601	7.97	231.93	199.75	1469.79
T cell maturation stages	Maturation stages of T cell	EM CD8+ %P			1618	1.53	43.79	42.95	87.38
T cell maturation stages	Maturation stages of T cell	EM CD8+ %GP			1618	0.60	13.65	12.50	52.50
T cell maturation stages	Maturation stages of T cell	TD CD8+ AC	CD8+	T lymphocytes	1601	4.91	146.15	104.62	3365.75
T cell maturation stages	Maturation stages of T cell	TD CD8+ %P			1618	1.46	25.90	23.18	91.12

stages	cell								
T cell maturation stages	Maturation stages of T cell	TD CD8+ %GP			1618	0.40	8.46	6.80	68.40
T cell maturation stages	Maturation stages of T cell	CD45RA- CD8+ AC	CD8+	T lymphocytes	1601	10.41	263.87	229.39	1581.87
T cell maturation stages	Maturation stages of T cell	CD45RA- CD8+ %P			1618	1.85	50.42	50.17	92.48
T cell maturation stages	Maturation stages of T cell	CD45RA- CD8+ %GP			1618	0.70	15.58	14.50	55.80
T cell maturation stages	Maturation stages of T cell	CD45RA+ CD8+ AC	CD8+	T lymphocytes	1611	16.40	248.10	218.38	3168.82
T cell maturation stages	Maturation stages of T cell	CD45RA+ CD8+ %P			1628	6.90	46.27	46.30	97.40
T cell maturation stages	Maturation stages of T cell	CD45RA+ CD8+ %GP			1628	1.60	14.18	13.40	64.30
T cell maturation stages	Treg	CD45RA- CD28- CD8+ AC	CD28- CD8+	CD8+	1610	1.00	110.98	77.54	1699.45
T cell maturation stages	Treg	CD45RA- CD28- CD8+ %P			1610	0.14	50.68	51.30	97.13
T cell maturation stages	Treg	CD45RA- CD28- CD8+ %GP			1610	0.12	19.43	16.65	75.45
T cell maturation stages	Treg	CD45RA- CD28+ CD8+ AC	CD45RA- CD8+	CD8+	1611	8.56	167.79	152.37	706.08
T cell maturation stages	Treg	CD45RA- CD28+ CD8+ %P			1610	3.89	65.14	66.99	97.86
T cell maturation stages	Treg	CD45RA- CD28+ CD8+ %GP			1611	1.97	34.32	32.80	81.97
T cell maturation stages	Treg	CD45RA+ CD28- CD8+ AC	CD28- CD8+	CD8+	1611	2.09	110.94	75.45	2686.90
T cell maturation stages	Treg	CD45RA+ CD28- CD8+ %P			1610	3.25	49.69	49.42	100.00
T cell maturation stages	Treg	CD45RA+ CD28- CD8+ %GP			1611	0.85	19.46	16.54	81.48
T cell maturation stages	Treg	CD45RA+ CD28+ CD8+ AC	CD45RA+ CD8+	CD8+	1611	3.33	137.16	110.59	874.96
T cell maturation stages	Treg	CD45RA+ CD28+ CD8+ %P			1611	0.91	56.29	56.97	97.95

T cell maturation stages	Treg	CD45RA+ CD28+ CD8+ %GP			1611	0.26	26.81	24.16	79.36
T cell maturation stages	Maturation stages of T cell	naive (CD4+ CD8+) AC	CD4+ CD8+	T lymphocytes	1601	0.00	2.18	1.39	55.26
T cell maturation stages	Maturation stages of T cell	naive (CD4+ CD8+) %P			1618	0.00	12.03	8.56	65.22
T cell maturation stages	Maturation stages of T cell	naive (CD4+ CD8+) %GP			1618	0.00	0.12	0.10	1.70
T cell maturation stages	Maturation stages of T cell	CM (CD4+ CD8+) AC	CD4+ CD8+	T lymphocytes	1601	0.00	3.14	2.00	44.29
T cell maturation stages	Maturation stages of T cell	CM (CD4+ CD8+) %P			1618	0.00	16.74	14.00	75.00
T cell maturation stages	Maturation stages of T cell	CM (CD4+ CD8+) %GP			1618	0.00	0.19	0.10	1.80
T cell maturation stages	Maturation stages of T cell	EM (CD4+ CD8+) AC	CD4+ CD8+	T lymphocytes	1601	0.00	14.82	9.33	373.91
T cell maturation stages	Maturation stages of T cell	EM (CD4+ CD8+) %P			1618	0.00	59.64	60.71	100.00
T cell maturation stages	Maturation stages of T cell	EM (CD4+ CD8+) %GP			1618	0.00	0.87	0.60	23.80
T cell maturation stages	Maturation stages of T cell	TD (CD4+ CD8+) AC	CD4+ CD8+	T lymphocytes	1601	0.00	4.35	0.91	714.68
T cell maturation stages	Maturation stages of T cell	TD (CD4+ CD8+) %P			1618	0.00	11.60	6.00	89.39
T cell maturation stages	Maturation stages of T cell	TD (CD4+ CD8+) %GP			1618	0.00	0.23	0.10	24.40
T cell maturation stages	Maturation stages of T cell	CD45RA- (CD4+ CD8+) AC	CD4+ CD8+	T lymphocytes	1601	0.12	17.96	12.35	376.73
T cell maturation stages	Maturation stages of T cell	CD45RA- (CD4+ CD8+) %P			1618	3.85	76.37	80.00	100.00
T cell maturation stages	Maturation stages of T cell	CD45RA- (CD4+ CD8+) %GP			1618	0.00	1.06	0.80	24.00

T cell maturation stages	Maturation stages of T cell	naive (CD4- CD8-) AC	CD4- CD8-	T lymphocytes	1601	0.00	3.20	2.08	27.07
T cell maturation stages	Maturation stages of T cell	naive (CD4- CD8-) %P			1618	0.00	5.17	3.41	36.36
T cell maturation stages	Maturation stages of T cell	naive (CD4- CD8-) %GP			1618	0.00	0.19	0.10	1.50
T cell maturation stages	Maturation stages of T cell	CM (CD4- CD8-) AC	CD4- CD8-	T lymphocytes	1601	0.00	1.08	0.78	24.61
T cell maturation stages	Maturation stages of T cell	CM (CD4- CD8-) %P			1618	0.00	1.87	1.29	21.43
T cell maturation stages	Maturation stages of T cell	CM (CD4- CD8-) %GP			1618	0.00	0.06	0.00	1.00
T cell maturation stages	Maturation stages of T cell	EM (CD4- CD8-) AC	CD4- CD8-	T lymphocytes	1601	0.48	51.94	36.72	789.37
T cell maturation stages	Maturation stages of T cell	EM (CD4- CD8-) %P			1618	1.12	62.79	65.71	97.71
T cell maturation stages	Maturation stages of T cell	EM (CD4- CD8-) %GP			1618	0.00	3.04	2.30	25.90
T cell maturation stages	Maturation stages of T cell	TD (CD4- CD8-) AC	CD4- CD8-	T lymphocytes	1601	0.59	24.01	14.95	378.24
T cell maturation stages	Maturation stages of T cell	TD (CD4- CD8-) %P			1618	1.40	30.18	25.98	97.02
T cell maturation stages	Maturation stages of T cell	TD (CD4- CD8-) %GP			1618	0.00	1.39	0.90	19.10
T cell maturation stages	Maturation stages of T cell	CD45RA- (CD4- CD8-) AC	CD4- CD8-	T lymphocytes	1601	0.48	53.03	37.78	789.37
T cell maturation stages	Maturation stages of T cell	CD45RA- (CD4- CD8-) %P			1618	1.49	64.65	67.70	98.03
T cell maturation stages	Maturation stages of T cell	CD45RA- (CD4- CD8-) %GP			1618	0.00	3.10	2.40	25.90
cDCs	cDC	cDC AC	no	no	1603	0.18	24.61	21.95	116.55
cDCs	cDC	CD62L- cDC AC	cDCs	no	1603	0.04	4.52	1.45	93.46

cDCs	cDC	CD62L- cDC %P			1627	0.80	13.51	7.20	100.00
cDCs	cDC	CD62L+ cDC AC	cDCs	no	1603	0.00	20.09	18.96	69.71
cDCs	cDC	CD62L+ cDC %P			1627	0.00	86.49	92.80	99.70
cDCs	cDC	myeloid cDC AC	cDCs	no	1603	0.18	15.34	12.58	106.02
cDCs	cDC	myeloid cDC %P			1627	8.20	55.61	55.10	95.00
cDCs	cDC	CD62L- myeloid cDC AC	cDCs	no	1603	0.04	4.34	1.26	93.38
cDCs	cDC	CD62L- myeloid cDC %P			1627	0.30	11.68	5.80	84.90
cDCs	cDC	CD62L+ myeloid cDC AC	cDCs	no	1603	0.00	11.00	10.14	51.92
cDCs	cDC	CD62L+ myeloid cDC %P			1627	0.00	43.92	44.90	81.10
cDCs	cDC	plasmacytoid cDC AC	cDCs	no	1603	0.00	9.27	8.39	37.06
cDCs	cDC	plasmacytoid cDC %P			1627	0.00	36.45	37.00	78.30
cDCs	cDC	CD62L- plasmacytoid cDC AC	cDCs	no	1603	0.00	0.18	0.11	3.56
cDCs	cDC	CD62L- plasmacytoid cDC %P			1627	0.00	0.72	0.50	12.10
cDCs	cDC	CD62L+ plasmacytoid cDC AC	cDCs	no	1603	0.00	9.09	8.19	36.62
cDCs	cDC	CD62L+ plasmacytoid cDC %P			1627	0.00	35.73	36.30	77.30
cDCs	cDC	CD86+ cDC AC	cDCs	no	1603	0.05	3.91	1.52	79.26
cDCs	cDC	CD86+ cDC %P			1627	0.20	12.05	7.50	84.80

Tabella 2. Descrizione statistica degli immunofenotipi.

Da destra verso sinistra sono mostrati per ciascun tratto, la categoria biologica, il pannello sul quale è stato misurato, il nome del tratto, le popolazioni parentali e gran parentali (quando opportuno), il numero di individui in cui il tratto è stato misurato e la sua descrizione statistica.

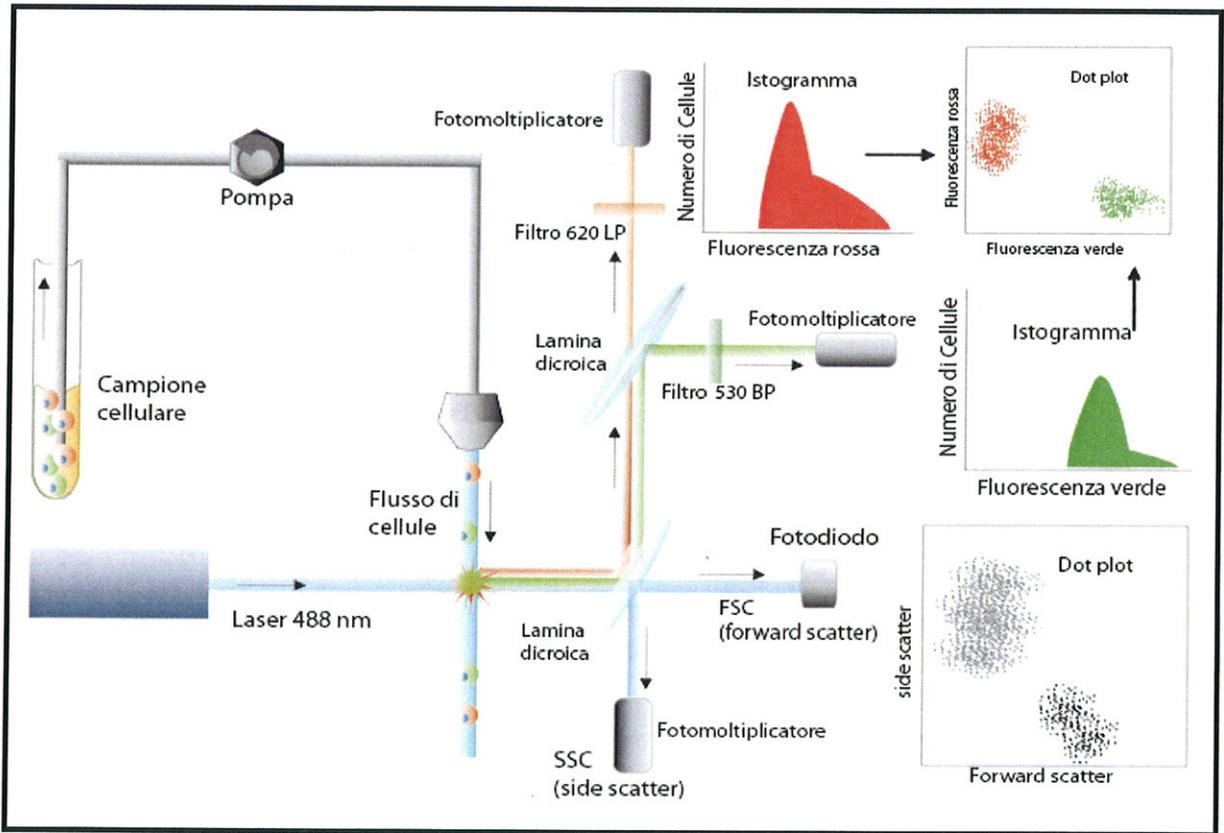


Figura 1. *Rappresentazione schematica del funzionamento di un citometro a flusso (biotec.casaccia.enea.it).*

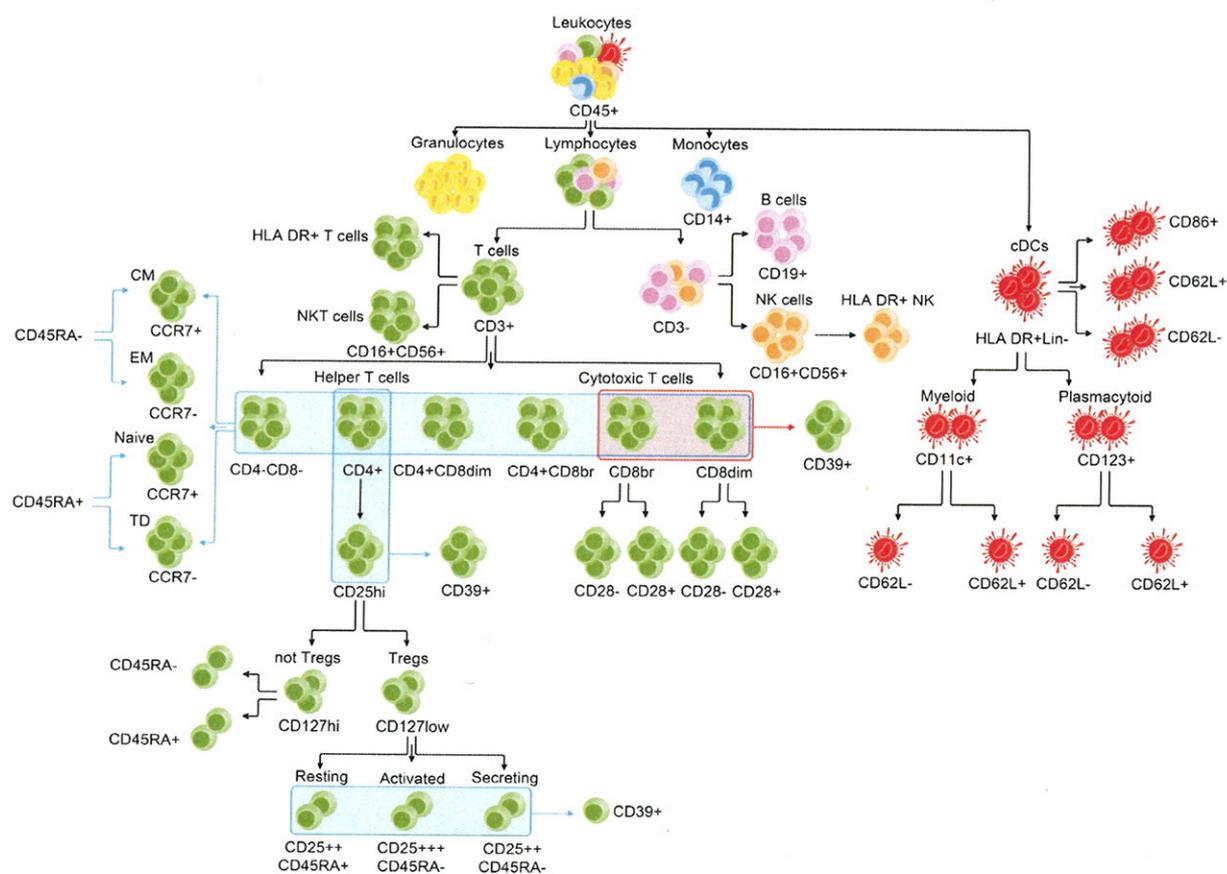


Figura 2. Sottopopolazioni leucocitarie oggetto di studio.

La figura rappresenta il diagramma dei tipi cellulari analizzati mediante citometria a flusso. Le frecce indicano i livelli gerarchici di separazione delle popolazioni cellulari leucocitarie circolanti. Le cellule illustrate includono i granulociti (giallo), i monociti (blu chiaro), le cellule dendritiche (rosso), le cellule B (magenta), le cellule T (verde) e le cellule natural killer (arancione). Il nome e, quando rilevante, l'antigene identificativo sono indicati accanto a ciascuna popolazione. Le cellule presenti all'interno del rettangolo blu chiaro sono fenotipicamente caratterizzate dall'antigene indicato dalla freccia blu chiaro presente a fianco; per esempio, i 6 sottotipi CD3⁺ (CD4-CD8-, CD4+, CD4+CD8dim, CD4+CD8br, CD8br, e CD8dim) sono mostrati all'interno di un ampio rettangolo blu, e sono stati suddivisi in cellule naive, cellule della memoria centrale, cellule della memoria effettiva e cellule "terminally differentiated", generando un totale di 24 tratti. Il rettangolo rosso indica che, le popolazioni cellulari incluse al suo interno, sono state analizzate insieme per l'antigene CD39, il marcatore indicato dalla freccia rossa. Per semplicità, sono stati

mostrati 46 dei 92 tratti cellulari primari descritti nel testo. Il numero totale dei tratti risulta essere 272 quando ciascuna popolazione all'interno di un rettangolo blu viene moltiplicato per il numero di sottotipi cellulari indicati dalle frecce blu chiaro.

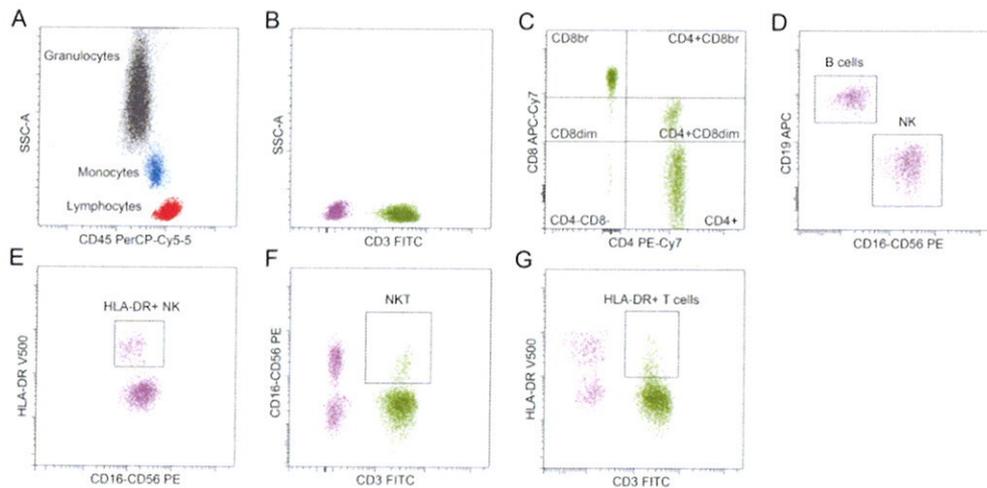


Figura 3. *Pannello T-B-NK.*

Questa figura illustra la strategia di analisi adottata per il pannello delle T-B-NK su un campione di sangue rappresentativo. A) I linfociti (rosso), i monociti (blu) ed i granulociti (grigio) sono stati identificati grazie alle loro caratteristiche morfologiche ed al livello di espressione dell'antigene CD45. B) I linfociti sono stati suddivisi nelle popolazioni CD3⁺ (cellule T, verde) e CD3⁻ (rosa). C) Le cellule T sono state ulteriormente distinte in base ai livelli di espressione degli antigeni CD4 e CD8. Il plot è stato diviso in 6 popolazioni cellulari. A partire dal quadrante in basso a sinistra e procedendo in senso orario, le popolazioni identificate sono le seguenti: CD4-CD8⁻, CD4-CD8dim (CD8dim), CD4-CD8bright (CD8br), CD4+CD8⁺, CD4+CD8dim, CD4+CD8⁻ (CD4⁺). D) Le cellule CD3 negative sono state divise in cellule B ed NK in base ai livelli di espressione degli antigeni CD19 e CD16 e/o CD56, rispettivamente. E) Il marcatore di attivazione HLA-DR è stato analizzato sulle cellule NK. F) Le cellule NKT sono state isolate considerando la loro positività per gli antigeni CD3, CD16 e/o CD56. G) Le cellule T attivate sono state identificate per la loro positività agli antigeni CD3 e HLA-DR.

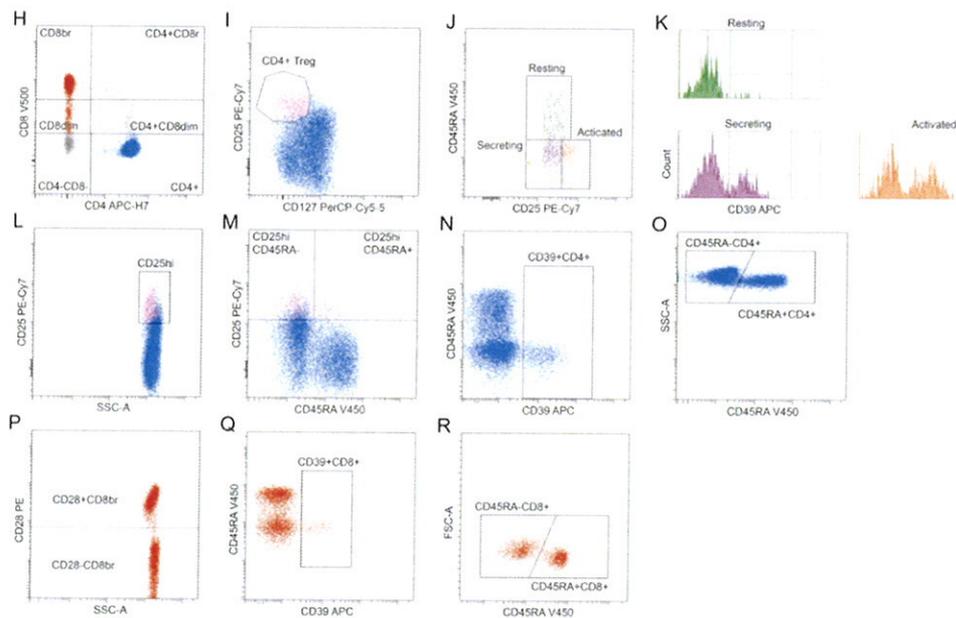


Figura 4. *Pannello cellule T regolatorie.*

Questa figura illustra la strategia di analisi adottata per il pannello delle cellule T regolatorie su un campione di sangue rappresentativo. H) Le cellule T sono state suddivise in base al livello di espressione degli antigeni CD4 e CD8. I) Il sottotipo CD4+ (blu) è stato analizzato per i marcatori CD127 e CD25. Le cellule CD25hiCD127low sono state considerate CD4+Treg (rosa). J) Le cellule CD4+ Treg sono state divise in resting (CD45+CD25++, verde), attivate (CD45-CD25+++, arancione) e secretorie (CD45-CD25++, viola). K) Le sottopopolazioni resting, attivate e secretorie sono state analizzate in base all'espressione del marcatore di attivazione CD39. L) Le cellule CD4+ sono state misurate in base al livello di espressione dell'antigene CD25 e definite cellule altamente positive per l'antigene CD25 (CD25hi). Questa popolazione include le cellule CD4+Tregs (rosa) e altre sottotipi CD4+. M) Le cellule CD4+Tregs (rosa) sono state sottratte dal numero totale di cellule CD4+CD25hi. Le popolazioni risultanti sono state così suddivise, in base all'espressione del CD45RA, in cellule CD45RA+CD25hiCD4+ non Treg e CD45RA-CD25hiCD4+ non Tregs, rispettivamente. P) Le cellule CD8br sono state discriminate in base all'espressione dell'antigene CD28. La stessa analisi è stata utilizzata sui sottotipi cellulari CD8br e CD8dim considerati uniti (cellule CD8+) (non mostrate in questa figura). Q-R) Le cellule CD8+ sono state analizzate in base all'espressione degli antigeni CD39 e CD45RA.

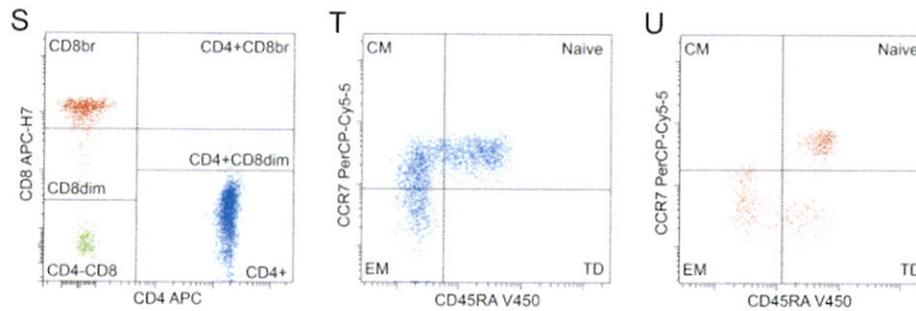


Figura 5. *Pannello maturazione T cellulare.*

Questa figura illustra la strategia di analisi adottata per il pannello della maturazione T cellulare su un campione di sangue rappresentativo. I sottotipi di cellule T descritte nel plot S sono state analizzate considerando i livelli di espressione degli antigeni CD45RA (asse delle X) e CCR7 (asse delle Y). Per semplicità, sono state mostrate solo le cellule CD4+ (plot T) e le cellule CD8br (plot U). Dividendo il plot T e U in 4 quadranti possiamo individuare: le cellule naive nel quadrante in alto a destra (CCR7+CD45RA-), le cellule della memoria centrale (CM) nel quadrante in basso a sinistra (CCR7-CD45RA-), le cellule della memoria effettrice (EM) nel quadrante in basso a sinistra (CCR7-CD45RA-) ed infine le cellule “terminally differentiated” nel quadrante in basso a destra (CCR7-CD45RA+).

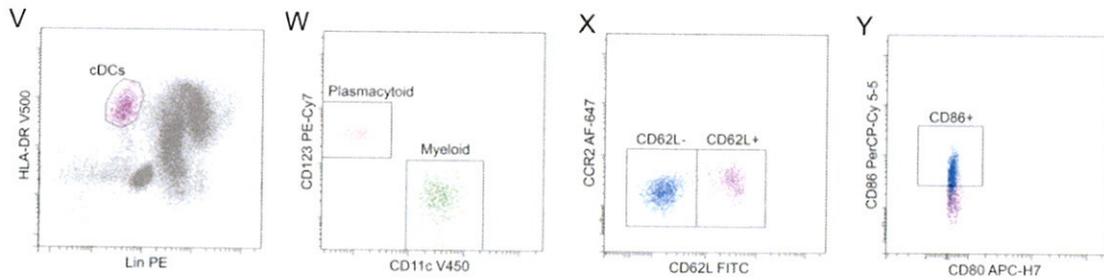


Figura 6. *Pannello cellule dendritiche circolanti.*

Questa figura illustra la strategia di analisi adottata per il pannello delle cellule dendritiche circolanti su un campione di sangue rappresentativo. V) Le cellule cDCs sono state discriminate in base alla loro negatività per il cocktail Lin 1 e per la loro positività per l'antigene HLA-DR. W) Le cDCs sono state suddivise in mieloidi (verde), positive per l'antigene CD11c e in plasmacitoidi (rosa chiaro), positive per l'antigene CD123. X) Le cDCs sono state analizzate per la molecola di adesione cellulare CD62L e per il recettore per le chemochine CCR2. La stessa strategia di analisi è stata applicata sui sottotipi di cellule mieloidi e plasmacitoidi (plots non mostrati). Y) Le cDCs sono state infine analizzate per i marcatori CD86 e CD80.

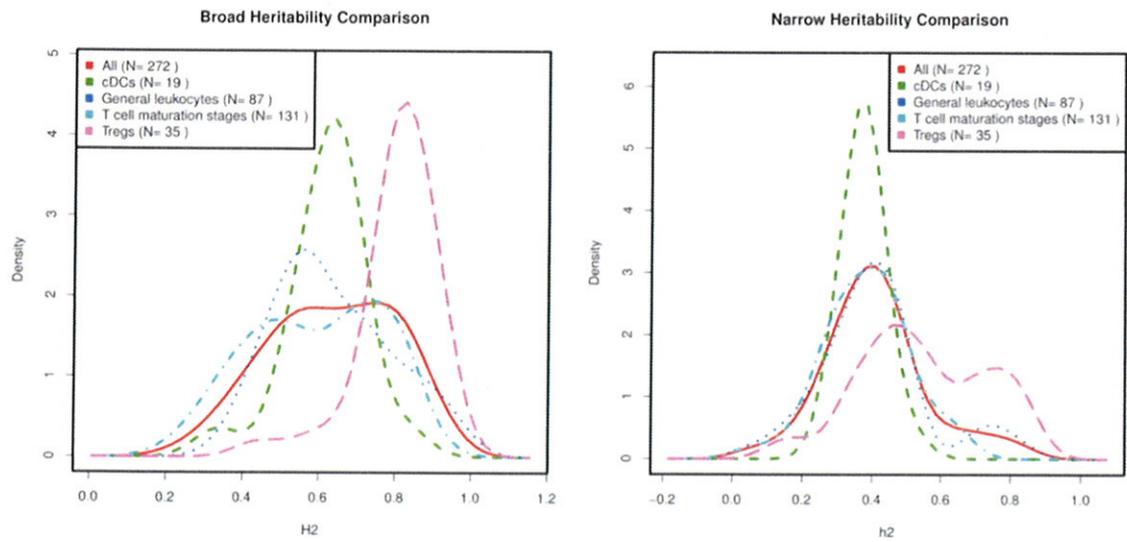


Figura 7. *Ereditabilità dei tratti immunofenotipici.*

I due istogrammi mostrano la distribuzione dell'ereditabilità ampia ("broad") e stretta ("narrow") per ciascun tratto quantitativo e per ciascuna delle 4 categorie (cDCs, leucociti generali, stadi della maturazione T cellulare, Tregs) in cui i tratti sono stati raggruppati.

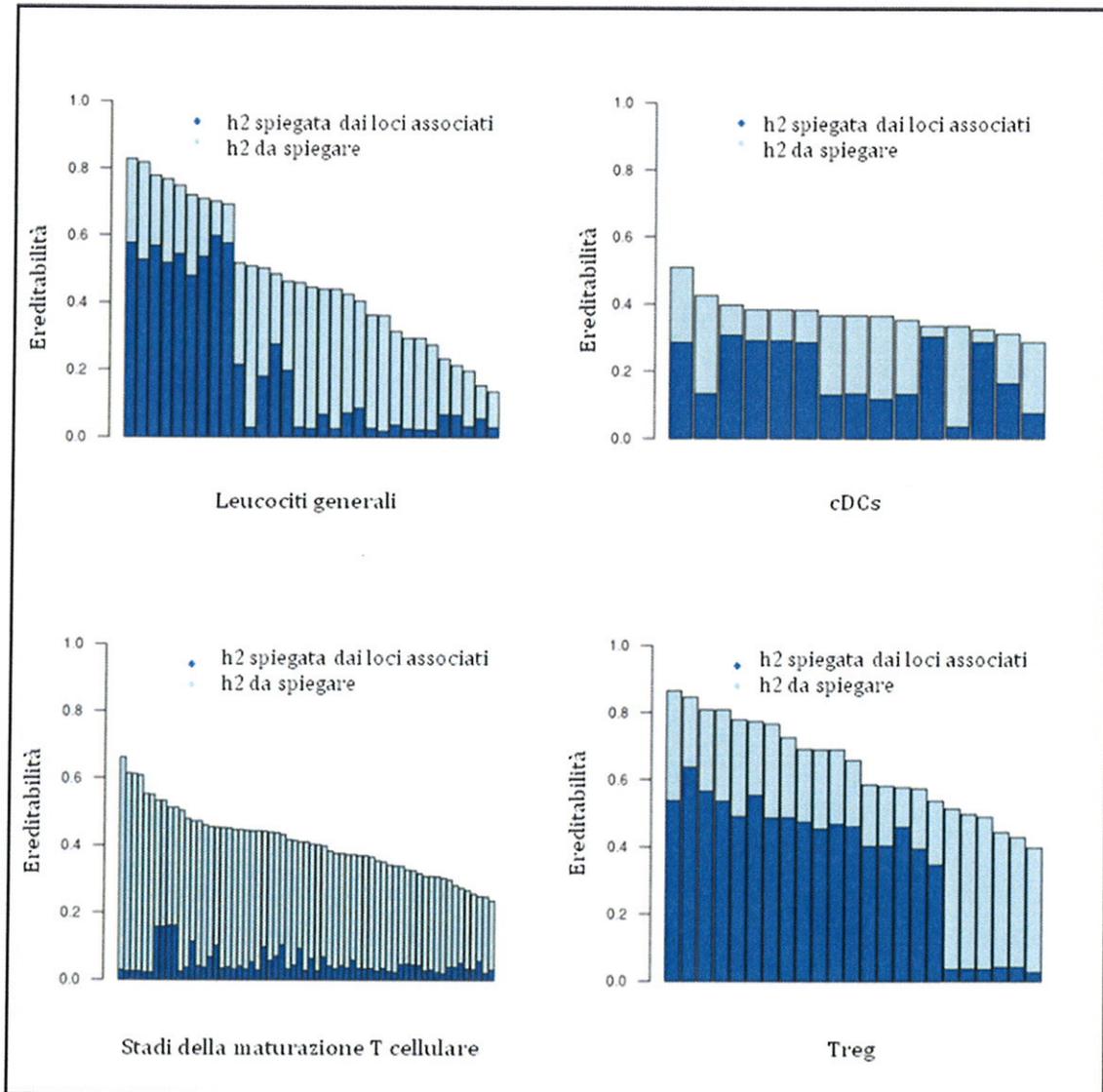


Figura 8. *Ereditabilità spiegata.*

I grafici a barre mostrano l'ereditabilità (h^2) di ciascun tratto, e distinguono la porzione dell'ereditabilità spiegata dai loci identificati (blu scuro), e la porzione dell'ereditabilità che rimane da spiegare (blu chiaro). Le barre rappresentano i 272 tratti immunofenotipici nella loro corrispondente categoria biologica.

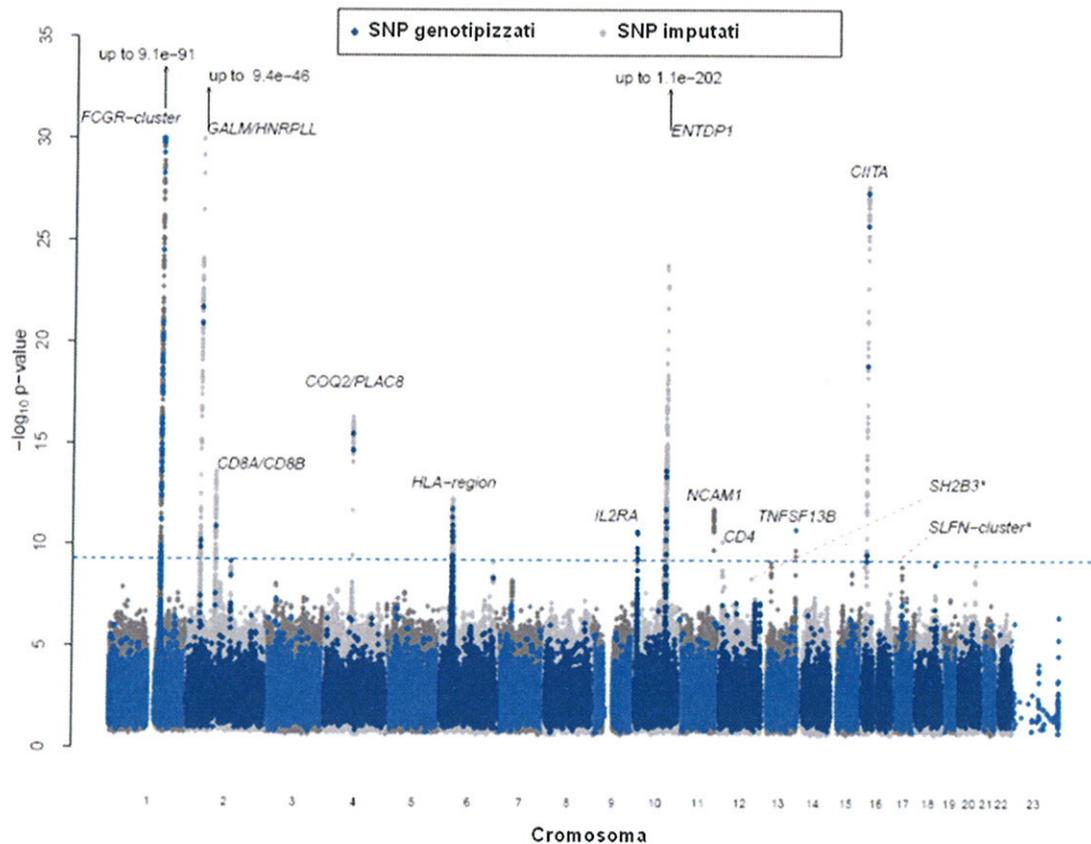
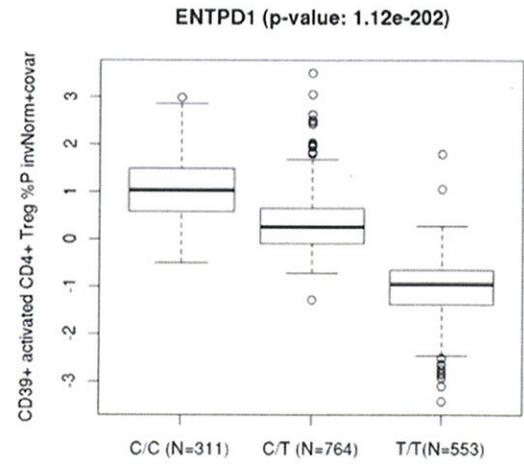
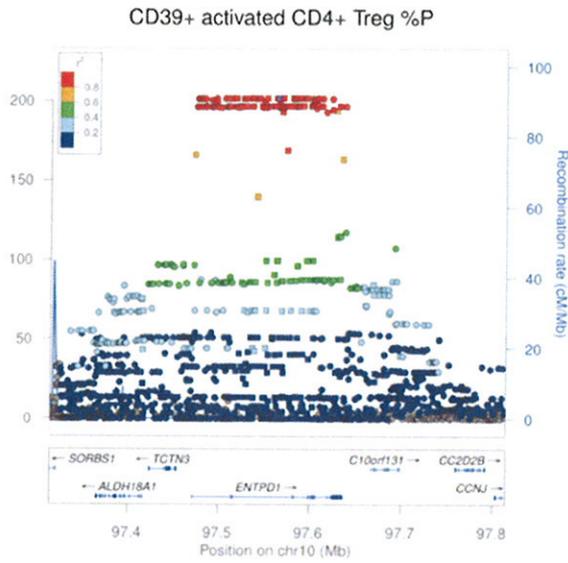


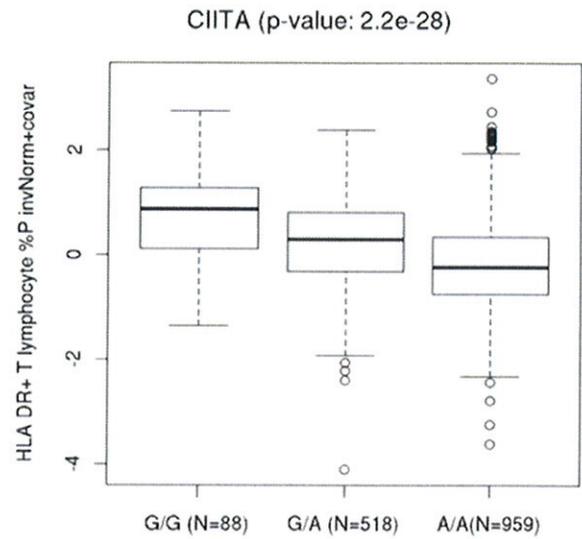
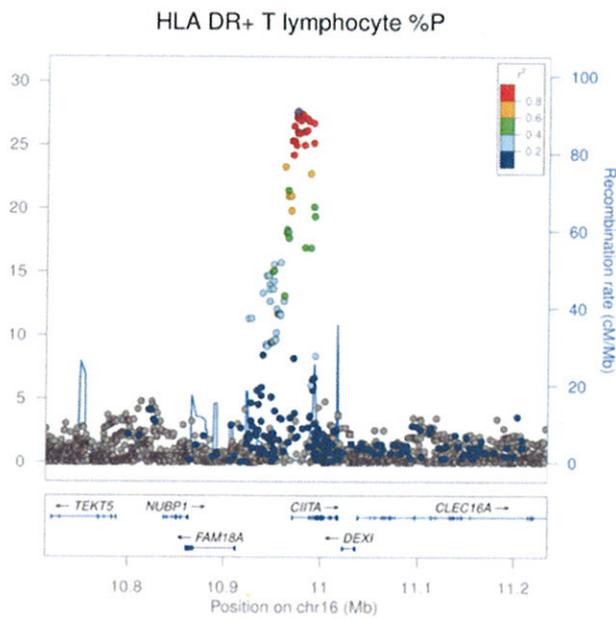
Figura 9. Manhattan plot.

Rappresentazione grafica delle associazioni genome-wide ottenute nello studio. Nell'asse delle X sono indicati i cromosomi, mentre nell'asse delle Y la significatività statistica espressa in $-\log_{10}$ del p-value. Per ogni SNP è stato indicato il miglior p-value osservato fra tutti i tratti. In blu sono rappresentati gli SNP genotipizzati, mentre in grigio sono indicati quelli imputati mediante sequenze sarde. La linea tratteggiata orizzontale rappresenta la soglia di significatività statistica utilizzata (5.26×10^{-10}). Il miglior gene candidato è indicato vicino al picco di associazione corrispondente. I loci sotto la soglia di significatività statistica e precedentemente descritti in letteratura, sono contrassegnati da un asterisco.

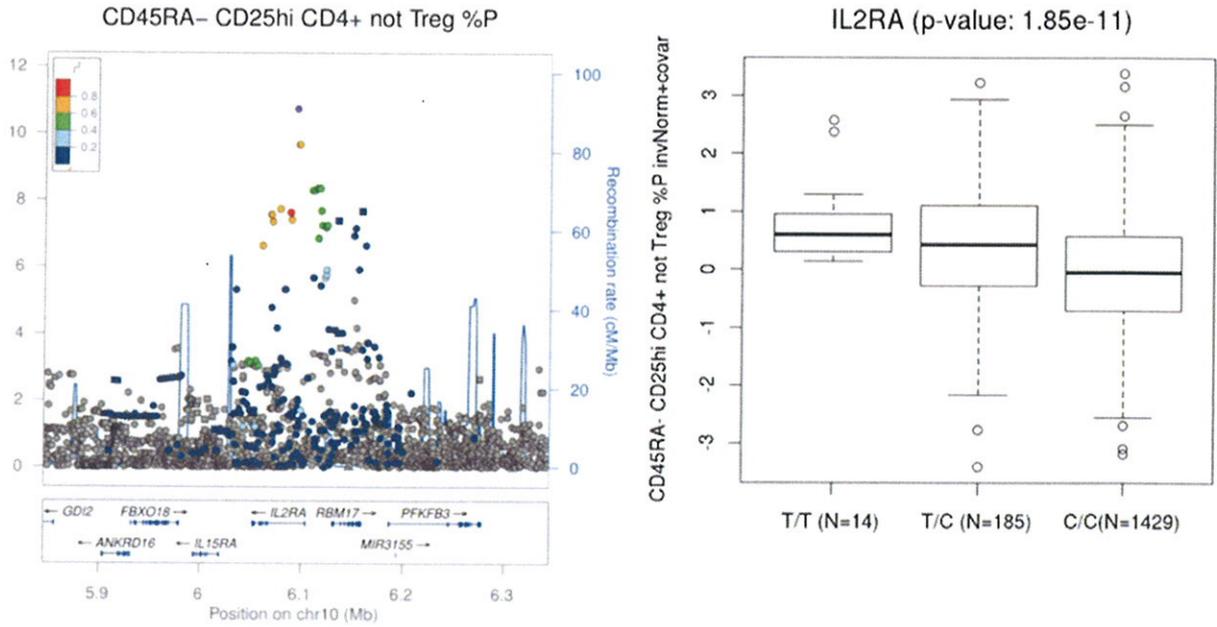
a) Regione del gene *ENTPD1*



b) Regione del gene *CIITA*



c) Regione del gene *IL2RA*



d) Regione dei geni *ATNX2/SH2B3*

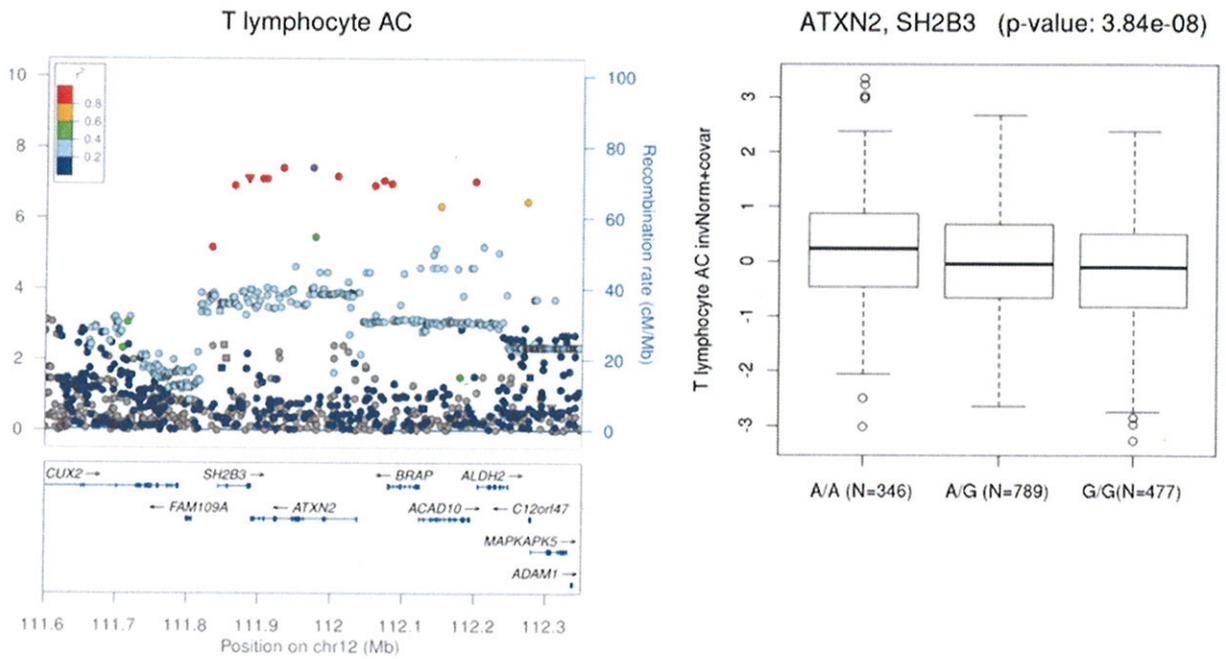


Figura 10. Plot regionali e boxplot per alcune varianti associate a tratti immunofenotipici.

Per ciascuna regione di associazione considerata, il pannello a sinistra (plot regionale) rappresenta la robustezza dell'associazione. L'asse delle X mostra la regione genica (onhg19/GRCh37 genomic build) in cui è localizzato lo SNP più significativo, il quale è indicato con un cerchio viola, mentre l'asse delle Y mostra il livello di significatività statistica (p-value). Gli altri SNPs presenti nella stessa regione e distinti da differenti colori riflettono il loro LD con il top SNP. I geni e la direzione della trascrizione, sono mostrati nel riquadro più in basso. I plot sono stati prodotti utilizzando la versione standalone del pacchetto LocusZoom [Pruim *et al.*, 2010]. Il pannello a destra, Box-plot, mostra la distribuzione dei livelli immunofenotipici all'interno di ciascuna classe genotipica ottenuta tenendo conto dei tratti normalizzati corretti per età ed genere. I box plot mostrano chiaramente l'effetto additivo osservato.

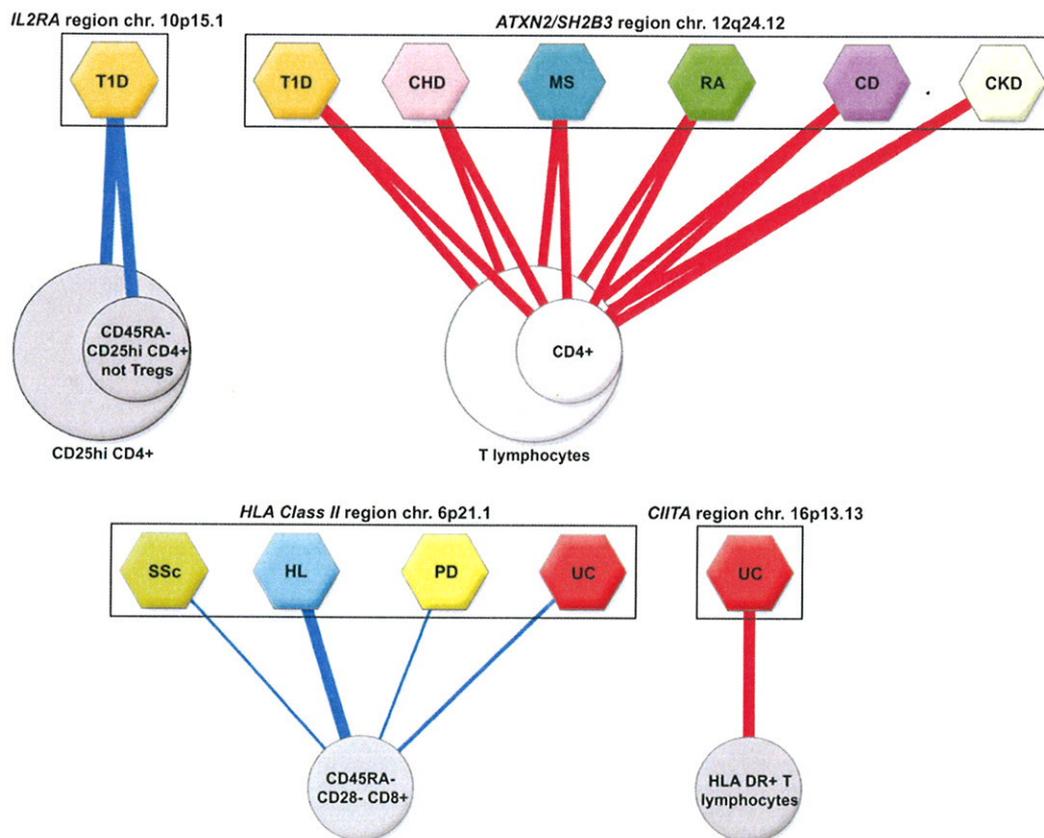


Figura 11. *Associazioni coincidenti con le malattie.*

Questa figura rappresenta in maniera schematica le associazioni coincidenti tra le cellule immuni e le malattie. I cerchi rappresentano le popolazioni cellulari ed i sottotipi cellulari sono mostrati come cerchi interni; la gradazione di grigio dei cerchi è proporzionale all'effetto relativo dell'associazione tra lo SNP e il tratto immunofenotipico (la tonalità più scura corrisponde ad un effetto più forte). Le malattie sono rappresentate da esagoni, il nome all'interno dell'esagono è un acronimo della malattia ed ogni colore rappresenta una patologia differente. Lo spessore di ciascun collegamento è proporzionale all' r^2 tra lo SNP associato con le malattie e quello associato con il tratto immunofenotipico; il colore indica se l'allele aumenta (rosso) o diminuisce (blu) il rischio della malattia. Gli acronimi relativi alle malattie sono i seguenti (in ordine alfabetico): CD= celiac disease (malattia celiaca), CKD= Chronic Kidney Disease (malattia renale cronica), CHD=Chronic Heart Disease (malattia cardiaca cronica), HL= Hodgkin's Lymphoma (linfoma di Hodgkin's), MS=Multiple Sclerosis (sclerosi multipla), PD= Parkinson's Disease

(malattia di parkinson), RA= Rheumatoid Arthritis (artrite reumatoide), SSc= Systemic Sclerosis (sclerosi sistemica), T1D= Type 1 Diabetes (diabete di tipo 1) e UC= Ulcerative Disease (colite ulcerosa).

Referenze Bibliografiche

Albanesi C., Scarponi C., Bosisio D., Sozzani S., Girolomoni G. Immune functions and recruitment of plasmacytoid dendritic cells in psoriasis. *Autoimmunity* (2010).

Auer P.L., Johnsen J.M., Johnson A.D., Logsdon B.A., Lange L.A., Nalls M.A., Zhang G., Franceschini N., Fox K., Lange E.M. et al. Imputation of exome sequence variants into population-based samples and blood-cell-trait-associated loci in African Americans. NHLBI GO Exome Sequencing Project. *Am J Hum Genet* (2012).

Banchereau J. and Virginia P. Type I Interferon in Systemic Lupus Erythematosus and Other Autoimmune Diseases. *Immunity* (2006).

Barrett J.C., Lee J.C., Lees C.W., Prescott N.J., Anderson C.A., Phillips A., Wesley E., Parnell K., Zhang, H., Drummond H., et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat Genet* (2009).

Battaglia M. and Roncarolo M.G. The Fate of Human Treg Cells. *Immunity* (2009).

Bayes B., Pastor M.C., Lauzurica R., Granada, M.L., Salinas I., and Romero R. Do anti-CD25 monoclonal antibodies potentiate posttransplant diabetes mellitus? *Transplant Proc* (2007).

Bluestone J.A., Mechanisms of tolerance. *Immunological Reviews* (2011).

Bonasio R., M Lucila S., Patrick S., Nir G., Andrew H.L. & Ulrich H. Clonal deletion of thymocytes by circulating dendritic cells homing to the thymus. *Nature Immunology* (2006).

Borsellino G., Kleinewietfeld M., Di Mitri D., Sternjak A., Diamantini A., Giometto R., Hopner S., Centonze D., Bernardi G., Dell'Acqua M.L., et al. Expression of ectonucleotidase CD39 by Foxp3⁺ Treg cells: hydrolysis of extracellular ATP and immune suppression. *Blood* (2007).

Buchholz V.R., Neuenhahn M., and Busch D.H. CD8⁺ T cell differentiation in the aging immune system: until the last clone standing. *Curr Opin Immunol* (2011).

Carrasco J., Daniele G., Aline V.P., Thierry B., and Pierre V.B. CD45RA on human CD8 T cells is sensitive to the time elapsed since the last antigenic stimulation. *Blood* (2006).

Caux C., Massacrier C., Vanberliet B., Dubois B., Van Kooten C., Durand I. and Banchereau J. Activation of human dendritic cells through CD40 cross-linking. *JEM* (1994).

Chalmin F., Gregoire M., Melanie B., Angelique C., Frederique V., Aziz H., Sylvain L., Valentin D., Julie V., David M., Simon C.R., Gerard E., Jean Rene' P., Christophe B., Bernhard R., Lionel A., Cedric R., and Francois G. Stat3 and Gfi-1 Transcription Factors Control Th17 Cell Immunosuppressive Activity via the Regulation of Ectonucleotidase Expression. *Immunity* (2012).

Charles J., Laurence C., Dimitri S., Jérémy D.D., Caroline A., Rémy G., Marie-Christine J., Marie-Jeanne R., Jean-Claude B., Joel P., and Marie-Thérèse L. Plasmacytoid dendritic cells and dermatological disorders: focus on their role in autoimmunity and cancer. *Eur J Dermatol* (2010).

Chatenoud, L. Immune therapy for type 1 diabetes mellitus-what is unique about anti-CD3 antibodies? *Nat Rev Endocrinol* (2010).

Chen L., *et al.*, Natural killer dendritic cells are an intermediate of developing dendritic cells. *J. Leukoc Biol* (2007).

Chen W.M., and Abecasis G.R. Family-based association tests for genome-wide association scans. *Am J Hum Genet.* (2007).

Cortes A., and Brown M.A. Promise and pitfalls of the Immunochip. *Arthritis Res Ther* (2011).

Davis M.M.. A prescription for human immunology. *Immunity* (2008).

Dendrou C.A., Plagnol V., Fung E., Yang J.H., Downes K., Cooper J.D., Nutland S., Coleman G., Himsworth M., Hardy M., *et al.* Cell-specific protein phenotypes for the autoimmune locus IL2RA using a genotype-selectable human bioresource. *Nat Genet* (2009).

Devalliere J., and Charreau B. The adaptor Lnk (SH2B3): an emerging regulator in vascular cells and a link between immune and inflammatory signaling. *Biochem Pharmacol* (2011).

Enciso-Mora V., Broderick P., Ma Y., Jarrett R.F., Hjalgrim H., Hemminki K., Van den Berg A., Olver B., Lloyd A., Dobbins S.E., *et al.* A genome-wide association study of Hodgkin's lymphoma identifies new susceptibility loci at 2p16.1 (REL), 8q24.21 and 10p14 (GATA3). *Nat Genet* (2010).

Ferreira M.A., Mangino M., Brumme C.J., Zhao, Z.Z., Medland S.E., Wright M.J., Nyholt D.R., Gordon S., Campbell M., McEvoy B.P., *et al.* Quantitative trait loci for CD4:CD8 lymphocyte ratio are associated with risk of type 1 diabetes and HIV-1 immune control. *Am J Hum Genet* (2010).

Fletcher J.M., Loneragan R., Costelloe L., Kinsella K., Moran B., O'Farrelly C., Tubridy N., Mills K.H. CD39⁺Foxp3⁺ regulatory T Cells suppress pathogenic Th17 cells and are impaired in multiple sclerosis. *J Immunol* (2009).

Fontenot J.D., Marc A.G. and Alexander Y. R. Foxp3 programs the development and function of CD4⁺CD25⁺ regulatory T cells. *Nature Immunology* (2003).

Gorlova O., Martin J.E., Rueda B., Koeleman B.P., Ying J., Teruel M., Diaz-Gallo L.M., Broen J.C., Vonk M.C., Simeon C.P., *et al.* Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome-wide association strategy. *PLoS Genet* (2011).

Hamza T.H., Zabetian C.P., Tenesa A., Laederach A., Montimurro J., Yearout D., Kay D.M., Doheny K.F., Paschall J., Pugh, E., *et al.* Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat Genet* (2011).

Hindorff L.A., M.J.E.B.I., Wise A., Junkins H.A., Hall P.N., Klemm A.K., and Manolio T.A. A Catalog of Published Genome-Wide Association Studies., Available at: www.genome.gov/gwastudies. Accessed [06/11/2012].

<http://www.immunobase.org/>

Hori S., Nomura T. & Sacaguchi S. Control of regulatory T cell development by the transcription factor Foxp3. *Science* (2003).

Howie B., Fuchsberger C., Stephens M., Marchini J., and Abecasis G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics* (2012).

- Huang J., Ellinghaus D., Franke A., Howie B., and Li Y. 1000 Genomes-based imputation identifies novel and refined associations for the Wellcome Trust Case Control Consortium phase I Data. *Eur J Hum Genet* (2012).
- Kristiansson K., Jussi N. and Leena P. Isolated populations and complex disease gene identification. *Genome Biology* (2008).
- Kreuzaler M., Rauch, M., Salzer U., Birmelin J., Rizzi M., Grimbacher B., Plebani A., Lougaris, V., Quinti I., Thon V., *et al.* Soluble BAFF levels inversely correlate with peripheral B cell numbers and the expression of BAFF receptors. *J Immunol* (2012).
- Lanzavecchia A. Antigen-specific interaction between T and B cells. *Nature* (1985).
- Li G., Ruan X., Auerbach R.K., Sandhu K.S., Zheng M., Wang, P., Poh H.M., Goh Y., Lim J., Zhang J., *et al.* Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* (2012).
- Li Y., Willer C., Sanna S., and Abecasis G. Genotype imputation. *Annu Rev Genomics Hum Genet* (2009).
- Lowe C.E., Cooper J.D., Brusko T., Walker N.M., Smyth D.J., Bailey R., Bourget K., Plagnol V., Field S., Atkinson M., *et al.* Large-scale genetic fine mapping and genotype-phenotype associations implicate polymorphism in the IL2RA region in type 1 diabetes. *Nat Genet* (2007).
- Maecker H.T., McCoy J.P. and Nussenblatt R. Standardizing immunophenotyping for the Human Immunology Project. *Nat Rev Immunol* (2012).
- Marwaha A.K., Crome S.Q., Panagiotopoulos C., Berg K.B., Qin H., Ouyang Q., Xu L., Priatel J.J., Levings M.K., Tan R. Cutting edge: Increased IL-17-secreting T cells in children with new onset type 1 diabetes. *J Immunol.* (2010).
- Marrosu M.G., Murru R., Murru M.R., Costa G., Zavattari P., Whalen M., Cocco E., Mancosu C., Schirru L., Solla E., *et al.* Dissection of the HLA association with multiple sclerosis in the founder isolated population of Sardinia. *Hum Mol Genet* (2001).
- McGovern D.P., Gardet A., Torkvist L., Goyette P., Essers J., Taylor K.D., Neale B.M., Ong R.T., Lagace C., Li C., *et al.* Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat Genet* (2010).

- Miyara M., Yoshioka Y., Kitoh A., Shima T., Wing K., Niwa A., Parizot C., Taflin C., Heike T., Valeyre D., *et al.* Functional delineation and differentiation dynamics of human CD4⁺ T cells expressing the FoxP3 transcription factor. *Immunity* (2009).
- Nalls M.A., Couper D.J., Tanaka T., Van Rooij F.J., Chen M.H., Smith A.V., Toniolo D., Zakai N.A., Yang Q., Greinacher A., *et al.* Multiple loci are associated with white blood cell phenotypes. *PLoS Genet* (2011).
- Negro R., Gobessi S., Longo P.G., He Y., Zhang Z.Y., Laurenti L., Efremov D.G.
- Overexpression of the autoimmunity-associated phosphatase PTPN22 promotes survival CCL cells by selectively activating AKT. *Blood* (2012).
- Ohnmacht C., Pullner A., King S.B., Drexler I., Meier S., Brocker T., and Voehringer D. Constitutive ablation of dendritic cells breaks self-tolerance of CD4 T cells and results in spontaneous fatal autoimmunity. *J Exp Med* (2009).
- Okada Y., Hirota T., Kamatani Y., Takahashi A., Ohmiya H., Kumasaka N., Higasa K., Yamaguchi-Kabata Y., Hosono N., Nalls M.A., *et al.* Identification of nine novel loci associated with white blood cell subtypes in a Japanese population. *PLoS Genet* (2011).
- Pilia G., Chen W.M., Scuteri A., Orru M., Albai G., Dei M., Lai S., Usala G., Lai M., Loi P., *et al.* Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* (2006).
- Pilling D., Arne N. Akbar, Paul A. Bacon and Mike S. CD4⁺ CD45RA⁺ T cells from adults respond to recall antigens after CD28 ligation. *International Immunology* (1996).
- Pulte E.D., Broekman M.J., Olson K.E., Drosopoulos J.H., Kizer J.R., Islam N., and Marcus A.J. CD39/NTPDase-1 activity and expression in normal leukocytes. *Thromb Res* (2007).
- Sallusto F., Danielle L., Reinhold F., Martin L. & Antonio L. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature* (1999).
- Sébastien A., David F.T. , Sébastien Lacroix-D., Srini V.K. and Jagadeesh B.. Surveillance of Antigen-Presenting Cells by CD4⁺CD25⁺ Regulatory T Cells in Autoimmunity. *Am J Pathol* (2009).

- Sun J.C. & Lanier L.L. NK cell development, homeostasis and function: parallels with CD8⁺ T cells. *Nat Rev Immunol* (2011).
- Sansoni P., Vescovini R., Fagnoni F., Biasini C., Zanni F., Zanlari, L., Telera A., Lucchini G., Passeri G., Monti D., *et al.* The immune system in extreme longevity. *Exp Gerontol* (2008).
- Schiemann B., Gommerman J.L., Vora K., Cachero T.G., Shulga-Morskaya S., Dobles M., Frew E., and Scott M.L. An essential role for BAFF in the normal development of B cells through a BCMA-independent pathway. *Science* (2001).
- Takada K., and Jameson S.C. Self-class I MHC molecules support survival of naive CD8 T cells, but depress their functional sensitivity through regulation of CD8 expression levels. *J Exp Med* (2009).
- Vendrame, F., Pileggi A., Laughlin E., Allende G., Martin-Pagola A., Molano R.D., Diamantopoulos S., Standifer N., Geubtner K., Falk B.A., *et al.* Recurrence of type 1 diabetes after simultaneous pancreas-kidney transplantation, despite immunosuppression, is associated with autoantibodies and pathogenic autoreactive CD4 T-cells. *Diabetes* (2010).
- Voight B.F., Kang H.M., Ding J., Palmer C.D., Sidore C., Chines P.S., Burt N.P., Fuchsberger C., Li, Y., Erdmann J., *et al.* The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet* (2012).
- Wu Z., Yates A.L., Hoyne G.F., and Goodnow C.C. Consequences of increased CD45RA and RC isoforms for TCR signaling and peripheral T cell deficiency resulting from heterogeneous nuclear ribonucleoprotein L-like mutation. *J Immunol* (2010).
- Yang G.X., Zhe-Xiong L., Kentaro K., Yuki M., Aftab A.A., Yong-Jun L., Susumu I. and M. Eric G. Plasmacytoid Dendritic Cells of Different Origins Have Distinct Characteristics and Function: Studies of Lymphoid Progenitors versus Myeloid Progenitors. *J Immunology* (2005).
- Zhernakova A., Elbers C.C., Ferwerda B., Romanos J., Trynka G., Dubois P.C., De Kovel C.G., Franke L., Oosting M., Barisani D., *et al.* Evolutionary and functional analysis of celiac risk loci reveals SH2B3 as a protective factor against bacterial infection. *Am J Hum Genet* (2010).

Ringraziamenti

Ringrazio tutti i volontari del progetto ProgeNIA che hanno partecipato allo studio con grande entusiasmo e pazienza.

Il progetto descritto in questa tesi ed i risultati emersi sono il frutto del lavoro di tante persone che con impegno e dedizione hanno contribuito alla sua realizzazione, quindi vorrei rivolgere un ringraziamento sentito a tutte loro.

In particolare, vorrei esprimere la mia sincera gratitudine al Prof. Francesco Cucca per avermi dato la possibilità di “vivere” questo meraviglioso viaggio raggiungendo questo traguardo, per la fiducia ed i consigli, e per le opportunità che mi ha offerto in questi anni. Gli sono grata per avermi dato la possibilità di conoscere e far parte della realtà del “progetto ProgeNIA” e per avermi sempre incoraggiato ad essere propositiva e positiva nei confronti di questo affascinante e complesso “mondo della ricerca”.

Un ringraziamento che nasce dal cuore è rivolto ai miei Tutori, la Dott.ssa Valeria Orrù ed il Dott. Edoardo Fiorillo. Senza di loro questo percorso di crescita professionale ma anche di vita non sarebbe stato così ‘carico di emozioni’. In questi anni mi hanno seguito con costanza, grande serietà e professionalità. Sono davvero grata a loro per la disponibilità, i preziosi consigli e la loro bontà d’animo. Desidero ringraziarli anche per il grande aiuto fornitomi durante la stesura della tesi.

Rivolgo un ringraziamento speciale a tutte le persone con cui ho lavorato per portare avanti questo progetto: i miei tutori Valeria Orrù ed Edoardo Fiorillo; i colleghi di laboratorio Mariano Dei, Sandra Lai, Antonella Mulas, Magdalena Zoledziewska, Monia Lobina, Maria Grazia Piras, Michele Marongiu e l'intero staff medico e paramedico ProgeNIA. Un pensiero speciale lo voglio rivolgere in particolare a Monia, Michele, Maria Grazia e Valentina. Grazie per le splendide giornate, le grasse risate ed il supporto morale.

Ringrazio inoltre Serena Sanna, Maristella Steri e Gabriella Sole che hanno effettuato le analisi statistiche sui tratti immunofenotipici. In particolare, ringrazio Maristella Steri per la revisione critica della tesi.

Un ringraziamento a tutti gli altri autori del lavoro: Carlo Sidore, Fabio Busonero, Matteo Floris, Wiesława I. Mentzen, Silvana A. M. Urru, Stefania Olla, Andrea Maschio, Maristella Pitzalis, Maria Francesca Urru, Marco Marcelli, Roberto Cusano, Francesca Deidda, Manuela Oppo, Rosella Pihu, Frederic Reinier, Riccardo Berutti, Luca Pireddu, Ilenia Zara, Eleonora Porcu, Alan Kwong, Christine Brennan, Brendan Tarrier, Robert Lyons, Hyun M. Kang, Sergio Uzzau, Rossano

Atzeni, Maria Valentini, Davide Firinu, Lidia Leoni, Gianluca Rotta, Silvia Naitza, Andrea Angius, Mauro Congia, Michael B. Whalen, Chris M. Jones, David Schlessinger, Gonçalo R. Abecasis.

Ci tengo a ringraziare tutte le persone con cui ho condiviso i primi anni di lavoro a Cagliari ed i primi mesi di Dottorato. Ricorderò sempre con affetto i bei giorni trascorsi insieme. In particolare: Silvia Naitza che mi ha permesso di far parte del team di ricerca del “Laboratorio 2” e di muovere i primi passi nel campo della ricerca ed i colleghi Sonia Sanna, Francesca Anedda, Alessia Loi, Christian Caria, Nazario Olla, Gabriella Sole, Elen Balaci, Elisa Lobina (ormai assente da anni dal gruppo ma che ricordo sempre con piacere), Alessandro Puddu e Roberta Piras.

Voglio ricordare i colleghi Andrea Maschio ed Eleonora Porcu per essermi stati vicino anche se “a distanza”, grazie per le parole di incoraggiamento e di sostegno morale.

Un dolce e sentito pensiero è rivolto al prof. Antonio Cao, scomparso nel novembre 2012, per avermi permesso di iniziare “la mia prima vera esperienza lavorativa” presso l’Istituto di Ricerca Genetica e Biomedica del Consiglio Nazionale delle Ricerche di Cagliari.

Ed infine un grazie a Giuseppe Pilia che, con brillante intuito e grandi capacità scientifiche, è riuscito a ideare e concretizzare il progetto ProgeNIA.

“L’ultimo ringraziamento, non di certo per importanza, va alla mia famiglia e a Mattia, per essere stati costantemente al mio fianco, per aver sempre creduto in me e per aver condiviso questi importanti e bellissimi anni della mia vita appoggiando senza riserve tutte le mie scelte”.