



A.D. MDLXII

**Università degli Studi di Sassari**

**Corso di dottorato in Biologia, Biochimica e Biotecnologie Molecolari**

**Ciclo XX**

**Genetica della sclerosi multipla e del diabete di tipo 1:  
regione HLA di classe III**

**Tutor:  
dr.ssa Laura Morelli**

**Coordinatore:  
prof. Bruno Masala**

**dottoranda:  
dr.ssa Lucia Gemma Delogu**



---

----- **Indice**

---

**INDICE**

<b>Capitolo 1 – Introduzione.....</b>	<b>1</b>
a) Sclerosi multipla: caratteristiche generali.....	1
a.1) ereditarietà e basi genetiche.....	7
b) Diabete di tipo 1: caratteristiche generali.....	10
b.1) Ereditarietà e basi genetiche.....	14
c) HLA.....	17
c.1) TNF.....	20
c.2) LTA.....	22
 <b>Capitolo 2 - Scopo della ricerca.....</b>	 <b>25</b>
 <b>Capitolo 3 - Presupposti e disegno sperimentale.....</b>	 <b>27</b>
a) Sclerosi multipla: sequenziamento e tipizzazione.....	30
b) Diabete di tipo1: sequenziamento, ricerca in silico e tipizzazione.....	31
 <b>Capitolo 4 - Materiali e metodi.....</b>	 <b>32</b>
a) Selezione del campione.....	32
b) Estrazione del DNA.....	36
c) Estrazione dell' RNA.....	37
d) Retrotrascrizione.....	38
e) Sequenziamento.....	38
f) Digestione enzimatica.....	41
g) ARMS-PCR.....	42
h) Selezione SNPs di interesse: calcolo dello score di rilevanza.....	43
i) Analisi taqman: discriminazione all'elica.....	45
j) Analisi statistica.....	48
 <b>Capitolo 5 - Risultati e discussione.....</b>	 <b>49</b>
a) Sequenziamento.....	49
b) Allineamento sequenze e valutazione dello score.....	52
c) TDT – transmission disequilibrium test.....	56
d) Conclusioni.....	57
 <b>Riassunto.....</b>	 <b>59</b>
<b>Bibliografia.....</b>	<b>62</b>
<b>Sitografia.....</b>	<b>74</b>



## **Capitolo 1 - Introduzione**

### **a) Sclerosi multipla: caratteristiche generali**

La sclerosi multipla (SM) è una malattia autoimmune, infiammatoria e demielinizzante che colpisce il sistema nervoso centrale (SNC). La demielinizzazione consiste nella scomparsa della mielina, sostanza costituita da acidi grassi che riveste gli assoni delle fibre nervose. Nel SNC, la mielina viene prodotta dagli oligodendrociti e risulta costituita da lamelle lipidiche sovrapposte che formano manicotti interrotti da nodi, i nodi di Ranvier (Fig.1). La mielina è fondamentale nella trasmissione degli impulsi nervosi poiché, attraverso un meccanismo di tipo “saltatorio” da un nodo di Ranvier all’altro, rende la conduzione dello stimolo lungo la fibra nervosa più rapida ed efficace rispetto alle fibre non mielinizzate.

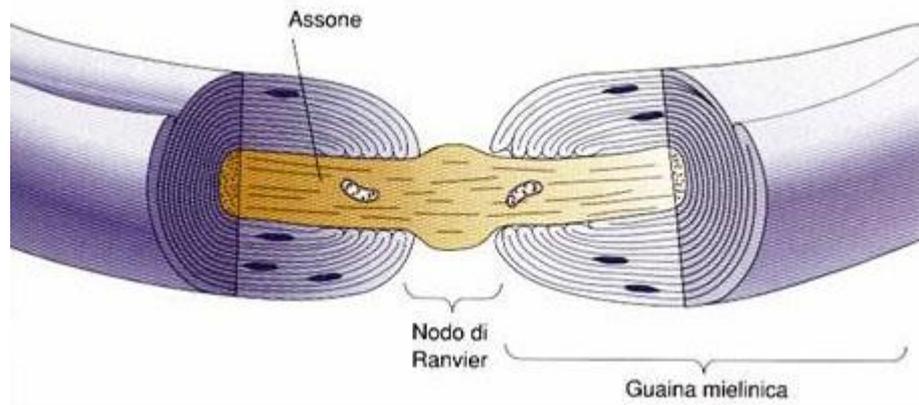


Fig.1 - Assone e guaina mielinica

La distruzione delle guaine mieliniche causa il blocco o il rallentamento della normale conduzione degli impulsi nervosi comportando il manifestarsi di una estrema varietà di sintomi propri di questa malattia. Le zone in cui la mielina è stata danneggiata o distrutta vengono anche dette “placche” e tali aree nel tempo vanno incontro ad un processo di cicatrizzazione. Da qui nasce il nome sclerosi multipla: “sclerosi” per la presenza di lesioni cicatrizzate (le placche), “multipla” per il fatto che le lesioni possono interessare varie zone del sistema nervoso centrale. I sintomi interessano le diverse funzioni dell'organismo regolate dal sistema nervoso centrale, ovvero il movimento (facile affaticamento, paralisi di uno o più arti, disturbi della coordinazione), le sensibilità (sensazione di anestesia, formicolio), la vista (annebbiamento, visione doppia), l'equilibrio, la parola, la coordinazione, le funzioni sfinteriche. Le manifestazioni della malattia variano a seconda delle aree dell'encefalo e del midollo spinale colpite. Si tratta di una patologia cronica e dal decorso variabile. Studi

epidemiologici ed esperienza derivante dalla pratica clinica hanno permesso di individuare diversi modelli a seconda dell'andamento della malattia. Il più comune è quello recidivante/remittente che colpisce il 90% delle persone ed è caratterizzato da una sequenza di ricadute e miglioramenti. Il restante 10% contrae una *forma progressiva* dal momento in cui appaiono i primi sintomi. All'interno di questi modelli abbastanza generali, si possono riconoscere quattro forme di sclerosi multipla.

*Benigna* - interessa circa il 10% dei malati ed è caratterizzata da uno o due attacchi che non lasciano segni di invalidità. In questo caso le condizioni si stabilizzano per lunghi periodi di tempo, arrivando a superare i 10 anni.

*Recidivante-remittente* - riguarda il 40% dei malati i quali, dopo ogni ricaduta mostrano un certo recupero. Il susseguirsi delle recidive, variando per gravità e durata, può aggravare i sintomi fino a lasciare segni d'invalidità. Queste ricadute possono verificarsi senza apparenti motivazioni, riconducibili comunque a stress, infezioni e traumi.

*Progressiva secondaria* - il 40% dei malati, pari alla metà delle persone con sclerosi multipla recidivante-remittente, ad un certo punto smettono di entrare nella fase di miglioramento e le loro condizioni iniziano a deteriorarsi costantemente.

*Primaria Progressiva* - Colpisce il 10% dei malati. Le persone affette da questa forma mostrano un costante peggioramento

senza fasi di remissione che li porta, in un brevissimo lasso di tempo, a completa invalidità.

Più dati fanno riflettere sulla necessaria assistenza di cui hanno bisogno i malati di SM. Il 50% non è più in grado di camminare autonomamente dopo 10 anni di malattia. Inoltre tutti i sintomi più diffusi sono altamente invalidanti. Per esempio la spossatezza è presente nell'89% dei casi, i deficit sensoriali nell'87%, l'atassia nell'82%, i disturbi della vescica nel 71%. Secondo statistiche britanniche, il 15,4% dei pazienti non è in condizione di lavorare, poco meno del 18% ha bisogno dell'assistenza di terzi per almeno 3 ore al giorno e, soprattutto, più del 67 % non ha una vita sociale a meno che non ci siano terzi ad aiutarlo. Non meraviglia, dunque, che questa malattia abbia un alto impatto economico: in base ai dati forniti dall'AIMS, ogni anno il malato richiede cure e assistenza per 25.000 euro (Tassinari et al. 2001; Richards et al. 2002).

Per quanto concerne l'età di esordio della patologia, essa si presenta tra i 15 e i 50 anni; i sintomi caratterizzanti si manifestano più frequentemente tra i 20 e i 40 anni. Sono maggiormente colpite le donne rispetto agli uomini con un rapporto di 2 a 1. Negli Stati Uniti la SM colpisce circa 400.000 persone (Hauser and Goodkin 1998), in Italia ne colpisce circa 50.000. La SM è frequente tra le popolazioni caucasiche (soprattutto tra quelle residenti nel nord-ovest europeo), nel nord America, nel sud-est dell'Australia e in

Nuova Zelanda, sud-Africa e America meridionale. La malattia è rara in Asia e nelle regioni caraibiche. Diversi studi hanno dimostrato come la prevalenza della malattia cresca all'aumentare della distanza dall'equatore (Yu et al. 1989; Cook et al. 1998) (Fig.2).

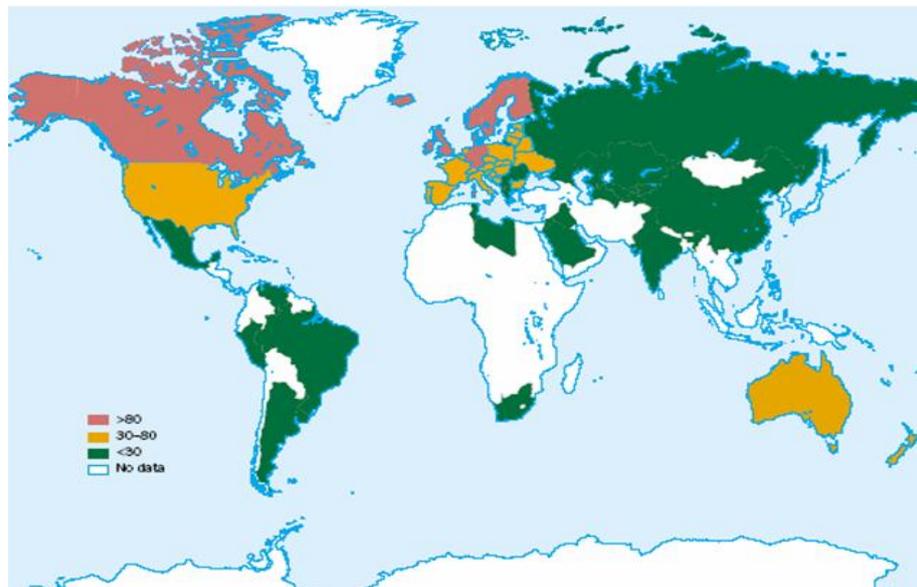


Fig. 2 - Prevalenza mondiale della sclerosi multipla

In Europa l'incidenza è alta nei paesi del nord soprattutto in Scandinavia e bassa nei paesi del sud, con una eccezione: la Sardegna (Fig.3). In Sardegna il rapporto di prevalenza è di 150 casi per 100.000 abitanti (Pugliatti et al. 2001). Questo dato è due volte superiore rispetto al resto della popolazione italiana e alla maggior parte delle popolazioni caucasiche; esso

è inoltre contrario al gradiente di prevalenza teorizzato in passato (Kurtzke, 1997).

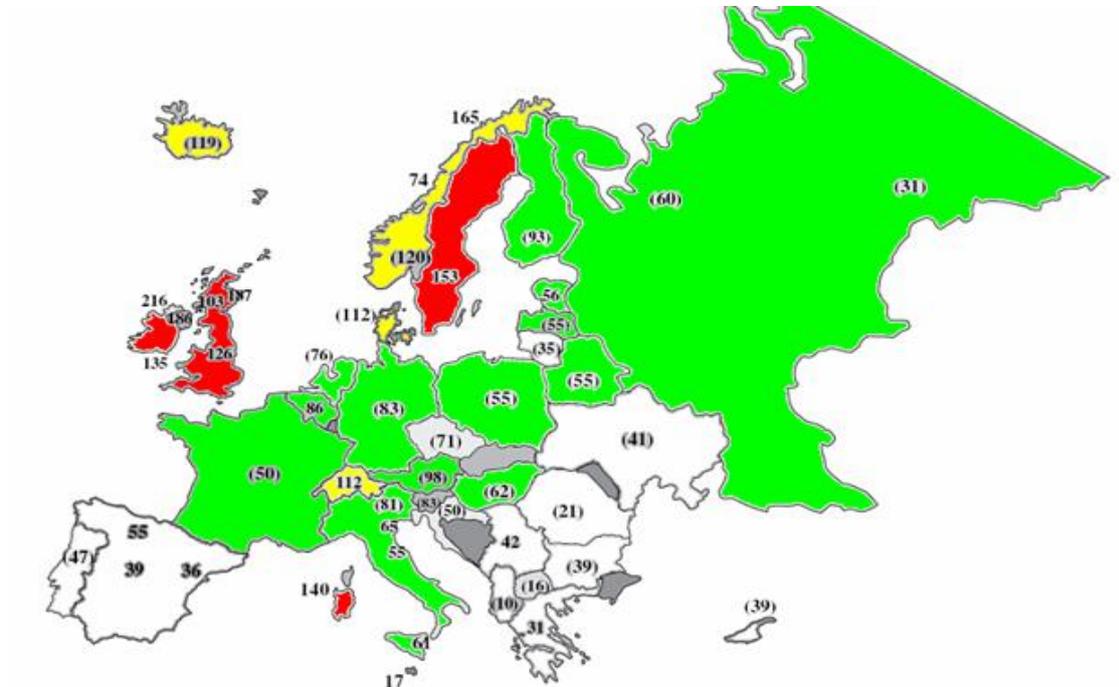


Fig.3 - Incidenza della sclerosi multipla in Europa

Occorre, comunque, interpretare con cautela i dati proposti a fondamento della teoria del gradiente di prevalenza e quindi tutte le considerazioni che in base ad esso possiamo trarre. La suddetta teoria venne recentemente messa in discussione da più ricercatori (22nd Congress of the European Committee for Treatment and Research in Multiple Sclerosis September 27-30, 2006). Sono molteplici i punti a sfavore della attendibilità della teoria:

- criteri diagnostici non omogenei in tutto il mondo; i criteri di Poser (Poser et al 1983) sono più restrittivi rispetto a quelli usati in precedenza;

- variabilità, nei diversi paesi, dell'accuratezza nell'accertamento dei casi;
- differenza nella probabilità di sopravvivenza a seconda della popolazione in esame;
- errati confronti tra popolazioni molto diverse tra loro per numero di abitanti.

Nonostante il crescente dibattito sulla distribuzione geografica della sclerosi multipla, è ormai assodato il fatto che la malattia sia fortemente presente in Sardegna (Pugliatti et al. 2002). Si registra inoltre un aumento dell'incidenza (Pugliatti et al. 2005) in questi ultimi anni anche se bisogna tenere conto dei passi avanti che sono stati fatti per quanto riguarda i criteri di diagnosi.

### **a.1) Ereditarietà e basi genetiche**

La sclerosi multipla è una malattia multifattoriale la cui insorgenza è dovuta alla presenza ed interazione di molteplici fattori sia genetici che ambientali. I numerosi studi effettuati sull'ereditarietà della patologia, condotti sui gemelli e sui fratelli di figli affetti, hanno messo in evidenza l'importanza del ruolo dei fattori genetici. L'aggregazione familiare nella SM è geneticamente determinata (Ebers et al. 1995; Mumford et al. 1994; Sadovnick et al. 1996).

I fratelli di individui affetti hanno un rischio di circa il 5% di ammalarsi nel corso della loro vita. L'evidenza dell'effetto

genetico deriva dagli studi sui gemelli effettuati in differenti popolazioni tra cui quella canadese e italiana. I dati ottenuti da queste ricerche variano leggermente a seconda della popolazione presa in esame: in generale essi mostrano un tasso di concordanza di circa il 30% nei monozigoti e di circa il 5% nei dizigoti (simile al rischio dei fratelli non gemelli) (Kinnun et al. 1986; Willer et al. 2003; Ristori et al. 2006). L'alto valore di concordanza nei gemelli e il basso rischio di ricorrenza familiare sottolineano il modello poligenico della malattia (Ebers 1988). Studi di linkage e associazione su differenti popolazioni hanno identificato il complesso maggiore di istocompatibilità MHC sul braccio corto del cromosoma 6 come uno dei determinanti genetici per la SM. I loci HLA-DR e -DQ posti in classe II giocano indubbiamente un ruolo centrale nella suscettibilità alla malattia. La precisa associazione a livello di popolazione con l'aplotipo DRB1\*1501, DQA1\*0102, DQB1\*0602 è stata ripetutamente dimostrata (Ligers et al. 2001, Barcellos et al. 2003, Oksenberg et al.2004). Uno studio effettuato su una regione di 11,4 Mb ha rilevato la presenza di cinque aplotipi DQB1-DRB1 positivamente associati con la SM nella popolazione Sarda (Marrosu et al. 2001).

Nonostante l'associazione con l'HLA sia stata ben determinata, il contributo che essa apporta alla suscettibilità è modesto e ciò riconduce all'ipotesi in base alla quale la SM è caratterizzata da una notevole eterogeneità genetica che va oltre la regione HLA (Dyment et al. 1997). Grazie a studi

*genome-wide* sono stati trovati numerosi loci associati con la patologia tra cui quelli sui cromosomi 1, 2, 3, 5, 10, 17; tuttavia tutti questi loci presentano un effetto modesto sulla suscettibilità (Sawcer et al. 1996, Ebers et al. 1996, Corradu et al. 2001, Dyment et al. 2001).

## **b) Diabete di tipo 1: caratteristiche generali**

Il diabete di tipo 1 (T1D) è una malattia cronica autoimmune causata da una combinazione di geni predisponenti, in presenza di fattori ambientali permissivi ancora sconosciuti.

La malattia è causata dalla distruzione irreversibile delle cellule  $\beta$  pancreatiche da parte di linfociti T autoreattivi, con conseguente deficiente sintesi dell'ormone insulina e permanente dipendenza dalla somministrazione dell'insulina esogena. La malattia rappresenta la forma di diabete clinicamente più grave e ad insorgenza più precoce.

Nel T1D si verifica una grave turba del metabolismo glucidico con conseguente incapacità dell'organismo a mantenere il livello di glucosio del sangue al di sotto di un certo valore. L'Organizzazione Mondiale della Sanità ha stabilito tale valore come una glicemia dopo pasto di 200 mg/dl o come una glicemia a digiuno di 126 mg/dl, riscontrati in almeno due rilevazioni indipendenti. Esistono diverse forme di diabete che presentano profonde differenze per patogenesi, quadro clinico all'esordio, età media d'insorgenza, severità del decorso e necessità di misure terapeutiche specifiche. In generale, il diabete si origina o per la ridotta produzione d'insulina, ormone chiave nella regolazione del metabolismo glucidico, o per una difettosa utilizzazione dell'insulina stessa legata ad alterazioni recettoriali e post-recettoriali a livello degli organi bersaglio dell'insulina. Esempi paradigmatici di queste due forme di diabete sono rappresentati rispettivamente dal diabete di tipo 1 (T1D), caratterizzato da una severa

deficienza nella produzione d'insulina e dal diabete di tipo 2 (T2D); in cui l'insulina, pur essendo presente e in certe fasi della malattia perfino aumentata, non agisce per una cosiddetta insulino-resistenza a livello degli organi bersaglio. Il T1D è più comunemente causato dall'autodistruzione, operata da linfociti T autoaggressivi, delle cellule  $\beta$  del pancreas, le uniche dell'organismo capaci di produrre efficientemente l'insulina. La malattia è in realtà il risultato finale di un processo infiammatorio cronico e si manifesta clinicamente quando circa il 90% delle cellule  $\beta$  pancreatiche è andato distrutto e quindi solo il 10% è ancora funzionante. A questo punto del processo autodistruttivo la produzione d'insulina da parte delle  $\beta$  cellule residue non è più sufficiente per regolare in maniera ottimale l'ingresso e l'utilizzazione del glucosio all'interno delle cellule dell'organismo. Compare quindi il quadro sintomatologico tipico della malattia, legato all'iperglicemia plasmatica e al concomitante squilibrio metabolico. Nel soggetto diabetico si manifesta minzione frequente, poliuria (cioè aumento della quantità di urina emessa), enuresi o emissione involontaria di urina dopo aver raggiunto il controllo degli sfinteri, polidipsia (sensazione di sete intensa che porta ad un frequente bisogno di bere), diminuzione della massa muscolare, stanchezza e dimagrimento. La causa di questa sintomatologia è una cheto-acidosi metabolica, legata allo scompenso metabolico di base, associata ad una disidratazione iperosmolare, in quanto la concentrazione osmotica del plasma aumenta, per un

incremento della concentrazione del glucosio extra-cellulare. Ciò determina richiamo di acqua dalle cellule al plasma con conseguente grave disidratazione intra-cellulare e marcata sofferenza di vari organi, fra i quali il sistema nervoso centrale, che in situazioni estreme determina il coma diabetico.

Fino a pochi anni fa si riteneva che il T1D fosse una patologia essenzialmente pediatrica e che fosse inevitabilmente associata ad una totale dipendenza dalla somministrazione di insulina esogena fin dall'esordio clinico della malattia. In realtà si è visto che questa forma autoimmune di T1D pur essendo più frequente nei bambini può insorgere in qualunque epoca della vita. In generale la malattia tende a scompensarsi molto più velocemente nel bambino, e richiede obbligatoriamente una terapia insulinica sostitutiva per la sopravvivenza. Nell'adulto, il quadro clinico è spesso più attenuato e in qualche caso all'esordio clinico della malattia e nei mesi immediatamente successivi non si manifesta la necessità di una terapia sostitutiva con insulina esogena. Tale forma di diabete autoimmune dell'adulto è stata denominata "*Latent Autoimmune Diabetes in Adults*" (L.A.D.A.) e può essere confusa da un punto di vista clinico con il T2D (la diagnosi corretta di L.A.D.A. viene effettuata in base alla presenza di autoanticorpi contro la beta-cellula tipici del diabete autoimmune). Nel loro insieme queste osservazioni evidenziano la presenza di una variabilità clinica nell'ambito del T1D.

Per quanto concerne l'incidenza del T1D, essa è differente in Europa a seconda della popolazione considerata: è alta nel nord, in particolare in Scandinavia, e bassa nell'area mediterranea. Così come per la sclerosi multipla, anche per il diabete di tipo 1 esiste un'eccezione a tale gradiente di incidenza: la Sardegna (fig.4). Gli alti valori di incidenza della Sardegna sono simili a quelli della Finlandia (Green et al. 1992; Muntoni et al. 1992). La malattia è inoltre molto rara in Asia e relativamente rara nelle popolazioni africane.

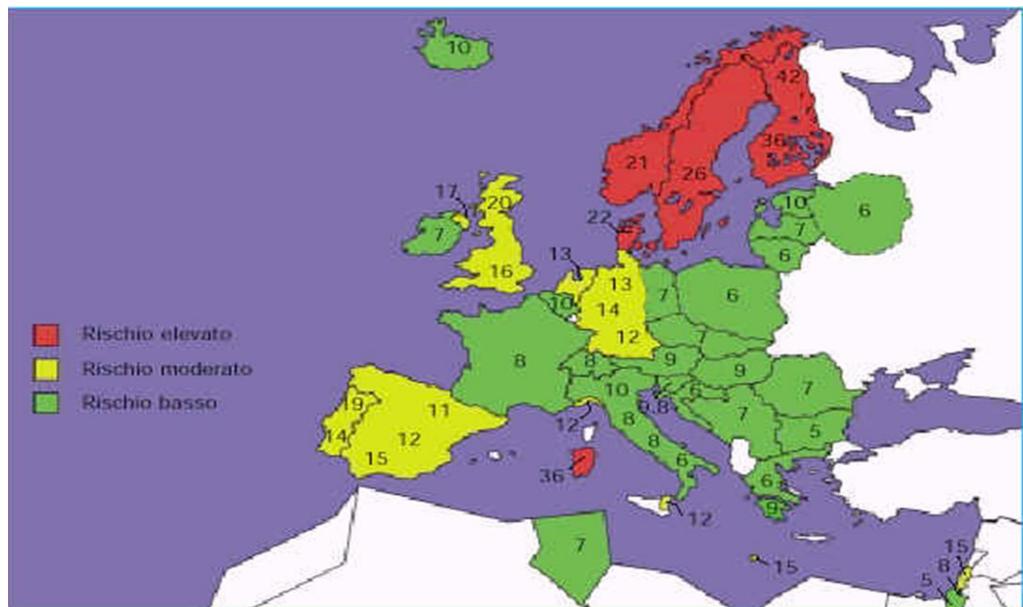


Fig.4 - Incidenza del T1D in Europa e in altri paesi Extraeuropei  
(numero di casi su 100.000 individui ogni anno nel range d'età 0-15 anni)

## **b.1) Ereditarietà e basi genetiche**

Anche il T1D, come la sclerosi multipla, è una malattia multifattoriale causata dall'interazione di fattori genetici predisponenti e di fattori ambientali. Il T1D è una malattia multifattoriale più accessibile all'analisi genetica rispetto alla sclerosi multipla, per la sua diagnosi clinica inequivocabile e per il suo esordio giovanile che consente una più facile raccolta dei DNA parentali.

Per quanto riguarda l'ereditarietà, siamo oggi a conoscenza di una grande quantità di dati che confermano l'importanza della predisposizione genetica nella comparsa della malattia. Il rischio medio di T1D per un individuo scelto a caso nella popolazione Europea è pari a circa lo 0.4%. Questo rischio incrementa al 6% quando l'individuo in questione è fratello di un paziente diabetico. Se questo ipotetico fratello condivide con il malato una totale identità genotipica a livello della regione HLA/MHC il suo rischio di contrarre la malattia incrementa al 12-15%. Infine il rischio di un gemello monozigote (MZ) di un paziente diabetico è pari al 70% (Kyvik et al. 1995).

La malattia mostra una marcata tendenza a segregare nelle famiglie in quanto il rischio di malattia è 15 volte più alto in un fratello di un malato (6%) rispetto a quello della popolazione generale (0.4%) (Todd et al. 1990). I dati indicano inoltre che geni contenuti nella regione HLA sono importanti in quanto il rischio di un fratello di un diabetico è

più alto se condivide con il probando un'identità genetica a livello della regione HLA (12-15%). Anche geni al di fuori della regione HLA sono importanti in quanto il rischio dei gemelli MZ, che condividono una completa identità non solo a livello della regione HLA ma a livello di tutto il genoma, è ancora più alto (70%). Infine, queste osservazioni suggeriscono che sono implicati anche fattori ambientali o epigenetici e quindi la penetranza del T1D è incompleta, in quanto il rischio empirico per gemelli MZ di pazienti T1D è inferiore al 100% (70%). La quantificazione del rischio nei gemelli MZ è particolarmente importante, in quanto fornisce una stima diretta della penetranza per l'intero corredo genetico di suscettibilità (Tisch, R., McDevitt, H. 1996; Vyse, T. J., Todd, J. A., 1996).

Le basi genetiche del diabete di tipo 1 sono estremamente complesse. Tuttavia è ormai chiaro il un ruolo assolutamente fondamentale svolto dal complesso maggiore d'istocompatibilità (MHC), sul braccio corto del cromosoma 6.

Esiste, infatti, una principale regione genica di suscettibilità alla malattia (*IDDM1*) che comprende specifici aplotipi predisponenti definiti da varianti alleliche a livello dei loci DRB1-DQB1 (Zavattari et al. 2001; Zavattari et al. 2000). Questi loci presentano un associazione primaria con il T1D (Herr et al. 2000; Cucca et al. 1995). Tuttavia analisi genetiche sistematiche effettuate negli ultimi anni, hanno rivelato che le varianti ai loci *DRB1* e *DQB1*, presenti in

classe seconda della regione HLA, da sole non spiegano l'intera predisposizione genetica alla malattia conferita dalla regione HLA considerata nel suo complesso (Zavattari et al. 2001). Numerosi studi di linkage e associazione hanno posto in evidenza altri loci localizzati al di fuori della regione HLA, nel tentativo di spiegare le complesse basi genetiche della malattia (Bergholdt, R. et al. 2005; European Consortium for IDDM Genome Studies, 2001; Davies et al. 1994.; Concannon et al. 1994;). È stato inoltre posto in evidenza il legame di geni quali il TCF1, il PTPN22 e OAS1 con la malattia (Yoshiuchi et al. 2001; Bottini et al. 2004,; Kawasaki et al 2006; Tessier et al. 2006).

### c) HLA

La regione HLA rappresenta la versione umana del complesso di maggiore istocompatibilità (MHC) e contiene circa 200 geni localizzati nel contesto di una sequenza lineare di circa 3500000 paia di basi di DNA sul cromosoma 6 (6p21.31) (Altman, D.M. & Trowsdale, J. 1989). Le proteine codificate da molti di questi geni svolgono importanti funzioni nella regolazione delle risposte immunitarie e pertanto sono state e sono tuttora sottoposte a forti pressioni selettive. Queste ultime, unitamente alle esigenze funzionali legate alla necessità di una coordinazione ottimale delle risposte immuni da parte dell'ospite, hanno influenzato non solo la sequenza nucleotidica e quindi proteica ma anche la posizione e l'ordine relativo di questi geni lungo il braccio corto del cromosoma 6. I geni presenti in questa regione sono infatti organizzati in specifici sub-gruppi a formare i cosiddetti clusters genetici. Ciò consente un sistema sofisticato e integrato di regolazione dell'espressione genica. Partendo dal centromero verso il telomero del braccio corto del cromosoma 6 troviamo i geni di classe II, III e I. (fig.5; fig.6).

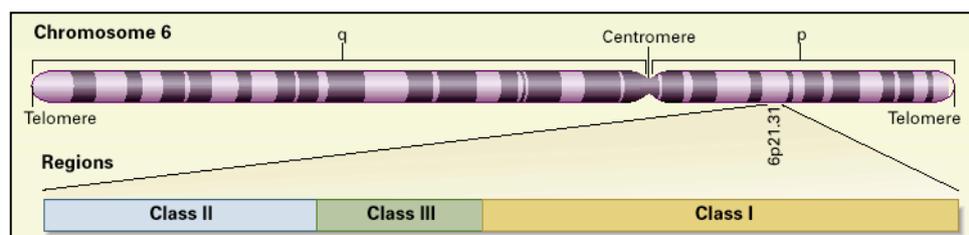


Fig.5 - Localizzazione classi HLA sul braccio corto del cromosoma 6.

Nella classe II e nella classe I sono contenuti una serie di geni che codificano per proteine altamente polimorfiche denominate molecole HLA classiche, in quanto in grado di presentare peptidi ai linfociti T. In particolare, le molecole HLA classiche sono codificate nella subregione di classe I dai loci A, B, C. Le molecole di classe I sono espresse in tutte le cellule e hanno la funzione di presentare peptidi virali ai linfociti citotossici (o CTL) denominati anche CD8+ per la presenza di tale molecola co-accessoria sulla superficie cellulare. Nella subregione di classe II sono presenti numerosi geni tra i quali DP, DQ, DR. Le molecole di classe seconda sono espresse in cellule specializzate, denominate APC o “antigen presenting cells” che sono rappresentate da macrofagi, cellule dendritiche, monociti e linfociti B. I prodotti di tali geni di classe II sono deputati alla presentazione di peptidi di origine batterica e protozoaria ai linfociti T helper o CD4+.

Fra la classe II e la classe I, in classe III, sono presenti un gruppo di geni che codifica per numerose molecole coinvolte nei processi infiammatori tra cui differenti citochine che agiscono da “mediatori” tra le cellule e fungono da segnali intercellulari. Le citochine prodotte dalla classe III hanno quindi la funzione di condizionare il comportamento di altre cellule-specifiche verso cui sono indirizzate.

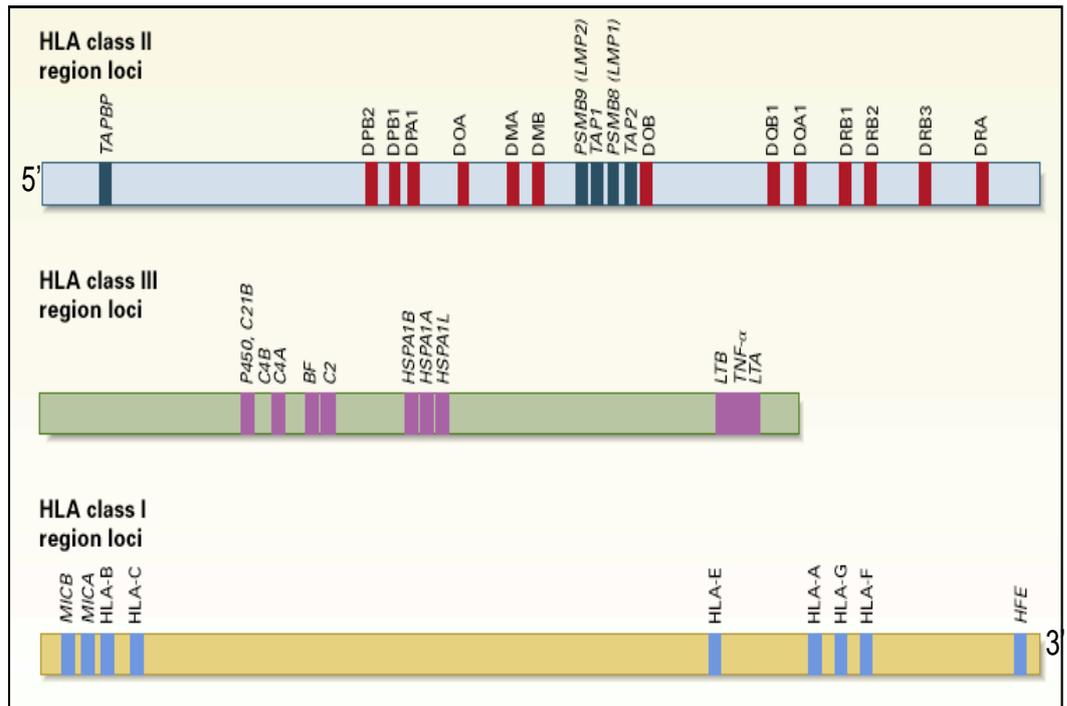


Fig.6 - Disposizione dei loci nelle differenti classi HLA.

Tra i geni di classe terza, il TNF e l'LTA, detto anche TNF $\beta$ , costituiscono l'oggetto di studio di questa ricerca (Fig.7). Questi geni presentano una struttura molto simile. Si estendono entrambi per circa 3 Kb, gli ultimi esoni presentano una omologia significativa (56%), le proteine che producono svolgono delle attività biologiche molto simili e condividono il 30% degli amminoacidi (Pennica et al. 1984, Nedwin et al. 1985).

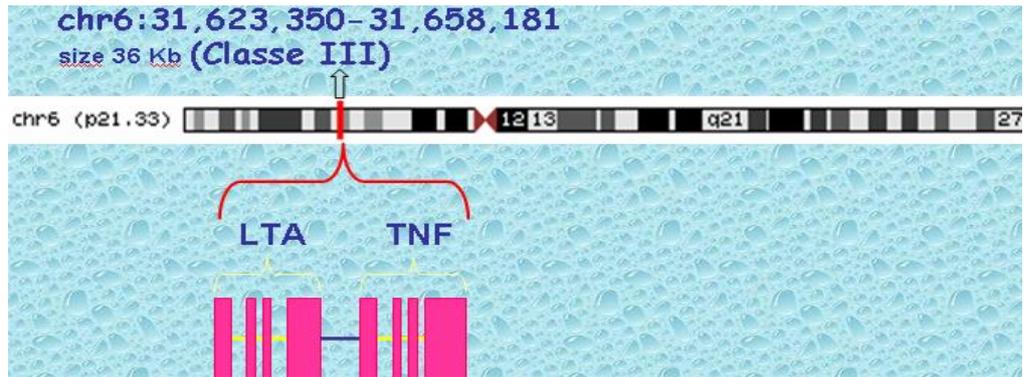


Fig.7 - Localizzazione dei geni di interesse sul cromosoma 6.

### c.1) TNF

Il TNF (Tumor Necrosis Factor) codifica per una citochina proinfiammatoria multifunzionale secreta principalmente dai monociti macrofagi e la sua azione si esplica sul metabolismo lipidico, la coagulazione, l'insulino-resistenza e le funzioni dell'endotelio (Shirai et al. 1985).

Il TNF una volta immesso in circolo si lega ai suoi recettori TNFRSF1A/TNFR1 e TNFRSF1B/TNFR2 che presentano una bassa affinità.

Le azioni del TNF sono differenti a seconda che questa proteina sia presenti in alte o basse dosi.

#### **Azioni del TNF a basse dosi: (10<sup>-9</sup>)**

- Aumenta l'espressione di molecole di adesione sulla superficie delle cellule endoteliali vascolari
- Attiva i leucociti
- Stimola fagociti mononucleati a produrre chemochine
- Aumenta l'espressione di molecole MHC classe I

- Agisce come fattore di crescita per i fibroblasti favorendo la deposizione di tessuto connettivo.

Se lo stimolo per la produzione di TNF è particolarmente intenso questa citochina agisce con un meccanismo di tipo endocrino:

- si comporta come un ormone
- attiva la coagulazione
- inibisce la replicazione delle cellule staminali nel midollo osseo.

Una somministrazione prolungata produce cachessia: distrugge, quindi, adipociti e cellule muscolari.

#### **Azioni del TNF ad alte dosi: (10-7)**

- Riduce la per fusione tissutale
- Causa trombosi intravascolare
- Causa gravi disturbi metabolici, quali un drastico abbassamento della glicemia.

Per quanto riguarda la struttura del gene TNF esso si estende per 2,8 Kb ed è costituito da 4 esoni, il cDNA consta di 1.667 bp (fig 8). Presenta, inoltre un isoforma trascritta e non tradotta.

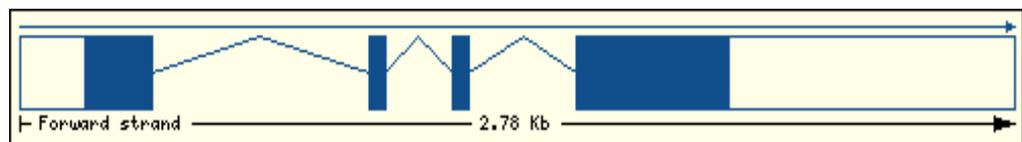


Fig.8 - Struttura del gene TNF: i rettangoli indicano gli esoni, le linee gli introni.

Il TNF è implicato in molte patologie quali la psoriasi, la tubercolosi, l'obesità, l'alzheimer (Boyman et al. 2004; Stein et al. 2005; Rosmond et al. 2001; Ramos et al.2006). Inoltre, come si spiegherà meglio successivamente, il TNF potrebbe avere un ruolo importante in patologie autoimmuni quali il diabete e la sclerosi multipla.

### **c.2) LTA**

L'LTA detta anche linfotossina  $\alpha$  o TNF $\beta$  appartiene, come il TNF, alla famiglia del tumor necrosis factor. È una citochina prodotta dai linfociti. LTA è altamente inducibile e secreta; esiste come molecola omotrimerica. Forma eterotrimeri con la linfotossina  $\beta$  che ancora la linfotossina  $\alpha$  alla superficie cellulare. La linfotossina  $\alpha$  condivide con il TNF i recettori TNFRSF1A/TNFR1 (Gray et al. 1984). In tabella sono riassunte le caratteristiche principali della proteina (Tab.1).

<b>Ave. residue weight:</b>	108.764
<b>Charge:</b>	8.5
<b>Isoelectric point:</b>	9.3578
<b>Molecular weight:</b>	22,296.69
<b>Number of residues:</b>	205

Tab.1 - Caratteristiche principali dalla proteina LTA

La fig. 9 rappresenta la struttura tridimensionale della proteina. L'LTA media una grande varietà di risposte infiammatorie, immunostimolatorie e antivirali, è coinvolta nella formazione

di organi linfoidi secondari durante lo sviluppo e gioca, inoltre, un ruolo chiave nell'apoptosi.

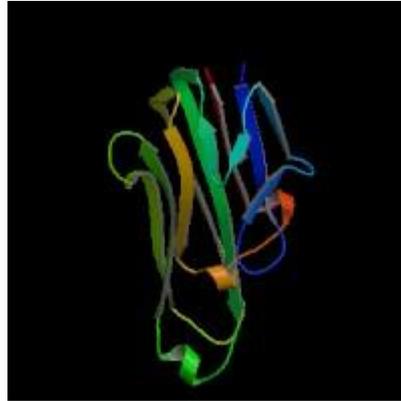


Fig.9 - Struttura tridimensionale dell'LTA vista frontalmente.

Per quanto concerne la struttura del gene esso si estende per 2 Kb ed è costituito da 4 esoni; il cDNA consta di 1.386 bp (Fig.10). Presenta, inoltre un'isoforma trascritta e non tradotta.

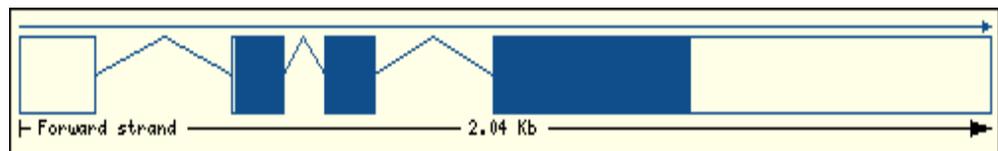


Fig.10 - Struttura del gene LTA: i rettangoli indicano gli esoni, le linee gli introni.

Diversi studi indicano l'associazione di questo gene con l'infarto al miocardio (Ozaki et al. 2002; PROCARDIS Consortium 2004). Tuttavia l'associazione di questo gene con l'infarto al miocardio non è ancora del tutto chiara e, come spiegato successivamente, non si esclude un suo ruolo in

malattie autoimmuni come la sclerosi multipla e il diabete di tipo 1.



## **Scopo della ricerca**

Il progetto di ricerca è volto allo studio di due malattie multifattoriali, la sclerosi multipla e il diabete di tipo 1, la cui insorgenza è, quindi, dovuta alla presenza ed interazione di numerosi fattori sia genetici che ambientali. L'arduo compito di identificare i determinanti genetici delle malattie non mendeliane è, indubbiamente, una meta fondamentale della genetica umana. Il riscontro che si potrebbe ottenere sulla società da tale mole di lavoro sarebbe estremamente importante in quanto, nel mondo occidentale, il principale contributo alla morbilità e mortalità deriva proprio dalla componente genetica di malattie complesse. L'identificazione dei fattori predisponenti potrebbe avere diverse fondamentali implicazioni in ambito diagnostico e rappresentare una importante "via di accesso" verso la terapia delle malattie. Si potrebbero sviluppare tests diagnostici, nuovi farmaci e, in generale si comprenderebbero meglio i meccanismi eziopatogenetici alla base della sclerosi multipla e del diabete di tipo 1.

Benché diversi studi abbiano rivelato l'associazione tra la regione HLA e le patologie in esame, si è ancora lontani dal

conoscere, comprendere appieno e capire le interazioni dei fattori genetici che ne determinano l'insorgenza.

Dunque, il principale scopo di questo lavoro è trovare le varianti genetiche che influenzano il rischio di ammalare di sclerosi multipla e diabete di tipo 1 attraverso lo studio di due geni candidati il TNF e LTA contribuendo quindi alla risposta che il mondo scientifico tenta di dare al paziente che si chiede "Perché mi sono ammalato?".



## **Presupposti e disegno sperimentale**

La scelta delle malattie oggetto del progetto è stata dettata dalle caratteristiche uniche del campione a nostra disposizione: la popolazione sarda. Questa è, infatti, il substrato più adatto per studiare la sclerosi multipla e il diabete di tipo 1 poiché, data la forte presenza di queste patologie, tale popolazione rappresenta un'eccezione al gradiente di incidenza. Il presente progetto di ricerca esamina l'SM e il T1D partendo dall'ipotesi che esistano varianti comuni in grado di predisporre ad entrambe le malattie. Infatti, in uno studio effettuato sulla popolazione sarda, è stato osservato che l'incidenza di diabete nei pazienti con sclerosi multipla è circa 3 volte superiore rispetto a quella dei fratelli o sorelle sani e maggiore di circa 5 volte rispetto alla popolazione generale. Inoltre, la presenza di sclerosi multipla in altri consanguinei aumenta il rischio di diabete di tipo 1 nei fratelli o sorelle di individui con sclerosi multipla (Marrosu et al. 2002).

L'attenzione del progetto si è focalizzata sulla regione HLA di classe III ed in particolare su due geni, il TNF e LTA, sulla base di accurate interpretazioni dei dati presenti in letteratura e di una serie di intuizioni ed ipotesi riguardanti sia il diabete che la sclerosi multipla.

Per quanto riguarda il diabete i principali loci di suscettibilità DR e DQ sono posizionati in classe II. Ma è anche noto che esistono effetti modificatori del rischio al di fuori dei loci DR-DQ. Con test di associazione condizionale è stato dimostrato che per i loci DRB1 e DQB1 con gli alleli \*0301-\*0201 esistono tre effetti modificatori del rischio dati dai loci DMB, DOB e TNFc i quali presentano una forte eterogeneità (Zavattari et al. 2001) (Fig.11). Il marker TNFc, ricade appunto in classe III e questo è uno dei motivi per cui l'attenzione del progetto si è rivolta verso questa localizzazione.

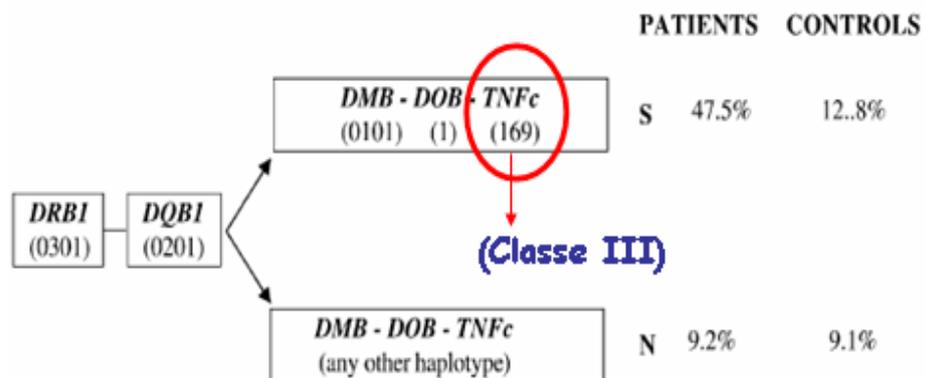


Fig.11 - Effetti modificatori del rischio conferito dai loci DRB1 e DQB1.

Per quanto riguarda, invece, la sclerosi multipla le considerazioni tratte dal lavoro di Marrosu et al. 2001 hanno indirizzato la ricerca verso la classe III. Questo lavoro indica la presenza di cinque aplotipi DQB1 e DRB1 in associazione con la malattia. Sono stati trovati in associazione anche due polimorfismi: il D6S1683 e l'82-1. La scelta della classe III si

è basata proprio sulla posizione di quest'ultimo che, infatti, ricade nella regione di interesse (Fig.12).

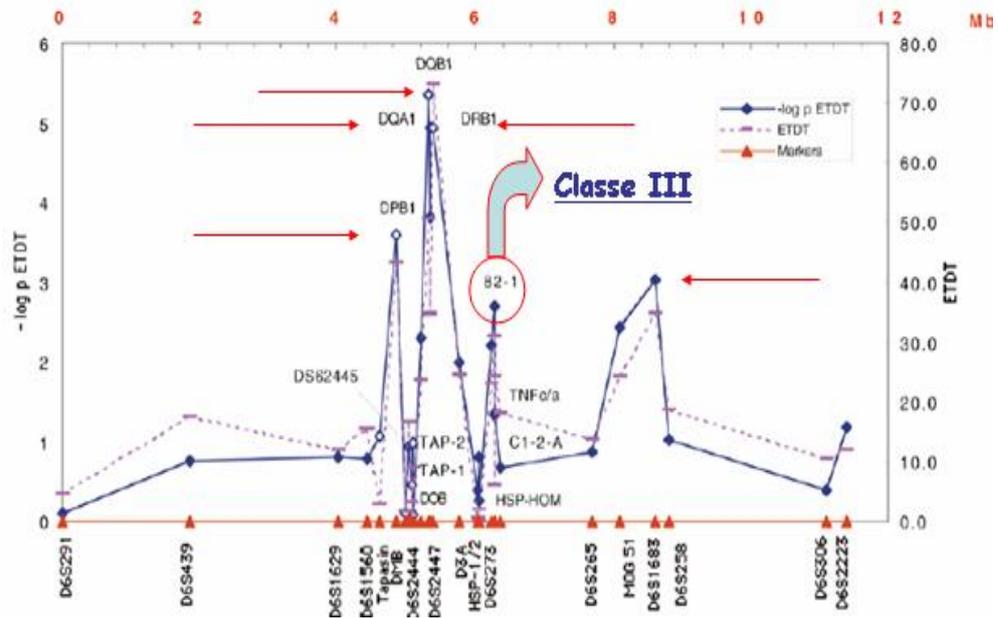


Fig.12 - Le frecce indicano i picchi dati dagli aplotipi di predisposizione ed il marker 82-1

All'interno di tale classe sono stati individuati dei geni candidati in base alla posizione dei markers TNFc e 82-1. Infatti nel gene LTA ricade il marker TNFc e il gene TNF è molto vicino al marker 82-1. La scelta dei geni di interesse si è basata anche sulle loro funzioni e caratteristiche descritte in letteratura.

Studi di tipo farmacologico, sui topi knockout suggeriscono la funzione neuroprotettiva della citochina per cui codifica facendo intuire, quindi, il ruolo che essa potrebbe avere nella sclerosi multipla (Nancy et al. 2004; Junliang et al. 1998).

Da uno studio sulla espressione del TNF e dell'LTA si evince che l'associazione della SM con l'HLA-DR2 potrebbe essere

legato ad una propensione delle cellule T a produrre maggiori quantità di queste citochine (Zipp et al. 1995). Inoltre è stata dimostrata l'associazione di un polimorfismo del promoter del TNF con la SM indipendente dall'associazione con la classe seconda dell'HLA (Fernandez-Arquero et al. 1999).

Infine, è stato dimostrato che una sovrapproduzione del TNF è implicata in diverse malattie autoimmuni tra cui il diabete di tipo 1 e la sclerosi multipla (Ruuls and Sedgwick 1999).

L'idea che possa esistere un comune denominatore tra le due malattie e i dati presenti in letteratura ci spingono verso l'idea che TNF ed LTA siano due forti geni candidati e che possano rappresentare delle varianti in grado di modificare il rischio al diabete di tipo 1 o alla sclerosi multipla o ad entrambe le malattie.

Per confutare questa ipotesi il disegno sperimentale adottato si avvale di più steps di lavoro differenti per SM e T1D.

#### **a) Sclerosi multipla: sequenziamento e tipizzazione.**

- Sono stati selezionati individui che presentavano un particolare aplotipo di forte predisposizione e sono stati sequenziati per i geni TNF ed LTA a partire da cDNA per coprire le regioni esoniche e da DNA per coprire le regioni in 5', 3' e la regione dell'LTA contenente un'isoforma.
- Dopo aver individuato dei polimorfismi (SNPs) ne sono stati scelti alcuni, sulla base delle loro caratteristiche,

per tipizzarli con l'uso di sonde TaqMan su un vasto campione di famiglie sarde con un figlio affetto.

- È stata valutata la loro potenziale associazione con la malattia grazie al TDT (transmission disequilibrium test).

#### **b) Diabete di tipo 1: sequenziamento, ricerca in silico e tipizzazione.**

- Per quanto riguarda il diabete, oltre al sequenziamento di 8 genitori di individui affetti, è stata svolta una ricerca in silico riguardanti gli aplotipi PGF, COX e QBL sequenziati in studi precedenti. È stata fatta una comparazione tra le sequenze, presenti in rete, dei differenti aplotipi.
- Valutando diverse caratteristiche, è stato assegnato uno score di importanza ad ogni polimorfismo presente in sequenza. In base allo score sono stati scelti gli SNPs da tipizzare con l'uso delle sonde TaqMan.
- Come per la SM, è stata valutata la loro potenziale associazione con la malattia grazie al TDT (trasmission disequilibrium test).



## **Materiali e metodi**

### **a) Selezione del campione**

Nel corso del progetto sono state fatte due selezioni del campione: una per il sequenziamento dei due geni e una per la tipizzazione su larga scala di polimorfismi di interesse.

Per poter comprendere il modo in cui sono stati selezionati i campioni per il sequenziamento occorre aprire una piccola parentesi riguardante gli aplotipi HLA.

I geni MHC sono codominanti, vengono quindi espressi i prodotti proteici dei geni di entrambe i cromosomi.

Sappiamo che un aplotipo è una particolare combinazione di varianti alleliche in una specifica regione di un cromosoma e sulla base degli alleli dei marcatori HLA si possono suddividere gli aplotipi HLA in predisponesti e non predisponesti.

Per quanto riguarda la suscettibilità genetica alla sclerosi multipla e al diabete esistono aplotipi che risultano predisponesti o meno o per una o per l'altra malattia. L'aplotipo DR3 B8 o COX (poco rappresentato in Sardegna)

risulta essere non predisponente per la sclerosi multipla e predisponente per il diabete. Il DR3B18 o QBL che è molto diffuso risulta essere predisponente per entrambe le malattie mentre il DR2 PGF è molto predisponente alla sclerosi multipla e non predisponente al diabete. Con le sigle” Cox QBL e PGF si indicano le linee cellulare tramite cui sono stati studiati questi aplotipi.

Per selezionare una casistica su cui studiare i quattro geni di interesse ci si è basati proprio sulle loro caratteristiche aplotipiche in modo da capire come si presentano i geni i esame per differenti aplotipi, predisponenti e non, con l’obiettivo di analizzare le varianti alleliche in grado di modificare il rischio di predisposizione alle malattie.

Sono, dunque, state selezionate due casistiche: una per il diabete e una per la sclerosi multipla.

Per il T1D, la casistica è costituita da 8 genitori di individui affetti in modo da capire cosa potrebbero trasmettere ai figli malati. I campioni presentano diversi aplotipi in maniera tale da poter individuare punti in comune o divergenti per i geni di interesse di classe III.

La casistica è così composta:

- 2 individui che presentano sul cromosoma trasmesso l’aplotipo DR3-B18-TNFC2 e su quello non trasmesso l’aplotipo DR3-non B18-TNFC2
- 2 individui che si presentano come DR3-non B18 e non-DR3-B18
- 2 individui con genotipo DR3-B18 e DR3-B8

- 2 individui che presentano sul cromosoma trasmesso l'aplotipo DR3-TNFC2-B18 e su quello non trasmesso l'aplotipo DRX (qualsiasi aplotipo che non sia DR3)-TNFC1-non B18.

Con TNFC si indica un marcatore (SNP) presente in classe III e le sigle 1 e 2 indicano la presenza dei differenti alleli.

Per la sclerosi multipla sono stati inclusi 4 individui affetti con aplotipo nonDR3-B18.

Il campione per la tipizzazione dei polimorfismi di interesse sia per la SM che per il T1D è costituito da famiglie sarde con un figlio affetto e altri individui non affetti.

Una descrizione del campione utilizzato per la SM e per diabete è data dalle tabelle 2 e 3.

	n		N totale	y		Y totale	Totale complessivo
Ped ID	f	m		f	m		
1		396	396				396
2	472		472	3		3	475
3				349	140	489	489
4	169	118	287	6		6	293
5	49	37	86	1	1	2	88
6	20	12	32				32
7	5	2	7				7
8	2	1	3				3
Totale complessivo	717	566	1283	359	141	500	1783

Tab.2 casistica sclerosi multipla

IDPed	n		n Totale	y		y Totale	Totale complessivo
	f	m		f	m		
1		799	799		33	33	832
2	814		814	18		18	832
3				364	468	832	832
4	270	248	518	56	55	111	629
5	92	75	167	8	14	22	189
6	27	23	50	3	1	4	54
7	13	9	22	2		2	24
8	3	7	10				10
9	2	2	4				4
10	1	1	2				2
11	1		1		1	1	2
12		1	1				1
Totale complessivo	1223	1166	2389	451	571	1022	3411

Tab 3. casistica diabete di tipo 1

Nella Colonna Ped ID vengono indicati gli individui che compongono le famiglie, quindi 1 è sempre il padre, 2 sempre la madre, il 3 è sempre l'affetto e gli altri numeri indicano i fratelli o sorelle del malato (che possono essere sani o affetti a loro volta) .

Le colonne 2, 3 e 4 sono relative ai non affetti: i maschi vengono indicati con “m” e le femmine con “f”. Le colonne 5, 6 e 7 sono relative agli individui affetti; dove con ”y” si indica lo stato positivo di affezione e con “n” si indica lo stato di individuo sano.

Per quanto riguarda l'SM il numero totale di individui genotipizzati è di 1783 di cui non affetti 1283 e affetti 500. Come si nota dalla riga 3, indicante sempre il figlio affetto, il totale è di 489 di cui 349 femmine e 140 maschi. Su 293 famiglie con 4 componenti 2 presentano più di un figlio affetto, nel dettaglio, sono 6 e tutte femmine. Come si può notare dalla tabella, il numero dei genitori non è uguale al quello degli individui affetti, poiché non per tutte le famiglie erano disponibili i DNA di entrambi i genitori. Siccome il TDT, il test usato nelle analisi dei dati, utilizza solo famiglie complete il numero totale di famiglie analizzate è di 396.

Per quanto riguarda il diabete, il numero totale di individui genotipizzati è di 3411, di cui 2389 non affetti e 1022 affetti. Il primo figlio affetto (riga 3) è dato in 364 casi da femmine e in 468 da maschi, per un totale di 832. Ben 629 famiglie presentano 2 figli affetti (riga 4) di cui 56 femmine e 55 maschi. Da notare, inoltre, che 18 madri su 814 e 33 padri su 799 sono affetti. Tutte le famiglie dei diabetici risultano complete ovvero con entrambi i genitori, per cui il numero totale di famiglie analizzate tramite TDT è di 832.

## **b) Estrazione del DNA**

Il DNA è stato estratto con la metodica dell'estrazione salina, eseguita su 10 ml di sangue periferico trattato con EDTA pH8, al fine di impedirne la coagulazione.

I passaggi fondamentali sui quali tale metodica si fonda sono qui di seguito schematizzati:

- 1) al fine di ottenere la lisi dei globuli rossi, viene eseguito un lavaggio con il Lysis Buffer;
- 2) la precipitazione dei globuli bianchi viene ottenuta tramite precipitazione, centrifugando il campione con Fisio Buffer a 4000 rpm;
- 3) dopo aver risospeso il pellet in Buffer A, si procede alla lisi dei globuli bianchi tramite l'aggiunta di SDS e Proteinasi K;
- 4) la precipitazione delle proteine viene ottenuta con NaCl soprassaturo;
- 5) il DNA viene ottenuto per precipitazione con isopropanolo assoluto, visualizzato come "nubecola" e quindi separato per essere poi risospeso in TE.

### **c) Estrazione dell'RNA**

L'RNA è stato estratto da sangue periferico trattato con EDTA pH8, al fine di impedirne la coagulazione.

I passaggi fondamentali sui quali tale metodica si fonda sono qui di seguito schematizzati:

- sono stati aggiunti 500 µl di Trizol in ciascuna eppendorf contenente il pellet di cellule, il pellet è stato vortexato e lasciato per 15 minuti a RT.
- sono stati aggiunti 100µl di cloroformio e dopo aver agitato vigorosamente per 15 secondi, i campioni sono stati lasciati a RT per 2-3 minuti.
- centrifugata a 4°C per 15 minuti a 12000rcf.

- in eppendorf nuove è stata trasferita solo la fase acquosa aspirandola dal menisco, facendo attenzione a non prelevare le fasi sottostanti che sono state eliminate.
- sono stati aggiunti 25µl di isopropanolo, dopo aver agitato sono stati lasciati in ghiaccio per 10 minuti
- centrifugata per 10 minuti a 4°C a 12000rcf e rimozione del sopranatante.
- è stato aggiunto al pellet 0.5ml di etanolo al 75%.
- centrifugata a 12000rcf per 5 minuti a 4°C.
- è stato rimosso il surnatante e lasciato evaporare l'etanolo residuo posizionando l'eppendorf contenente il pellet o in camera calda o sotto la lampada, per circa 30 minuti.
- il pellet di RNA è stato risospeso con 40 µl di acqua distillata.

#### **d) Retrotrascrizione**

La retrotrascrizione, per ottenere dall'RNA dei filamenti di cDNA, è stata effettuata con l'utilizzo di un apposito kit che prevede l'uso di random primers in grado di dare una buona resa di reazione e un prodotto di amplificazione affidabile.

#### **e) Sequenziamento**

Per effettuare le sequenze, al fine di individuare polimorfismi di interesse per la SM e il T1D, il cDNA sia del TNF che dell'LTA è stato suddiviso in 4 frammenti. I primers sono stati disegnati a cavallo degli esoni e overlappanti tra loro.

Per la reazione di PCR sono stati utilizzati i seguenti primers:

LTA, cDNA:

<i>Forward-t.m.</i>	<i>Reverse-t.m.</i>
Gcttcgtgctttggactacc-60°	Acctcattggagaccccag-60°
Cctgggccttggttct-53°	Aggtccagctcttctcctcc-63°
Cccaccagtggcatctactt-60°	Cctgatgtctgtctggctga-60°
Gccattccaacagctcaagt-58°	Cgatgaagcccaataaacct-56°

TNF, cDNA:

<i>Forward-t.m.</i>	<i>Reverse-t.m.</i>
Ctcctcagcaaggacagca- 63°	Cccaggcagtcagatcatctt-61°
Cctcttctcttctctgatcg-60°	Ctatctgggaggggtcttc-60°
Ccgtctctaccagacca-58°	Aagaattcaaactggggcct-56°
Acctgggattcaggaatgtgt-60°	Ccctactattcagtggcgaga-61°

Per coprire le regioni in 5' e 3' e la regione dell'LTA che comprendeva l'isoforma, le sequenze sono state effettuate su DNA utilizzando i seguenti primers:

LTA:

GGCCTCAGCCTTTCCTGCCT = for 5'
CAGGGCGGGAGGTCTGTCTT = rev 5'
CCCATGACACCACCTGAACG = for
TGGAGAGCAAACACGGACCG = rev
AGGAACCCAAGCATCCACCC = for
TCTTCTTTGGAGCCTTCGCTCTG =rev
GGACCACACCTCCCTGAACCA = for 3'
GGGCTCAGAGAGCTTCAGGGA = rev 3'

Per coprire le regioni in 5' e 3' del TNF le sequenze sono state effettuate su DNA utilizzando i seguenti primers:

AAGAATCATTCAACCAGCGG= for 5'
CAGAGGGAAGAGGTGAGTGC = rev 5'
ACAGACACAACCTCCCCT= for 3'
CCTGAACCAGCCACTTAACC= rev 3'

Dopo aver effettuato le PCR, si procede con la purificazione enzimatica tramite Exo-Sap. Per le reazioni di sequenza sono stati usati gli stessi primers utilizzati in PCR; le reazioni sono state condotte sia in forward strand che in reverse strand. I prodotti delle reazioni di sequenza sono stati purificati tramite precipitazione con etanolo. Le reazioni di sequenza sono state effettuate con l'utilizzo del metodo Sanger tramite big-dye terminators marcati con fluorocromi differenti a seconda della base del codice genetico.

Il DNA generato in sequenza si presenta sotto forma di frammenti dati da un'interruzione controllata della replicazione enzimatica. Si utilizza per questo scopo la DNA polimerasi I, che ha la funzione di copiare una particolare sequenza di DNA a singolo filamento. La sintesi utilizza un primer complementare al filamento di DNA amplificato e quattro desossiribonucleotidi trifosfato (dATP, dCTP, dGTP, dTTP), che permettono l'allungamento del filamento in formazione. La miscela di incubazione deve contenere inoltre

un analogo 2', 3' dideossiribonucleotide di ogni base (ddATP, ddTTP, ddCTP, ddGTP) marcato per ciascuna base con una diversa fluorescenza. L'incorporazione dei dideossi interrompe a diversi livelli la crescita della nuova catena perché privi del terminale ossidrilico in 3' necessario per formare il successivo legame fosfodiesterico. Vengono così casualmente prodotti frammenti di lunghezza diversa tutti con un dideossi marcato all'estremità 3'. Le molecole di DNA sintetizzate a singola elica vengono poi caricate su un sequenziatore automatico capillare e separate per via elettroforetica grazie alla differenza di potenziale applicata ai capillari della macchina. Al passaggio delle molecole davanti alla lampada laser, i fluorocromi si ecciteranno, rilasciando un'energia che verrà captata e rielaborata da un software. I dati analizzati automaticamente verranno poi visualizzati come picchi di diverso colore, cui verranno attribuite le basi del codice genetico. Si otterrà così una sequenza di basi corrispondente alla sequenza stessa del frammento di PCR.

#### **f) Digestione enzimatica**

Nel gene LTA è stato individuato un polimorfismo la cui lettura in sequenza non era di facile interpretazione. È stato ritenuto opportuno valutare l'effettiva presenza dello SNP tramite digestione enzimatica con l'utilizzo dell'enzima FAU I. Questa tecnica consiste nell'analisi diretta dell'amplificato mediante elettroforesi dopo digestione con enzimi di restrizione. Tale approccio può essere utilizzato per tutte

quelle mutazioni puntiformi che aboliscono o creano un sito di riconoscimento per un enzima di restrizione. Dopo aver amplificato il segmento del gene da studiare, il prodotto ottenuto viene successivamente sottoposto a digestione con l'enzima appropriatamente selezionato.

Il pattern di frammenti ottenuti mediante elettroforesi su gel d'agarosio viene confrontato con opportuni marcatori di peso molecolare, cioè una serie di frammenti di DNA a lunghezza nota con cui si possono determinare le dimensioni dei nuovi frammenti generati dai tagli degli enzimi di restrizione.

### **g) ARMS-PCR**

Per testare l'effettiva presenza di due SNPs trovati in sequenza sull'LTA è stata utilizzata la tecnica dell'ARMS-PCR.

I primers interni usati per le reazioni sono i seguenti:

#### 1° SNP

GCACCCCAAGATGCATCTTGCCCACAGGAC  
TACCAATGAGGTGAGCAGCAGGTTTGACGT

#### 2° SNP

CACCTGAACGTCTCTTCCTCCCAAGGGAGT  
CAGAAGGAGGAGGTGTAGGGTGGTGCCTCG

L'ARMS PCR si basa sull'abilità della DNA polimerasi di incorporare specifici deossiribonucleotidi complementari nella sequenza del template. La metodica prevede l'utilizzo di due coppie di primers: la prima produrrà un frammento identico

alla nostra sequenza target che fungerà anche da controllo dell'avvenuta reazione di amplificazione, mentre gli altri due primers (detti primers interni) si accoppieranno rispettivamente a uno dei due primers esterni e riconosceranno il sito mutato o il sito wild-type. Ciò è possibile in quanto l'ultima base in 3' dei suddetti primers riconoscerà appunto il sito mutato o il sito normale. Con questa tecnica è stata confermata la presenza degli SNPs individuati in sequenza.

#### **h) Selezione SNPs di interesse: calcolo dello score di rilevanza**

La selezione degli SNPs di interesse da tipizzare su larga scala su tutti i campioni a nostra disposizione, è stata eseguita dopo una attenta analisi delle informazioni presenti nei browser in rete sugli aplotipi DR2 (PGF), DR3-B8 (COX) e DR3-B18 (QBL).

Le sequenze presenti in rete sono state allineate grazie a un apposito software e in tal modo sono stati evidenziati i punti in cui si differenziano i 3 aplotipi. È stata creata una scala di importanza per tutti gli SNPs trovati nell'allineamento. Per ottenere uno score di rilevanza per ciascun polimorfismo abbiamo utilizzato differenti criteri di selezione (tab 4).

PGF=QBL
QBL≠COX
LOCALIZZAZIONE
CONSERVAZIONE
SPLICING
SOSTITUZIONE AMMINOACIDICA

Tab. 4 criteri di selezione SNPs

Ad ogni SNPs è stato assegnato un punteggio per ogni criterio.

- Uguaglianza allelica tra PGF DR2 e QBL DR3B18: è stato assegnato 1 nel caso in cui ciò fosse vero e 0 nel caso in cui fosse falso
- Differenza allelica tra COX DR3B8 e QBL DR3B18: è stato assegnato 1 nel caso in cui ciò fosse vero e 0 nel caso in cui fosse falso
- Localizzazione: è stato assegnato 2 nel caso di uno snp esonico, 1 per quelli intronici e nel promoter e 0 per gli altri.
- Conservazione: è stato assegnato 2 se maggiore di 0.5, 1 se minore e 0 se pari a zero.
- Splicing: sono stati utilizzati 3 siti internet di predizione dello splicing (<http://125.itba.mi.cnr.it/~webgene/wwwspliceview.html>; <http://www.cbs.dtu.dk/services/NetGene2/>; [http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)). È stato assegnato 3 nel caso in cui tutti i siti confermassero un sito accettore o donatore, 2 se solo due di questi

indicavano un sito, 1 se solo uno indicava un sito e 0 se non veniva indicato nessun sito.

- Sostituzioni amminoacidiche: è stato assegnato 3 nel caso in cui vi fosse una sostituzione non sinonima, 2 nel caso di una sostituzione sinonima e 0 per gli SNPs che non danno sostituzioni amminoacidiche.

Successivamente i punteggi sono stati sommati per ottenere uno score di rilevanza e in base ad esso scegliere i polimorfismi da tipizzare.

### **i) Analisi TaqMan: discriminazione allelica**

Gli SNPs selezionati per LTA e il TNF sono stati tipizzati sui campioni a nostra disposizione tramite l'utilizzo di uno strumento di Real Time Fast PCR in grado di eseguire delle letture in end-point (ovvero al termine della reazione di PCR).

La tipizzazione si è basata sull'utilizzo di oligonucleotidi allele-specifici (ASO) come sonde per l'ibridazione.

Tale strategia (che implica l'ibridazione con oligonucleotidi allele-specifici e la determinazione attraverso il trasferimento di energia di risonanza fluorescente, detta FRET) si basa sull'utilizzo di un oligonucleotide innovativo, il quale risulta essere complementare alla sequenza bersaglio da amplificare nella regione del polimorfismo di interesse ed è in grado di appaiarsi tra i due primers forward e reverse. La sonda è in grado di riconoscere specifiche sequenze nel DNA per la complementarietà delle basi, e di generare un segnale sequenza specifico, che viene rivelato durante la PCR,

permettendo quindi l'identificazione di una variazione puntiforme. Con la discriminazione allelica è, infatti, possibile identificare forme differenti dello stesso gene che differiscono per un solo SNP.

La FRET (Fluorescence Resonance Energy Transfer) è il metodo di rilevazione utilizzato dallo strumento. Il trasferimento di energia avviene solamente nel caso in cui siano soddisfatte due condizioni. In primo luogo, lo spettro di emissione della molecola fluorescente “donatore” (nel caso della metodica in esame, viene chiamata Reporter) deve sovrapporsi con la lunghezza d'onda d'eccitazione della molecola “accettore” (Quencher); in secondo luogo, le due molecole devono trovarsi molto vicine fra loro altrimenti il trasferimento di energia cade rapidamente all'aumentare della distanza (Fig.13).

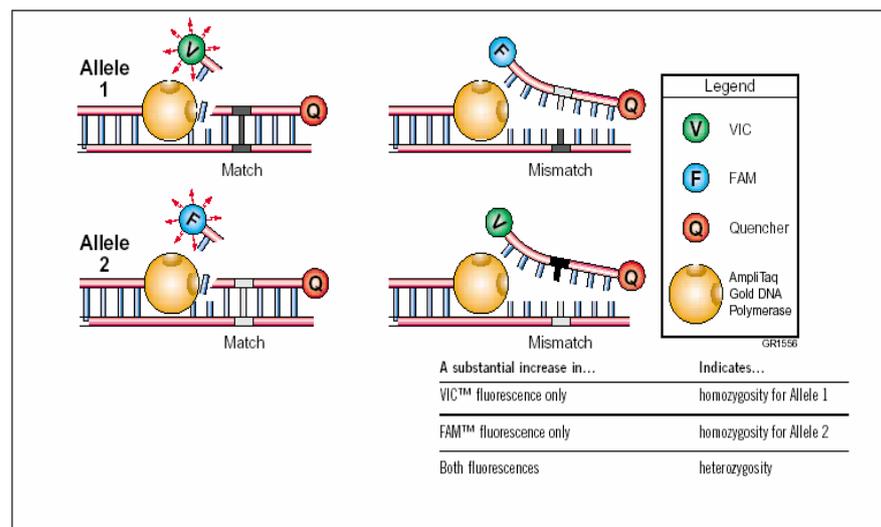


Fig.13 - Meccanismo di reazione della metodica di discriminazione allelica.

Tale metodica fa parte dei metodi definiti omogenei, in quanto non presenta fasi di separazione e l'andamento della tipizzazione dei campioni può essere monitorato in tempo reale durante la PCR. In essa sono contemporanee due fasi che nella maggior parte delle altre strategie di indagine risultano distinte: la PCR e i processi di analisi post-PCR. Questo consente una notevole riduzione dei tempi di esecuzione e di analisi. La discriminazione allelica attraverso il TaqMan è completamente automatizzabile e presenta un'altissima sensibilità e affidabilità.

Il protocollo di reazione utilizzato è esposto in tab. 5:

<i><b>Reagenti</b></i>	<i><b>Quantità</b></i>
DNA	5 a 4ng/ µl
Sonda	0,125 µl (40X)
Mix	2,5 µl (2X)
H <sub>2</sub> O	0,375 µl

Tab. 5 protocollo di reazione per tipizzazione taqman

Il DNA e la mix di reagenti sono stati dispensati, con un processo automatizzato, in apposite piastre su cui è avvenuta la reazione attraverso l'uso dello strumento Real Time Fast PCR.

## **j) Analisi statistica**

Dopo aver effettuato le tipizzazioni degli SNPs di interesse, i risultati sono stati analizzati tramite il test di disequilibrio della trasmissione (TDT) (Schaid 1998).

Il TDT è in grado di valutare se un dato marcatore sia o meno associato con la malattia in esame. Il test confronta il numero di genitori che trasmettono ai figli affetti un dato allele con quelli che trasmettono l'altro allele. Quanto più la trasmissione di un dato allele si discosta dal 50% tanto più il dato che si ottiene è rilevante. La soglia di significatività è data dal P-value; sono significativi tutti i valori di P-value inferiori a 0,05. Maggiore è il numero di famiglie tipizzate maggiore è il potere statistico del test e, quindi, l'attendibilità del dato ottenuto. Durante il test vengono scartate tutte le famiglie in cui non si riesca determinare la trasmissione allelica, ad esempio quando entrambi i genitori sono omozigoti per lo stesso allele. Quest'ultima caratteristica del test pone l'attenzione sulla necessità di tenere conto della eterozigosità dei marker da tipizzare in quanto marker poco polimorfici portano ad una drastica riduzione del numero di famiglie analizzabili e quindi ad un abbassamento del potere statistico.



## **Risultati e discussione**

### **a) Sequenziamento**

Per quanto riguarda il TNF, che consta di 1.667 bp di cDNA, dalla lettura delle sequenze non è emerso nessun risultato rilevante, in quanto non è stato trovato nessun polimorfismo nè per la casistica T1D nè per gli individui affetti da SM.

Nell'LTA, il cui cDNA consta di 1.386 bp, sono stati individuati tre SNPs (polimorfismi biallelici).

Il primo SNP (rs2239704) si trova nel 1° esone e ricade nella regione trascritta e non tradotta. Il secondo SNP (rs17859106) ricade nel 2° esone e nella forma wild-type (t) la tripletta a cui appartiene codifica per una cisteina, nella forma mutata (c) la tripletta codifica per una arginina.

Il terzo SNP (rs17859107) individuato si trova nel 3° esone: è la seconda base del codone e nella forma wild-type (c) la tripletta codifica per una treonina, nella forma mutata (c) la tripletta codifica per una asparagina.

Per poter avere un conferma dei risultati trovati in sequenza, la presenza degli SNPs è stata testata nei campioni anche con altre metodiche. Nello specifico, per lo SNP contenuto nel primo esone è stata effettuata una digestione enzimatica

tramite l'enzima FAU I. Mentre per gli altri due SNPs è stata utilizzata la tecnica della ARMS-PCR.

Con queste metodiche è stato possibile confermare i dati osservati in sequenza. Dalle sequenze e dalle tecniche utilizzate successivamente si deduce che gli SNPs identificati sono polimorfici nel campione preso in esame e presenti in una buona percentuale di esso.

Con la lettura delle sequenze effettuate su DNA per l'LTA sono stati individuati uno SNP in 5', due nel primo introne e uno nel secondo.

Nella tabella sottostante (tab.6), vengono riportati i risultati ottenuti sia per i genitori di individui affetti da T1D sia per i campioni di SM. Quindi, in totale sono stati individuati 7 SNPs, tutti noti, uno in 5', 3 esonici e tre intronici.

	5' rs1800683	Es 1- rs2239704	1° intr rs909253	1° intr rs746868	Es 2 rs2229094	2° intr rs3093542	Es 3 rs1041981
INDIVIDUI							
2 (T1D)(m)	g/g	c/a	a/a	c/c	t/c	g/g	c/c
3 (T1D) (m)	g/a	c/c	g/a	g/g	t/c	g/g	a/c
4 (T1D) (f)	g/g	c/a	a/a	g/c	t/c	g/g	c/c
5(T1D) (m)	g/a	c/c	g/a	g/g	t/c	g/g	a/c
6 (T1D) (f)	g/a	c/c	g/a	g/g	t/c	g/g	a/c
7(T1D) (f)	g/a	c/a	g/a	g/g	t/t	g/g	a/c
11(T1D)(m)	g/g	c/a	a/a	g/c	t/c	g/g	c/c
12(T1D)(m)	a/a	c/c	g/g	g/g	t/t	g/g	a/a
1 (SM) (f)	g/g	c/c	a/a	g/c	c/c	c/g	c/c
8 (SM) (m)	g/g	c/c	a/a	g/g	c/c	c/g	c/c
9 (SM) (f)	g/a	c/c	a/g	g/g	t/c	c/g	a/c
10(SM) (f)	g/a	c/c	a/g	g/c	c/c	c/g	a/a

Tab. 6 SNPs individuati in sequenza. In azzurro si evidenziano gli SNPs in linkage disequilibrium fra loro.

Tra tutti gli SNPs studiati, oltre a quelli che danno mutazioni aminoacidiche, risulta essere interessante l'rs3093542 in quanto si presenta in eterozigosi soltanto nei campioni SM mentre in tutti gli altri campioni si presenta in omozigosi G/G.

## **b) Allineamento delle sequenze e valutazione dello score**

La scelta degli SNPs da tipizzare si è basata non solo sulle informazioni ricavate dal sequenziamento, ma anche da una attenta valutazione delle informazioni ottenibili dalla rete e riguardanti gli aplotipi PGF, COX e QBL.

La regione genica presa in esame abbraccia oltre al TNF e l'LTA anche i geni e le regioni a monte e a valle di questi. La regione del cromosoma 6p21.33, su cui è stato fatto l'allineamento di sequenze comprende i geni NFKBIL1, LTA, TNF, LTB che si trova nell'intervallo 31,622,969 bp-31,659,000 bp, per un totale di 36032 bp. Dal confronto tra i 3 aplotipi per questa regione di DNA, si deducono una serie di informazioni riguardanti i polimorfismi presenti in questa regione e la rilevanza che potrebbero avere.

Da una visione generale dell'allineamento si ricavano le seguenti informazioni:

- 73 polimorfismi totali (SNPs e indels), di cui:
- 32 sono in zone intergeniche
- 12 con alleli identici in PGF e QBL
- 23 con alleli diversi tra COX e QBL (non considerando il DR2)
- 3 ricadono negli esoni e danno sostituzioni amminoacidiche

Considerando invece gli SNPs con alleli identici in PGF e QBL e che differiscono da COX si traggono le seguenti informazioni:

15→SNPs totali

8→ nell' NFKBIL1

2→ nella zona intergenica tra NFKBIL1 e LTA

3→ nell'LTA (piu' altri 4 noti, trovati in sequenza, ma che non rispettano la sopra citata caratteristica)

1→ nel TNF

1→ nella zona intergenica tra TNF e LTB

Infine, esaminando gli SNPs con alleli differenti tra QBL e COX e con alleli identici tra PGF e COX si deducono le sotto indicate informazioni:

23→ SNPs in totale

6→ nell'NFKBIL1

8→ nella zona intergenica tra NFKBIL1 e LTA

4→ nell'LTA

4→ nel TNF

Il calcolo dello score di rilevanza ha posto in evidenza i polimorfismi che, per le loro caratteristiche, sono i più importanti nella regione di classe III studiata. Tra tutti i polimorfismi esaminati, solamente 2 presentano uno score interessante: rispettivamente di 7 punti per il rs2229094 e di 9 punti per il rs1041981.

Nelle tabelle 7 e 8 sono riportate le informazioni ricavate riguardanti questi SNPs.

LOCALIZZAZIONE	2° esone del gene LTA chr6:31,648,535			
CONSERVAZIONE (score)	0			
SOSTITUZIONE AMMINOACIDICA	T=cisteina C=Arginina			
PROMOTORE	-			
SPLICING	BDGP		C	-
			T	-
	Splice view		C	-
			T	-
	Netgene2		C	Attivazione di un sito accettore confidence score 0.07
			T	-

Tab. 7 caratteristiche SNP rs2229094 C/T

LOCALIZZAZIONE	3° esone del gene LTA__chr6: 31,648,763			
CONSERVAZIONE (score)	0.850394			
SOSTITUZIONE AMMINOACIDICA	C = Treonina A = Asparagina			
PROMOTORE	-			
SPLICING	BDGP		C	-
			A	-
	Splice view		C	-
			A	-
	Netgene2		C	-
			A	-

Tab 8. caratteristiche SNP rs1041981 C/A

Di seguito vengono riportati i punti ottenuti dai due SNPs per ogni criterio considerato e il totale dello score di rilevanza (tab 9).

	rs2229094	rs1041981
PGF=QBL		1
QBL≠COX	1	1
LOCALIZZAZIONE	2	2
CONSERVAZIONE		2
SPLICING	1	3
SOST.		
AMMINOACIDICA	3	
PROMOTER		
SCORE TOTALE	7	9

Tab. 9 punteggi relativi agli SNPs rs2229094 e rs1041981

Gli SNPs rs1041981 e rs2229094, inoltre, sono stati trovati nella fase di sequenziamento sia sui genitori di diabetici che sugli affetti da sclerosi multipla. Data la loro rilevanza emersa dall'allineamento e data la loro certa presenza nella popolazione Sarda accertata dal sequenziamento, si è ritenuto essenziale tipizzare questi polimorfismi su grande scala utilizzando una cospicua casistica sia per il diabete di tipo 1 sia per la sclerosi multipla.

Lo SNP rs3093542, presente nel 2° introne dell'LTA, non rientra nello studio di valutazione dello score in quanto dalle informazioni presenti nei browser si presenta con lo stesso allele in tutti e tre gli aplotipi. Quest'ultima informazione ha messo in evidenza l'estrema importanza che questo SNP può avere poiché, come spiegato precedentemente, in sequenza si presentava in eterozigosi solo negli individui affetti da sclerosi multipla e in omozigosi in tutti gli altri.

Dunque, gli SNPs più rilevanti per la regione in esame sono: rs2229094, rs1041981 e rs3093542.

Per il suo alto score è stata data la priorità nelle tipizzazioni allo SNP rs1041981, trovato in sequenza e con score pari a 9.

### c) **Trasmission disequilibrium test**

Lo SNP rs1041981 è stato tipizzato sulla vasta casistica precedentemente descritta. I dati ottenuti dalla genotipizzazione sono stati analizzati tramite transmission disequilibrium test (TDT). Per il test ci si è rifatti all'utilizzo del software online: <http://bioinformatics.holstegelab.nl/home/index.php>.

I valori di P-value ottenuti sono di seguito riportati (tab 9).

	<b>p-value</b>
Diabete di tipo 1	0.0612462
Sclerosi multipla	0.321277

Tab. 9 valori di p-value

Essendo la soglia di significatività inferiore a 0.5, i dati ottenuti indicano che lo SNP rs1041981 non è associato con le malattie prese in esame.

#### **d) Conclusioni**

I dati ottenuti dal TDT per il polimorfismo sopra descritto non escludono né affermano l'associazione del gene LTA con le patologie oggetto di studio. Il risultato ottenuto, infatti, può solo appurare che l'rs1041981 e la mutazione che esso comporta non ha un ruolo nella comparsa delle patologie. Tuttavia, per poter escludere o provare l'associazione dell'LTA con T1D e SM occorre, sicuramente, tenere in considerazione gli SNPs rs2229094 e rs3093542. La tipizzazione di questi SNPs è tutt'ora in corso e in base ad essa si potrebbe avere un quadro certo riguardante l'associazione di LTA. Inoltre le informazioni che potrebbero dare questi due SNPs non sarebbero assolutamente ridondanti in quanto non si trovano in linkage disequilibrium con rs1041981 né tra di loro. Sicuramente, quindi, il prossimo passo del progetto di ricerca sarà dato dalla tipizzazione di questi ultimi due SNPs su larga scala.

Per quanto riguarda il TNF, su questo gene non sono stati trovati polimorfismi in sequenza e la valutazione dello score di rilevanza non ha portato ad evidenziare nessun SNP rispetto ad altri. Tuttavia, i presupposti che indicano questo gene come un possibile candidato sono estremamente forti e convincenti. La strategia adottabile per capire il ruolo del TNF è data dalla scelta di tagSNPs presenti nel gene o vicino ad esso e che, una volta tipizzati, chiarirebbero l'azione di questo gene. La scelta di questi polimorfismi è in corso e si basa sui dati presenti nei

database <http://www.broad.mit.edu/mpg/tagger/> e  
<http://gvs.gs.washington.edu>.

I futuri sviluppi del progetto di ricerca sono dati dalla tipizzazione degli SNPs rs2229094 e rs3093542 e dalla individuazione, tramite appositi database, di tagSNPs significativi che possano chiarire il ruolo del TNF. Inoltre, non è possibile escludere a priori un ruolo dei geni a monte e a valle del TNF nelle patologie in esame. Per appurare questa affermazione, è sempre possibile la scelta di SNPs in base ai criteri dello score di rilevanza ed in base ai tagSNPs presenti nella regione.

Il progetto è, dunque, in fase di sviluppo e progredisce perseguendo il fine ultimo di chiarire quali siano i geni implicati nelle due patologie che più colpiscono la nostra isola: il diabete di tipo 1 e la sclerosi multipla.



## **Riassunto**

La linea di ricerca di questo progetto è basata sullo studio di due patologie multifattoriali: la sclerosi multipla (SM) e il diabete di tipo 1 (T1D), fortemente prevalenti in Sardegna ed entrambe autoimmuni.

Lo scopo del progetto è la ricerca di varianti genetiche che influenzano il rischio di ammalare di sclerosi multipla e diabete di tipo 1, partendo dal presupposto dell'esistenza di geni comuni che predispongono per entrambe le malattie.

Come documentato in letteratura, il principale locus genetico di suscettibilità alla sclerosi multipla e al diabete di tipo 1 è localizzato a livello della regione HLA.

Due lavori scientifici di rilievo (Zavattari et al. 2001, Marrosu et al. 2001) hanno indirizzato questo progetto di ricerca verso la classe III dell'HLA ed, in particolar modo su due geni all'interno di essa l' LTA e il TNF. Questi due geni appartengono entrambi alla famiglia del tumor necrosis factor e codificano per delle citochine.

Il lavoro si è svolto in più steps. Nel primo step è stata fatta una selezione della casistica utilizzabile per poter sequenziare i geni di interesse. La selezione si è basata sulle caratteristiche aplotipiche ed ha portato alla scelta di 8 genitori di affetti da T1D e di 4 individui affetti da SM.

Lo step successivo è stato quello di sequenziare i geni di interesse su cDNA e DNA. In questa fase del progetto sono stati individuati 7 SNPs nel gene LTA; nel gene TNF non è stato trovato alcun polimorfismo nelle regioni sequenziate.

Il terzo step del lavoro è stato quello della selezione della casistica necessaria per tipizzare gli SNPs di interesse su larga scala. Sono quindi state selezionate 396 famiglie per la SM e 832 per il T1D; tutte le famiglie erano costituite da entrambi i genitori, un figlio affetto e, dove presenti, i fratelli non affetti.

Inoltre è stato fatto uno studio su tutti gli SNPs presenti non solo nei geni di interesse, ma anche nei geni e nelle regioni intergeniche a monte e a valle di questi. Tale studio ha preso in esame tramite differenti criteri di selezione ogni singolo SNP al fine di poter individuare i più importanti presenti nella regione basandosi su uno score di rilevanza.

Il quinto step è stato dato dalla tipizzazione dello SNP rs1041981 che tra tutti quelli esaminati rivestiva la maggiore importanza sia dai risultati ottenuti in sequenza sia dallo studio basato sullo score di rilevanza.

Infine è stata effettuata l'analisi statistica tramite il transmission disequilibrium test che ha escluso un ruolo dell'rs1041981 sul diabete e sulla sclerosi multipla, in quanto è stato ottenuto per entrambe le malattie un P-value superiore a 0.5. Il risultato ottenuto non esclude un ruolo dell'LTA, ruolo che deve essere appurato tramite la tipizzazione di altri SNPs di interesse individuati e la cui genotipizzazione è in corso.

I presupposti che indicano il TNF e l'LTA come geni candidati sono assolutamente forti e convincenti e spingono il progetto di ricerca verso la chiarificazione del ruolo di questi geni nelle malattie esaminate. Le prospettive future sono quindi date dalla genotipizzazione su larga scala di altri SNPs rilevanti presenti nei geni di interesse e a monte e a valle di essi.



## **Bibliografia**

Altman, D.M. & Trowsdale, J. Major histocompatibility complex structure and function. *Curr Opin Immunol* 2, 93-98 (1989).

Analysis of a non-functional HNF-1-alpha (TCF1) mutation in Japanese subjects with familial type 1 diabetes. *Hum. Mutat.* 18: 345-351, 2001

Barcellos, L. F.; Oksenberg, J. R.; Begovich, A. B.; Martin, E. R.; Schmidt, S.; Vittinghoff, E.; Goodin, D. S.; Pelletier, D.; Lincoln, R. R.; Bucher, P.; Swerdlin, A.; Pericak-Vance, M. A.; Haines, J. L.; Hauser, S. L. :

Bergholdt, R.; Nerup, J.; Pociot, F. : Fine mapping of a region on chromosome 21q21.11-q22.3 showing linkage to type 1 diabetes. *J. Med. Genet.* 42: 17-25, 2005.

Bottini et al.: A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nature Genet.* 36: 337-338, 2004

Boyman, O.; Hefti, H. P.; Conrad, C.; Nickoloff, B. J.; Suter, M.; Nestle, F. O. : Spontaneous development of psoriasis in a new animal model shows an essential role for resident T cells and tumor necrosis factor-alpha. *J. Exp. Med.* 199: 731-736, 2004.

Concannon, P.; Gogolin-Ewens, K. J.; Hinds, D. A.; Wapelhorst, B.; Morrison, V. A.; Stirling, B.; Mitra, M.; Farmer, J.; Williams, S. R.; Cox, N. J.; Bell, G. I.; Risch, N.; Spielman, R. S. : A second-generation screen of the human genome for susceptibility to insulin-dependent diabetes mellitus. *Nature Genet.* 19: 292-296, 1998.

Cook et al. 1998 Declining incidence of multiple sclerosis in Shetland Islands: an update. *Acta Neurol Scand* 77, 148-151.

Corradu, F.; Sawcer, S.; D'Alfonso, S.; Lai, M.; Hensiek, A.; Solla, E.; Broadley, S.; Mancosu, C.; Pugliatti, M.; Marrosu, M. G.; Compston, A. : A genome screen for multiple sclerosis in Sardinian multiplex families. *Europ. J. Hum. Genet.* 9: 621-626, 2001.

Cucca et al. The distribution of DR4 haplotypes in Sardinia suggest a primary association of insulin dependent diabetes mellitus with DRB1 and DQB1 loci. *Hum. Immunol* 43, 301-308 (1995)

Davies, J. L.; Kawaguchi, Y.; Bennett, S. T.; Copeman, J. B.; Cordell, H. J.; Pritchard, L. E.; Reed, P. W.; Gough, S. C. L.; Jenkins, S. C.; Palmer, S. M.; Balfour, K. M.; Rowe, B. R.; Farrall, M.; Barnett, A. H.; Bain, S. C.; Todd, J. A. : A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 371: 130-136, 1994.

Dyment, D. A.; Sadnovich, A. D.; Ebers, G. C. : Genetics of multiple sclerosis. *Hum. Molec. Genet.* 6: 1693-1698, 1997.

Dyment, D. A.; Willer, C. J.; Scott, B.; Armstrong, H.; Ligers, A.; Hillert, J.; Paty, D. W.; Hashimoto, S.; Devonshire, V.; Hooge, J.; Kastrukoff, L.; Oger, J.; and 24 others : Genetic susceptibility to MS: a second stage analysis in Canadian MS families. *Neurogenetics* 3: 145-151, 2001.

Ebers, G. C. : Personal Communication. London, Ontario, Canada, 8/29/1988.

Ebers, G. C.; Kukay, K.; Bulman, D. E.; Sadovnick, A. D.; Rice, G.; Anderson, C.; Armstrong, H.; Cousin, K.; Bell, R. B.; Hader, W.; Paty, D. W.; Hashimoto, S.; Oger, J.; Duquette, P.; Warren, S.; Gray, T.; O'Connor, P.; Nath, A.; Auty, A.; Metz, L.; and 10 others : A full genome search in multiple sclerosis. *Nature Genet.* 13: 472-476, 1996.

Ebers, G. C.; Sadovnick, A. D.; Risch, N. J.; Canadian Collaborative Study Group : A genetic basis for familial aggregation in multiple sclerosis. *Nature* 377: 150-151, 1995

European Consortium for IDDM Genome Studies: A genomewide scan for type 1-diabetes susceptibility in Scandinavian families: identification of new loci with evidence of interactions. *Am. J. Hum. Genet.* 69: 1301-1313, 2001.

Fernandez-Arquero, M.; Arroyo, R.; Rubio, A.; Martin, C.; Vigil, P.; Conejero, L.; Figueredo, M. A.; de la Concha, E. G. : Primary association of a TNF gene polymorphism with susceptibility to multiple sclerosis. *Neurology* 53: 1361-1363, 1999.

Gray, P. W.; Aggarwal, B. B.; Benton, C. V.; Bringman, T. S.; Henzel, W. J.; Jarrett, J. A.; Leung, D. W.; Moffat, B.; Ng, P.; Svedersky, L. P.; Palladino, M. A.; Nedwin, G. E. : Cloning and expression of cDNA for human lymphotoxin, a lymphokine with tumour necrosis activity. *Nature* 312: 721-724, 1984.

Green et al. Incidence of childhood-onset insulin-dependent diabetes mellitus: the EURODIAB ACE study. *Lancet* 1992:339-905

Hauser and Goodkin 1998: Multiple sclerosis and other demyelinating diseases. In: Fauci A., et al. (Eds), Harrison's Principle of internal medicine 14<sup>th</sup> ed. Mc Graw Hill, New York, pp. 2409-2418.

Herr et al. Evaluation of fine mapping strategies for a multifactorial disease locus: systematic linkage and association analysis of IDDM in the HLA region chromosome 6p21. Hum Mol Genet 9, 1291- 1301 (2000).

HLA-DR2 dose effect on susceptibility to multiple sclerosis and influence on disease course. Am. J. Hum. Genet. 72: 710-716, 2003.

Junliang et al.: TNF is a potent anti-inflammatory cytokine in autoimmune-mediated demyelination. Nature medicine vol.4 no.1;78-83 1998

Kawasaki et al.: Japanese Study Group on Type 1 Diabetes Genetics : Systematic search for single nucleotide polymorphisms in a lymphoid tyrosine phosphatase gene (PTPN22): association between a promoter polymorphism and type 1 diabetes in Asian populations. Am. J. Med. Genet. 140A: 586-593, 2006.

Kinnunen, E.; Koskenvuo, M.; Kapprio, J.; Aho, K. :

Kurtzke JF: The epidemiology of multiple sclerosis in Raine CS, McFarland HF, Tourtellotte WW(eds): Multiple sclerosis: Clinical and Pathogenetic Basis. London Chapman &Hall Medical 1997 pp91-139.

Kyvik, K.O., Green, A. & Beck-Nielsen, H. Concordance rates of insulin dependent diabetes mellitus: a population based study of young Danish twins. Br Med J 311, 913-917 (1995).

Ligers, A.; Dyment, D. A.; Willer, C. J.; Sadovnick, A. D.; Ebers, G.; Risch, N.; Hillert, J.; Canadian Collaborative Study Groups : Evidence of linkage with HLA-DR in DRB1\*15-negative families with multiple sclerosis. Am. J. Hum. Genet. 69: 9903, 2001.

Marrosu et al.: Dissection of the HLA association with multiple sclerosis in the founder isolated population of Sardinia. Hum mol Gen vol.10,No.25 2907-2916, 2001.

Marrosu MG et al Lancet 2002 ; 359: 1461-1465

Multiple sclerosis in a nationwide series of twins. Neurology 37: 1627-1629, 1986

Mumford, C. J.; Wood, N. W.; Kellar-Wood, H.; Thorpe, J. W.; Miller, D. H.; Compston, D. A. S. : The British

Isles survey of multiple sclerosis in twins. *Neurology* 44: 11-15, 1994.

Muntoni et al. High incidence rate of IDDM in Sardinia. *Diabetes Care* 1992; 15:1317

Nancy J. et al.: New Drugs for Rheumatoid Arthritis. *NEngl J Med* 350; 21. 2004.

Nedwin, G. E.; Naylor, S. L.; Sakaguchi, A. Y.; Smith, D.; Jarrett-Nedwin, J.; Pennica, D.; Goeddel, D. V.; Gray, P. W. : Human lymphotoxin and tumor necrosis factor genes: structure, homology and chromosomal localization. *Nucleic Acids Res.* 13: 6361-6373, 1985.

Oksenberg, J. R.; Barcellos, L. F.; Cree, B. A. C.; Baranzini, S. E.; Bugawan, T. L.; Khan, O.; Lincoln, R. R.; Swerdlin, A.; Mignot, E.; Lin, L.; Goodin, D.; Erlich, H. A.; Schmidt, S.; Thomson, G.; Reich, D. E.; Pericak-Vance, M. A.; Haines, J. L.; Hauser, S. L. : Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. *Am. J. Hum. Genet.* 74: 160-167, 2004.

Pennica, D.; Nedwin, G. E.; Hayflick, J. S.; Seeburg, P. H.; Derynck, R.; Palladino, M. A.; Kohr, W. J.; Aggarwal, B. B.; Goeddel, D. V. : Human tumour

necrosis factor: precursor structure, expression and homology to lymphotoxin. *Nature* 312: 724-729, 1984.

Poser et al New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983; 13:227-231.

PROCARDIS Consortium : A trio family study showing association of the lymphotoxin-alpha N26 (804A) allele with coronary artery disease. *Europ. J. Hum. Genet.* 12: 770-774, 2004.

Pugliatti et al. Increasing incidence of multiple sclerosis in the province of Sassari, Northern Sardinia *Neuroepidemiology* 25: 129-134 2005

Pugliatti et al. Multiple sclerosis distribution in northern Sardinia: spatial cluster analysis of prevalence. *Neurology* 58: 277-282, 2002

Pugliatti et al. Multiple sclerosis epidemiology in Sardinia: evidence for a true increasing risk. *Acta Neurol Scand* 2001;103:20-26).

Ramos, E. M.; Lin, M.-T.; Larson, E. B.; Maezawa, I.; Tseng, L.-H.; Edwards, K. L.; Schellenberg, G. D.; Hansen, J. A.; Kukull, W. A.; Jin, L.-W. : Tumor necrosis factor-alpha and interleukin 10 promoter region

polymorphisms and risk of late-onset Alzheimer disease. Arch. Neurol. 63: 1165-1169, 2006.

Richards RG et al. A review of the natural history and epidemiology of multiple sclerosis: implication for resource allocation and health economic models. Health Technology Assessment 2002; Vol 6; No. 10

Ristori, G.; Cannoni, S.; Stazi, M. A.; Vanacore, N.; Cotichini, R.; Alfo, M.; Pugliatti, M.; Sotgiu, S.; Solaro, C.; Bomprezzi, R.; Di Giovanni, S.; Talamanca, L. F.; and 10 others : Multiple sclerosis in twins from continental Italy and Sardinia: a nationwide study. Ann. Neurol. 59: 27-34, 2006.

Roach et enetic mapping at 3-kilobase resolution reveals inositol 1,4,5-triphosphate receptor 3 as a risk factor for type 1 diabetes in Sweden. Am. J. Hum. Genet. 79: 614-627, 2006.

Rosmond, R.; Chagnon, M.; Bouchard, C.; Bjorntorp, P. : G-308A polymorphism of the tumor necrosis factor alpha gene promoter and salivary cortisol secretion. J. Clin. Endocr. Metab. 86: 2178-2180, 2001.

Ruuls, S. R.; Sedgwick, J. D. : Unlinking tumor necrosis factor biology from the major histocompatibility

complex: lessons from human genetics and animal models. *Am. J. Hum. Genet.* 65: 294-301, 1999.

Sawcer, S.; Jones, H. B.; Feakes, R.; Gray, J.; Smaldon, N.; Chataway, J.; Robertson, N.; Clayton, D.; Goodfellow, P. N.; Compston, A. : A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nature Genet.* 13: 464-468, 1996.

Schaid: Transmission disequilibrium, family controls and genetic expectations, *Am J Hum Genet* 63 935-942. 1998.

Shirai, T.; Yamaguchi, H.; Ito, H.; Todd, C. W.; Wallace, R. B. : Cloning and expression in *Escherichia coli* of the gene for human tumour necrosis factor. *Nature* 313: 803-806, 1985.

Stein, C. M.; Nshuti, L.; Chiunda, A. B.; Boom, W. H.; Elston, R. C.; Mugerwa, R. D.; Iyengar, S. K.; Whalen, C. C. : Evidence for a major gene influence on tumor necrosis factor-alpha expression in tuberculosis: path and segregation analysis. *Hum. Hered.* 60: 109-118, 2005.

Tassinari T, Parodi S, Badino R, Vercelli M. Mortality trend for multiple sclerosis in Italy (1974-1993). *Eur J Epidemiol* 2001;17(2):105

Tessier, M.-C.; Qu, H.-Q.; Frechette, R.; Bacot, F.; Grabs, R.; Taback, S. P.; Lawson, M. L.; Kirsch, S. E.; Hudson, T. J.; Polychronakos, C. : Type 1 diabetes and the OAS gene cluster: association with splicing polymorphism or haplotype? *J. Clin. Genet.* 43: 129-132, 2006.

Tisch, R.; McDevitt, H. : Insulin-dependent diabetes mellitus. *Cell* 85: 291-297, 1996.

Todd, J. A. : Genetic control of autoimmunity in type 1 diabetes. *Immun. Today* 11: 122-129, 1990.

Vyse, T. J.; Todd, J. A. : Genetic analysis of autoimmune disease. *Cell* 85: 311-318, 1996.

Willer, C. J.; Dyment, D. A.; Risch, N. J.; Sadovnick, A. D.; Ebers, G. C.; The Canadian Collaborative Study Group : Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc. Nat. Acad. Sci.* 100: 12877-00-12882, 2003.

Yu et al. 1989 Multiple sclerosis amongst Chinese in Hong Kong. *Brain* 112, 1445-1467.

Zavattari et al.: Conditional linkage disequilibrium analysis of a complex disease super locus, IDDM 1 in the HLA region, reveals the presence of independent

modifying gene effects influencing the type 1 diabetes risk encoded by the major HLA-DQB1, -DRB1 disease loci. Hum mol gen vol.10, No. 881- 889. 2001.

Zavattari, P. et al. Conditional linkage disequilibrium analysis of a complex disease superlocus, IDDM1 in the HLA region, reveals the presence of independent modifying gene effects influencing the type 1 diabetes risk encoded by the major HLA-DQB1, -DRB1 disease loci. Hum Mol Genet 10, 881-889. (2001).

Zavattari, P. et al. Confirmation of the DRB1-DQB1 loci as the major component of IDDM1 in the isolated founder population of Sardinia. Hum Mol Genet 9, 2967-2972. (2000).

Zipp, F.; Weber, F.; Huber, S.; Sotgiu, S.; Czlonkowska, A.; Holler, E.; Albert, E.; Weiss, E. H.; Wekerle, H.; Hohlfeld, R. : Genetic control of multiple sclerosis: increased production of lymphotoxin and tumor necrosis factor-alpha by HLA-DR2(+) T cells. Ann. Neurol. 38: 723-730, 1995.



## Sitografia

<http://125.itba.mi.cnr.it/~webgene/wwwspliceview.html>

<http://www.cbs.dtu.dk/services/NetGene2/>

[http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)

<http://bioinformatics.holstegelab.nl/home/index.php>

<http://www.broad.mit.edu/mpg/tagger/>

<http://gvs.gs.washington.edu>

<http://www.up.univ-mrs.fr/wabim/english/logligne.html>

<http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>

<http://genome.ucsc.edu/cgi-bin/hgGateway>

<http://tools.neb.com/NEBcutter2/cutshow.php?name=7b1e36d9->

<http://www.basic.northwestern.edu/biotools/oligoalc.html>

[http://fokker.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://fokker.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)

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

<http://watcut.uwaterloo.ca/watcut/watcut/template.php?act=help#snp>

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

<http://gvs.gs.washington.edu/GVS/PopStatsServlet?searchBy=Gene+Hugo&target=LTA&upstreamSize=0&downstreamSize=0&x=&y=>

<http://snpanalyzer.utmem.edu/>