

Genetics of the immune cell receptors TCRB and CCR5 in human disease

by Marc McWilliams Buhler

A thesis submitted to the University of Sydney in fulfillment of the requirements for the degree of Doctor of Philosophy

Institute of Immunology and Allergy Research,
Westmead Millennium Institute, Westmead Hospital
and the University of Sydney, New South Wales

April 2003

DECLARATION

Declaration

I hereby declare that the research data presented in this thesis is the result of original research conducted by myself, except where otherwise acknowledged. This thesis has not previously been submitted for a degree at this or any other university.

Marc McWilliams Buhler

Acknowledgements

I would like to thank my supervisor, Associate Professor Graeme J. Stewart, and my co-supervisor, Dr. Bruce H. Bennetts, for their constant support and encouragement. I thank as well Dr. Graham Mann for many useful discussions and much advice and Professor David Harris for his support. Thanks as well to Dr. Pat Concannon for DNA probes and advice on PCR of the TCRB alleles, to Professors Claude Bernard and Ian Mackay for RA samples, and to Dr. Judith Greer for MS, OND and control DNA samples. Further thanks to Drs. Maria Craig and Kim C. Donaghue and Professor Martin Silink of the New Children's Hospital at Westmead, to Dr. Jinny Willis and Professors Russell Scott and Brian Darlow of the Lipid and Diabetes Research Group at Christchurch Hospital, New Zealand, and to Prof. Brian D. Tait of the Victorian Transplant and Immunogenetics Service at the Royal Melbourne Hospital for type 1 diabetes samples.

Further thanks to Professor Nick Manolios of the Department of Rheumatology for SLE samples and for his support and advice, and thanks to Anné Proos and Viive Howell of the Kolling Institute at RNS Hospital for use of the CCR5 typing data from their caucasian student control group. Many thanks to Prerna Badhwar for persistent effort in helping to CCR5 genotype the type 1 diabetes samples, to Helen Smart for her effort in all things clerical and to many other friends and co-workers for pushing me along. In addition, I thank Joe Huang and Suzy Teutsch for using my TCRB typing data and Dr. Hassan Naif and Prof. Tony Cunningham for including CCR5 and microsatellite typing data of mine from the VaHa family in work which we have published and Dr. Karen Bythe for statistical advice.

For the study of CCR5 and microsatellites in Ashkenazi Jews I thank Professor Leslie Burnett, along with Anné Proos and Viive Howell, for use of samples from their DNA

bank and for many useful discussions and thank Dr. Graham Mann and James Indsto for their help with DNA fragment length analysis. I also thank Professor Montgomery Slatkin of U.C. Berkeley for useful discussions.

I would like to thank my wife, Jenny, and my parents, Roald and Shirrell, both for personal support and encouragement and also for their direct involvement. Jenny Buhler did much of the work on several of the figures in this thesis and much of the data management and a number of statistical calculations were made using the P-Stat software package and advice kindly provided by Roald and Shirrell Buhler at P-Stat, Inc. of Hopewell NJ. My two young sons, Wade and Owen, deserve to be thanked as well just for being there - especially Wade who will happily tell you that “philosophy” is when you “think about ideas”, which I have tried to do in this thesis.

Abstract

Early in the evolution of the vertebrates it is thought that two genomic duplications occurred, providing a basis for the evolution in body plan and neural crest of very early vertebrates and substantive material for further evolution of various gene families such as those making up a number of components of the adaptive vertebrate immune system. While the bony fish possibly had another, genome duplications are not generally a feature of vertebrate evolution and indeed the appearance of an antigen-adaptive immune recognition system may have served to limit the size that various vertebrate genomes, including that of the human, can in fact achieve. This initial step in vertebrate immune evolution, the establishment of recognition of non-self against the unique set of “self” epitopes for an individual, provided an immensely powerful weapon in immune function with the ability to tailor a defense against as-yet-unseen dangers at any time albeit with the pitfall of autoimmune disease. As the recognition sites of the antigen receptor molecules such as TcR are produced by clonal modification of the segments provided in the germline and are thus not in the genome itself, pathogens have not been able to hijack this one component of the immune system in the way so many other components have been put to use throughout evolution, nor do these components necessarily reveal themselves as associated with disease through genome screens. Importantly, overall immune function is determined not just by the potential repertoire of recognition receptors but also by the ability of immunocompetent cells to migrate in a tissue specific fashion through the use of various chemokines and their receptors. Typical of the hijacking of an immune system component by a pathogen is the use of a chemokine ligand gene in the viral ancestor to SIV and HIV, allowing for virus binding to immunocompetent cells as is seen in the use of the CCR5 chemokine receptor by macrophage-tropic HIV strains.

This thesis describes the allele and genotype frequencies for several TcR beta-chain variable segment polymorphisms in a population of MS patients compared with controls

before and after stratification for HLA-DR15, polymorphism in the Apo-1 / Fas promoter, the DRB1 Val⁸⁶/Val⁸⁶ genotype, *CCR5-Δ32* and the HLA-DRA promoter. The thesis continues with *CCR5-Δ32* genotyping in IDDM, MS and SLE cohorts and then examines the question of the population of origin of the delta-32 allele of the CCR5 receptor for chemokine. Here, a case / control comparison of 122 RR-MS patients with 96 normal individuals was made for allele and genotype frequencies and for haplotypes formed by pairs of TCRB markers. Further analysis was made after HLA-DR15 stratification. Linkage disequilibrium was found between pairs of alleles of *bv8s1*, *bv10s1*, *bv15s1* and *bv3s1* loci in both patients and controls. In the RR-MS cohort, an increase in the allele frequency of *bv8s1*2* was seen ($p = 0.03$) and the haplotype *bv8s1*2 / bv3s1*1* was increased ($p = 0.006$), and both were found to be statistically significant. In the DR15-positive group, association between MS and TCRB was seen with the *bv8s1*2* allele ($p = 0.05$) and the *bv8s1*2 / bv10s1* haplotypes ($p = 0.048$), while the haplotype associations seen among the DR15-negative patients included the *bv3s1*1* allele (*bv10s1*1 / bv3s1*1*, $p = 0.022$; *bv8s1*2 / bv3s1*1*, $p = 0.048$). While no associations were found after stratification for SDF1-3'A, Apo-1 / Fas or DRB1 there were modest interactions between *bv3s1*, *bv10s1* and *bv15s1* and the HLA-DRA promoter. These results support the involvement of the TCRB region in MS susceptibility.

The further study of autoimmune disease here includes genotype analysis of *CCR5-Δ32* in type 1 diabetes (IDDM) and SLE. CCR5 is the major co-receptor for viral entry used by macrophage-tropic HIV strains and protection from infection is seen in homozygotes for *CCR5-Δ32*. In diabetes, infiltration of pancreatic tissue by autoreactive T-cells involves secretion of multiple cytokines and chemokine receptor expression. Variation in the chemokine receptor CCR5 may result in differences in inflammatory cell migration in response to relevant chemokines. Adolescents with type 1 diabetes were genotyped for *CCR5-Δ32* ($n = 626$). The allele frequency was compared with that of

253 non-diabetic adolescents and with that of 92 adults with SLE. A reduced allele frequency was seen in type 1 diabetes compared with controls (0.092 vs 0.123, $p = 0.05$). This difference was not seen for the cohort of patients with SLE (freq = 0.114). A reduction in the number of *CCR5-Δ32/Δ32* homozygotes, who lack CCR5, in the type 1 diabetes cohort was also seen and while not statistically significant (2 observed compared to 5.25 expected; $p = 0.12$) is interesting. These results suggest a partial protection from type 1 diabetes for *CCR5-Δ32* homozygous individuals is possible and that CCR5 has a potential role in the pathogenesis of type 1 diabetes.

Global surveys of the *CCR5-Δ32* allele have confirmed a single mutation event in a Northeastern European population as the source of this allele. Here, Australian Ashkenazi Jews ($n = 807$) were found to have a *CCR5-Δ32* allele frequency of 14.6% while Australian Sephardic Jews ($n = 35$) had a frequency of 5.7% and non-Jewish Australian controls ($n = 311$) had an allele frequency of 11.25%. Data on birthplace of grandparents showed a gradient with highest *CCR5-Δ32* frequencies from Eastern European Ashkenazim (~19.5% for those whose four grandparents come only from Russia, Poland, Hungary, Austria and Czechoslovakia; $n = 197$) which differs significantly from the frequency seen in Ashkenazi Jews from Western Europe ($n = 101$, $p = 0.001$). Homozygotes for *CCR5-Δ32* were genotyped with 3p21 region microsatellites. This has defined an ancestral haplotype on which the mutation first occurred and helped to date this event to between 40 and 50 generations ago or just over a thousand years ago. The population gradient, combined with the dating of the mutation by microsatellite allele frequencies, suggests an origin for the *CCR5-Δ32* allele in a population ancestral to the Ashkenazim. The distribution in non-Jewish populations in northern Europe has led others to postulate spread of the mutation by Vikings. It is hypothesised here that the link between the two populations could be the kingdom of Khazaria with subsequent admixture into both Swedish Vikings and Ashkenazi Jews.

The basic driving force of evolution is through selection and the immune system has a role which, through the survival pressure exerted by viruses and other pathogens, has the potential to exert a great deal of selective force on the various components of this system. The effects of this pronounced selection on an immune system component can be seen for example in the increase of the *CCR5-Δ32* allele over the last thousand years to the current frequency. As mentioned, some immune system components are not affected by such straightforward selection. In the case of the TCRBV segments, effects on the immune repertoire can occur through MHC interaction at the point of thymic entry and in the effects of various superantigens, but the actual binding pockets that recognise antigen are themselves unable to be selected for (or against). The findings presented in this thesis provide support for the association of TCRBV gene segments with multiple sclerosis and also provide support for the further study of the role of the *CCR5-Δ32* allele in type 1 diabetes. Furthermore, data presented here suggests that the *CCR5-Δ32* allele had an origin in the Khazar Kingdom just over a thousand years ago, accounting for the allele frequencies in both the Ashkenazi Jews and in lands frequented by the Vikings. The definition of an extended ancestral haplotype for the *CCR5-Δ32* allele shows how the effect of selection of an allele of one gene can carry with it specific alleles of a large number of other genes as well.

Publications arising from work contained in this thesis

Chapter 3

Buhler, M. McW., BH Bennetts, RNS Heard, GJ Stewart (2000). T cell receptor β chain genotyping in Australian relapsing-remitting multiple sclerosis patients. *Multiple Sclerosis* 6:140-147.

This paper contains the data presented and discussed in chapter 3.

Teutsch, SM, QR Huang, M. Buhler, BH Bennetts, RNS Heard, N Manolios, GJ Stewart (2000). Evaluation of the Apo-1/Fas promoter Mva 1 polymorphism in multiple sclerosis. *Multiple Sclerosis* 6:14-18.

In this paper, the TCRBV genotyping was used for stratification and assessment of potential genetic interactions between the T-cell receptor and the apo-1/Fas gene.

Teutsch, SM, BH Bennetts, MM Buhler, RNS Heard, GJ Stewart (1999). The DRB1 Val86/Val86 genotype associates with multiple sclerosis in Australian patients. *Human Immunology* 60:715-722.

In this paper, the TCRBV genotyping was used for stratification and assessment of potential genetic interactions between the T-cell receptor and DRB1 Val/Gly aa 86 polymorphism.

Bennetts, BH, SM Teutsch, MM Buhler, RNS Heard, GJ Stewart (1999). HLA-DMB gene and HLA-DRA promoter region polymorphisms in Australian multiple sclerosis patients. *Human Immunology* 60:886-893.

In this paper, the TCRBV genotyping was used for stratification and assessment of potential genetic interactions between the T-cell receptor, HLA DMB and HLA-DRA promoter genes; modest interactions were noted between the TCRBV3S1, TCRBV10S1 and TCRBV15S1 loci and HLA-DRA promoter genotypes.

Bennetts, BH, SM Teutsch, MM Buhler, RNS Heard, GJ Stewart (1997). The CCR5 deletion mutation fails to protect against multiple sclerosis. *Human Imm.* 58:52-59.

In this paper, the TCRBV genotyping was used for stratification and assessment of

potential genetic interactions between the T-cell receptor and CCR5-Δ32 mutation.

Chapter 4

Buhler, MM, M. Craig, KC Donaghue, P Badhwar, J Willis, N Manolios, BD Tait, M Silink, BH Bennetts, GJ Stewart (2002). CCR5 genotyping in an Australian and New Zealand Type 1 diabetes cohort. *Autoimmunity* 35:457-461.

This paper contains the data presented in Chapter 4.

Chapter 5

Buhler, M McW, A Proos, V Howell, BH Bennetts, L Burnett, GJ Stewart (submitted). Evidence of an Eastern European Ashkenazi Jewish Enrichment of the CCR5-delta32 allele: a clue to the origin? *This paper contains the data presented in Chapter 5.*

Naif, HM, AL Cunningham, M Alali, S Li, J Nasr, MM Buhler, D Schols, E de Clercq, GJ Stewart (2002). A human immunodeficiency virus type 1 isolate from an infected person homozygous for CCR5-Δ32 exhibits dual-tropism by infecting macrophages and MT2 cells via CXCR4. *Journal of Virology* 76:3114-3124.

This paper, aimed at identifying the co-receptor used in an HIV infected CCR5-Δ32 homozygote, involved the use of the microsatellite assays in the CCR5 region developed and described in Chapter 5 as well as microsatellite markers for CXCR4. The microsatellite data carried out by were critical for the identification of genetic identity between this individual and other family members studied.

Abstract accepted for oral presentation:

Marc M. Buhler, Anné Proos, Viive Howell, Bruce H. Bennetts, Leslie Burnett and Graeme J. Stewart. Could admixture of the CCR5-Δ32 allele into Ashkenazi Jews and Vikings be explained by an origin in the Kingdom of the Khazars?

This abstract was presented as a talk in the symposium on Empirical Population Genetics at the XIX International Congress of Genetics, Melbourne, 2003.

Table of Contents	
Declaration	ii
Acknowledgements	iii
Abstract	v
Publications arising from work contained in this thesis	ix
Table of Contents	xi
List of Figures	xiv
List of Tables	xv
Abbreviations	xvii
Chapter One - Introduction	1
1.1 Introduction	2
1.2 The MHC and the nature of the peptide being presented	3
1.3 Biology of the TCRB locus	6
1.4 Chemokine receptors and the immune response	10
1.5 Immune polymorphism, self polymorphism and effects on repertoire	13
1.6 Linkage disequilibrium, population admixture and selection	20
1.7 Microsatellite polymorphisms and mutation dating	22
1.8 Genomics and repetitive regions of DNA	27
1.9 Evolutionary phylogenetics of specific vertebrate immunity	29
1.10 Duplication versus a limit to “self” in immune evolution	32
1.11 The study of immunogenetics of human disease	33

Chapter Two - Materials and methods	35
2.1 Introduction	36
2.2 Tissue culture and DNA extraction	37
2.3 Ethics approval for subjects and study cohorts	38
2.4 Subjects for TCRB analysis	38
2.5 Subjects for <i>CCR5-Δ32</i> studies	39
2.5.1 Autoimmune diseases - type 1 diabetes, MS and SLE	39
2.5.2 <i>CCR5</i> in the Ashkenazi Jews	39
2.5.3 <i>CCR5</i> and <i>CXCR4</i> in the “VaHa” family	40
2.6 Polyacrylamide gel electrophoresis (PAGE)	40
2.7 RFLP analysis of <i>tcrbv8s1</i> and <i>tcrbv11s1</i>	41
2.8 PCR typing of <i>bv6s5</i> , <i>bv8s1</i> , <i>bv10s1</i> , <i>bv15s1</i> and <i>bv3s1</i>	43
2.9 Statistical analysis	49
2.10 Genotyping for <i>CCR5-Δ32</i>	50
2.11 Methods for the study of microsatellite polymorphisms	52
2.12 <i>CCR5</i> and microsatellite typing in the “VaHa” family	52
2.13 Physical map data of markers near <i>CCR5</i>	60
2.13.1 An analysis of the 3p21 physical map (pre-Build 22)	60
2.13.2 Locating <i>D3S663</i> on the NIH maps	61
2.13.3 Changes to the physical map of 3p21 in Build 22 to Build 24	62
2.13.4 Genetic and physical 3p21 map data compared	63
2.14 Dating of the origin of <i>CCR5-Δ32</i>	63
2.15 Typing for <i>SDF-1</i> in MS and controls	69
Chapter Three - The TCRB locus in MS	70
3.1 Introduction	71
3.2 Results:	78
3.2.1 Allele and genotype frequencies	78
3.2.2 Frequencies for haplotypes defined by two markers	81

3.2.3	Linkage disequilibrium (LD) analysis	81
3.3	Discussion	86
Chapter Four - CCR5 and autoimmune disease		90
4.1	Introduction	91
4.1.1	Type 1 diabetes	93
4.2	Results	95
4.3	Discussion	97
Chapter Five - The question of the origin of the <i>CCR5-Δ32</i> allele		104
5.1	Introduction	105
5.2	Results	111
5.2.1	Population frequencies of the <i>CCR5-Δ32</i> allele	111
5.2.2	The <i>CCR5-Δ32</i> frequency varies with the country of origin	113
5.2.3	Homozygous <i>CCR5-Δ32</i> individuals	118
5.2.4	Identifying the <i>CCR5-Δ32</i> founder haplotype	119
5.2.5	Estimating the age of the <i>CCR5-Δ32</i> mutation	127
5.2.6	Could the <i>CCR5-Δ32</i> founder have been Jewish?	136
5.3	Discussion	137
5.3.1	The Khazar Empire and the role of the Vikings	145
Chapter Six - Discussion		151
6.1	The study of immunogenetics of human disease	152
6.2	<i>CCR5-Δ32</i> polymorphisms in autoimmune disease	155
6.3	The origin of the <i>CCR5-Δ32</i> allele	156
6.4	Selection and the <i>CCR5-Δ32</i> allele	167
6.5	And what of other 3p21 region genes?	172
6.6	Future directions	174
References		177

List of Figures

<u>Figure:</u>	<u>page</u>
1 a-e Hypothetical alleles, haplotypes and mutation origin	24-26
2.1 RFLP autoradiograph of genomic DNA probed for V-beta 8.1	42
2.2 RFLP autoradiograph of genomic DNA probed for V-beta 11	42
2.3 PAGE (10% mini-gel) of bv6s5 PCR product digested with <i>Bst</i> N1	44
2.4a 10% mini-PAGE of bv8s1 PCR product digested with <i>Bam</i> HI	45
2.4b bv10s1 PCR product digested (<i>Hae</i> III) - 10% mini-PAGE gel	45
2.5 bv3s1 pcr product digested with <i>Pvu</i> II - 10% mini-PAGE gel	47
2.6 bv15s1 product digested with <i>Pvu</i> II - 10% mini-PAGE gel	48
2.7 <i>CCR5-Δ32</i> allele genotyping by 10% PAGE of PCR products	
a) Initial gel run of pcr products	50
b) 3 rd gel of <i>CCR5</i> pcr products	51
2.8 a) 3p21 markers & VaHa Family haplotypes	54
b) CXCR4 markers	55
2.9 NCBI "Build 22" map of 3p21 genes and markers	56-58
2.10 Microsatellite PCR product fragment length analysis	
a) D3S3559	65
b) Fragment length analysis of D3S663 product	66
c) Fragment length analysis of microsatellite D3S4580	67
3.1 Location of TCRB segments and haplotypes at 7q35	77
3.2 TCRBV allele and haplotype frequencies in RR-MS	80
5.1 (a & b) <i>CCR5</i> and selected markers in Builds 28 & Build 31	130-131

List of Tables

<u>Table</u>	<u>page</u>
1.1	81 possible 4-marker biallelic phenotypes 29
2.1	Primers for microsatellites used in the 3p21 region 59
2.2	Decay of haplotype frequency over genetic distance 68
2.3	Frequency of SDF1-3'A in Australian MS and controls 69
3.1	Frequency for TCRB alleles and genotypes in RR-MS and Controls 79
3.2	Frequencies for TCRB haplotypes bv8s1/bv3s1, bv8s1/bv10s1, bv10s1/bv3s1, bv6s5/bv8s1 and bv6s5/bv3s1 in Australian RR-MS and controls and the DR15 positive and negative subsets 83-85
4.1	Genotype and allele frequencies of <i>CCR5-Δ32</i> for 620 Australian and New Zealand type 1 diabetes patients and 253 adolescent controls 96
4.2	Allele frequencies compared with frequency of reported homozygous <i>CCR5-Δ32/Δ32</i> individuals for type 1 diabetes, MS, RA and SLE 96
5.1	<i>CCR5-Δ32</i> genotype and allele frequency for Australian Jewish and non-Jewish subgroups 113
5.2	<i>CCR5-Δ32</i> genotype and allele frequency compared with birthplace of maternal grandmother (MatGM) for Australian Ashkenazi Jews 115
5.3	Country of reported grandparents' birth for Ashkenazi Jews and <i>CCR5-Δ32</i> allele frequency for individuals whose grandparent's birthplace is indicated 116
5.4	Country of mother's birth and <i>CCR5</i> frequency for Ashkenazi Jews compared with published allele frequencies for general (non-Jewish) populations 117
5.5	Allele frequencies for microsatellite marker D3S4579 in <i>CCR5</i> wild-type Ashkenazi Jews and Australians, <i>CCR5-Δ32</i> homozygotes, and in heterozygotes 121
5.6	Allele frequencies for microsatellite marker D3S4580 in <i>CCR5</i> wild-type Ashkenazi Jews and Australians, <i>CCR5-Δ32</i> homozygotes, and in heterozygotes 122

5.7	Microsatellite marker “afmb362wb9” allele frequencies in Ashkenazi Jews typed for CCR5 and in non-Jewish heterozygotes	123
5.8	Microsatellite marker “gaat12d11” allele frequencies in Ashkenazi Jews, Ashkenazi heterozygotes, <i>CCR5-Δ32</i> homozygotes and non-Jewish heterozygotes	123
5.9	Microsatellite D3S3559 frequency for CEPH (GDB) data and for the Ashkenazi Jewish homozygote <i>CCR5 Δ32/Δ32</i> cohort	124
5.10	Allele frequencies and number of alleles for the D3S663 microsatellite marker in CCR5 wild-type Ashkenazi Jews and Australians, <i>CCR5-Δ32</i> homozygotes, and in heterozygotes	125
5.11	Allele frequencies and number of alleles for the D3S1578 microsatellite marker in CCR5 wild-type Ashkenazi Jews and Australians, <i>CCR5-Δ32</i> homozygotes, and in heterozygotes	126
5.12	Number of generations required to have various microsatellite alleles on the suggested ancestral haplotype reach the observed values	135
5.13	Possible ages of a <i>CCR5Δ32</i> -D3S663-D3S1578 haplotype	135
5.14	Elevated <i>CCR5-Δ32</i> frequency given in various published cohorts	138
6.1	List of selected genes included in the 3p21 region between afmb362wb9 and D3S1578	170-171

Abbreviations

$\Delta 32$	32 basepair deletion (delta 32) mutation of CCR5
ag	antigen
AIDS	acquired immune deficiency syndrome
APC	antigen presenting cell
blast	basic local alignment search tool
blat	basic local alignment tool
bp	basepair
contig	“contiguous” joining of adjacent sequences of DNA
cM	centimorgan
CP-MS	chronic-progressive MS
CNS	central nervous system
DC	dendritic cell
IDDM	type 1 diabetes (prev. insulin dependent diabetes mellitus)
EBV	Epstein-Barr virus
FCS	fetal calf serum
HCMV	human cytomegalovirus
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
Ig	immunoglobulin
Kbp	kilobasepair (1,000 bp)
LCL	lymphoblastoid cell line
LD	linkage disequilibrium
Mbp	mega-basepair (one million bp)
MCMV	murine cytomegalovirus
MHC	major histocompatibility region
MIP-1 α	macrophage inflammatory protein-1 α
MS	multiple sclerosis
OND	other neurological disease
PAGE	polyacrylamide gel electrophoresis
PBMC	peripheral blood mononuclear cell

PCR	polymerase chain reaction
RA	rheumatoid arthritis
RR-MS	relapsing-remitting multiple sclerosis
RP-MS	relapsing / progressive MS (secondary-progressive MS)
T1D	type 1 diabetes
TcR	T-cell antigen receptor
TCRB	T-cell receptor (beta) variable segment
TCRBV	T-cell receptor (beta) variable segment
Th1	helper T-cell type 1
Th2	helper T-cell type 2
SLE	systemic lupus erythematosus