Infection, Genetics and Evolution 9 (2009) 574–580

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

A 475 years-old founder effect involving *IL12RB1*: A highly prevalent mutation conferring Mendelian Susceptibility to Mycobacterial Diseases in European descendants

J. Yancoski [a], C. Rocco [a], A. Bernasconi [a], M. Oleastro [a], L. Bezrodnik [b], C. Vrátnica [c], F. Haerynck [d], S.D. Rosenzweig [a,*]

[a] Hospital Nacional de Pediatría J. P. Garrahan, Buenos Aires, Argentina
[b] Hospital de Niños Ricardo Gutierrez, Buenos Aires, Argentina
[c] Hospital de Niños Juan Pablo II, Corrientes, Argentina
[d] Ghent University Hospital, Ghent, Belgium

ARTICLE INFO

Article history:
Received 11 December 2008
Received in revised form 13 February 2009
Accepted 16 February 2009
Available online 9 March 2009

Keywords:
Hot spot
Salmonella
Bacillus Calmette Guerin
Interferon gamma
Linkage disequilibrium

ABSTRACT

Mutations in *IFNGR1*, *IFNGR2*, *IL12RB1*, *IL12B*, *STAT1* and *NEMO* result in a common clinical phenotype known as *Mendelian Susceptibility to Mycobacterial Diseases* (MSMD). Interleukin-12 receptor β1 (IL-12Rβ1) deficiency is the most common genetic etiology for MSMD. Known mutations affecting *IL12RB1* are recessively inherited and are associated with null response to both IL-12 and IL-23. Mutation *IL12RB1* 1623_1624delinsTT was originally described in 5 families from European origin (2 from Germany; 1 from Cyprus, France and Belgium). Interestingly, this same mutation was found in an unexpectedly high prevalence among IL-12Rβ1 deficient patients in Argentina: 5-out-of-6 individuals born to unrelated families carried this particular change. To determine whether mutation 1623_1624delinsTT represents a DNA mutational hotspot or a founder effect, 34 polymorphic markers internal or proximal to *IL12RB1* were studied in the Argentinean and the Belgian patients. A common haplotype spanning 1.45–3.51 Mb was shared by all chromosomes carrying mutation 1623_1624delinsTT, and was not detected on 100 control chromosomes. Applying a modified likelihood-based method the age of the most recent common ancestor carrying mutation 1623_1624delinsTT was estimated in 475 years (95% CI, 175–1275), which is the time when the Spaniards initiated the colonization of the Americas. Mutation 1623_1624delinsTT represents the first founder effect described on IL-12Rβ1, the most frequently affected gene in MSMD, and affecting patients with European ancestors. The reason(s) behind the persistency of this mutation across multiple generations, its relative high prevalence, and any potential selective advantage are yet to be established.

Published by Elsevier B.V.

1. Introduction

With more than 100 affected patients worldwide, IL-12Rβ1 deficiency is the most common genetic etiology for *Mendelian susceptibility to Mycobacterial Disease* (MSMD). Interleukin-12 receptorβ1 deficiency causes an illness characterized by a selective susceptibility to poorly pathogenic mycobacteria, and nontyphoid *Salmonella* species attesting to the importance of IL-12R in the host defense against intracellular pathogens (Altare et al., 1998; de Jong et al., 1998; Fieschi et al., 2003; van de Vosse and Ottenhoff, 2006; Filipe-Santos et al., 2006; Rosenzweig et al., 2006).

Mutations on the *IL12RB1* gene are associated with impaired response to IL-12 and IL-23 (Altare et al., 1998; de Jong et al., 1998; Hoeve et al., 2003). With only one exception, all IL-12Rβ1 deficient patients described displayed no detectable IL-12Rβ1 on their cell surface due to mutations that either interrupt the open reading frame (nonsense and frameshift mutations) or disrupt folding of the protein (missense mutations) (Altare et al., 1998; de Jong et al., 1998; Fieschi et al., 2003; Fieschi et al., 2004; van de Vosse and Ottenhoff, 2006; Filipe-Santos et al., 2006). Complex mutation 1623_1624delinsTT (GC deletion, TT insertion) results in a silent point mutation (V541 V) and the generation of a stop codon (Q542X) in *IL12RB*. This particular change had been previously described in 5 families from European origin (2 from Germany; and 1 from Cyprus, France and Belgium) (Fieschi et al., 2003; Haerynck et al., 2008). Unexpectedly, mutation 1623_1624delinsTT was found in 5-out-of-6 Argentinean IL-12Rβ1 deficient patients (3 homozygous and 2 heterozygous) from 6 unrelated families.

* Corresponding author at: Servicio de Inmunología, Hospital Nacional de Pediatría J. P. Garrahan, Combate de los Pozos 1881, Cdad. de Buenos Aires (1245), Argentina. Tel.: +54 9 11 5408 2641; fax: +54 11 4308 4076.
E-mail address: srosenzweig@garrahan.gov.ar (S.D. Rosenzweig).

1567-1348/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.meegid.2009.02.010

The finding of identical mutations in apparently unrelated patients from different regions in the planet suggested that recurrent mutations could arise (*mutational hot spot*). Alternatively, it was possible that despite the geographic separation the families shared a common ancestry (*mutational founder effect*). Herein, these two alternative hypotheses are explored.

2. Patients, materials and methods

2.1. Patient cohort

Six patients with complete IL-12Rβ1 deficiency and born to 6 unrelated non-consanguineous Argentinean families were referred to the Hospital Nacional de Pediatría J. P. Garrahan in Buenos Aires, Argentina and evaluated for this study. All the patients had been assessed for IFNγ/IL-12/IL-23 pathway integrity after they developed disseminated mycobacterial infections early on life. All the patients developed disseminated *M. bovis—BCG* infection after routine BCG vaccination. None of the patients expressed IL-12Rβ1 on the surface of their T cells or NK cells as determined by flow cytometry and their lymphocytes did not respond to IL-12 stimulation ex vivo. All the patients were screened for mutations on *IL-12Rβ1* by PCR amplification of genomic DNA (17 exons and exon/intron boundaries), followed by direct sequencing of the PCR products. Among this cohort, mutation 1623_1624delinsTT was found in 5 affected individuals, all of them born to European ancestors. Three patients were homozygous and 2 heterozygous for this trait. Another patient, Belgium-born and heterozygous for the same mutation, was also included in this study.

The patients' ancestry was established by direct questioning for at least 4 generations in each affected family with no evidence of consanguinity. However, encrypted relatedness could not be ruled out completely.

Thirty-four polymorphic sites on chromosome 19 were analyzed in these patients: 21 intragenic *IL12RB1* (15 exonic, 6 intronic) and 13 extrageneic. All markers were tested on gDNA. In addition, IL-12Rβ1 mRNA was extracted, cDNA converted, PCR amplified, subcloned and sequenced in heterozygous patients for individual-allele *IL12RB1*-exonic polymorphic markers determination.

Genomic DNA was obtained by saline extraction from PBMCs. Single nucleotide polymorphisms' genotyping was performed with DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences, Little Chalfont Buckinghamshire, UK) from PCR products obtained using specific primers (Appendix) and Taq polymerase (Invitrogen, Carlsbad, CA). Messenger RNA was extracted from PBMCs of 1623_1624delinsTT heterozygous patients according to the manufacturer's instructions (Trizol, Invitrogen, Invitrogen, Carlsbad, CA). Messenger RNA was reverse transcribed to cDNA using Oligo-dT (Invitrogen, Carlsbad, CA) and reverse transcriptase (Invitrogen, Carlsbad, CA) following producers' protocols. Interleukin-12 receptor β1 cDNA was amplified with Taq polymerase and specific primers. The obtained PCR product was cloned into a TOPO TA vector (TOPO TA Cloning kit, Invitrogen, Carlsbad, CA) and later used to transform TOP10 One Shot Chemically Competent cells (Invitrogen, Carlsbad, CA). Transformed bacteria were seeded on Luria Bertani (LB)-XGal-Ampicillin (Amp) plates, efficiently-transformed bacteria were grown in LB-Amp media, plasmidic DNA was extracted (Wizard Plus SV Minipreps, Promega, Madison, WI), sequenced on an ABI PRISM 3130 Genetic Analyzer, and analyzed using Gene Scan software (Applied Biosystem). Microsatellites were PCR-amplified with fluorescent-labeled specific F primers, run on an ABI PRISM 3130 Genetic Analyzer, and analyzed using the Gene Scan software.

2.2. Control population

Genomic DNA was extracted from 50 unrelated healthy Argentinean blood donors of European-descent (100 chromosomes). Samples were screened for SNPs G451C and A705G by restriction fragment length polymorphism (RFLP) analysis and/or direct sequencing. For SNP A705G allele frequency determination, exon 7 IL12RB1 was PCR-amplified with specific primers, the obtained product was PvuII-digested according to manufacturer's recommendations (New England Biolabs, Ipswish, MA), resolved in 2% agarose gels, stained with ethidium bromide and analyzed under UV light. Digestion with PvuII in the presence of allele A705 generates 2 fragments of 107 and 91 bp and ruled out the presence of the mutated haplotype. For G451C allele frequency determination, exon 4 IL12Rβ1 was PCR-amplified with specific primers, the obtained product was digested with restriction enzyme HphI (New England Biolabs, Ipswish, MA), PAGE resolved in NOVEX 20% TBE gels (Invitrogen, Carlsbad, CA), stained with ethidium bromide and analyzed under UV light. Digestion with HphI generates a set of four fragments (320, 172, 112 and 36 bp) in the presence of allele 451C, while a different set of fragments (284, 226, 94 and 36 bp) is observed in the presence of allele G451. Homozygosity for G451 ruled out the presence of the mutated haplotype. Six normal controls (12 chromosomes) were tested by direct sequencing of the entire coding region of IL12Rβ1 gene.

All patients and controls gave written consent for DNA/RNA extraction and analysis. Primer sequences are shown in the Appendix A, Table A1.

2.3. Mutation 1623_1624delinsTT age estimation

In order to estimate the age to the most recent common ancestor of haplotypes carrying the 1623_1624delinsTT mutation, a previously described likelihood method was modified and applied on a set of six recombination markers located at variant distances on both sides of the disease mutation (Genin et al., 2004). The approximation was introduced by Genin and collaborators as an interval censored survival analysis in which the mutation locus, denoted as Δ, is the starting point, the genetic distance is the time scale and the occurrence of recombination is the event. The model takes the following basic assumptions:

- monophyletic origin of the disease mutation
- independence of each sampled chromosome and equal time to a common ancestor
- independent recombination rates at each side of Δ
- negligible probability of double recombination

Given $M_1,\ldots,M_x,\ldots,M_k$ k ordered markers typed on one side of Δ, let $\theta_1,\ldots,\theta_x,\ldots,\theta_k$ be the recombination fractions from Δ (for convenience, $\theta_0 = 0$ and $\theta_{k+1} = 1$). Define $S(x)$, the probability that no recombination took place during n generations between Δ and M_x as

$$S(x) = (1-\theta_x)^n.$$

Therefore, the probability of recombination in the xth interval—between M_{x-1} and M_x—namely $f(x)$, is

$$f(x) = S(x-1) - S(x).$$

Taking advantage of this construction, missing information on M_{x-1} for any sampled chromosome could be handled by considering the most distant marker from Δ without recombination (M_{x-2}, M_{x-3}, etc). However, as the intervals include more missing markers the assumption of negligible double recombination becomes less sustainable, and further formulations would be needed. Under the described model, the likelihood function is

constructed independently for each side of Δ and for a sample of N chromosomes two situations should be noticed. In the first one, every chromosome shows evidence of recombination to one side of Δ. In this case, there are two groups of haplotypes. In one group, denoted as G_1, y chromosomes ($2 \le y \le N$) share all their alleles for markers $M_1,\ldots M_{k-1}$ and have a different allele at M_k. For the G_1 group, the likelihood is written

$$L_{G_1}(n) = f(k)^y + yS(k)\,f(k)^{y-1}, \quad (1)$$

taking into account the uncertainty on which the ancestral allele is. In the remaining $N-y$ chromosomes, G_2, a recombination occurred on the interval x_i ($1 \le i \le N-y$) closer to Δ ($x_i \le k$). For G_2 the likelihood is written

$$L_{G_2}(n) = \prod_{i \in G_2} f(x_i). \quad (2)$$

A less desirable, unless possible, scenario is found when there is no recombination before M_k. In other words, a set of z chromosomes, group G_3, present the same alleles for M_1 to M_k. Accordingly, likelihood is

$$L_{G_3}(n) = S(k)^z. \quad (3)$$

Note that when the G_3 group is not empty then G_1 is. Finally, chromosomes with missing data for markers $M_a,\ldots,M_k$ ($1 \le a \le k$) and without evidence of recombination for markers $M_1,\ldots,M_{a-1}$ conform G_4, with likelihood

$$L_{G_4}(n) = \prod_{j \in G_4} S(x_j), \quad (4)$$

where $x_j < k$ is the interval to the last marker with available information and no recombination.

In conclusion, the likelihood for n generations on the studied side is

$$L_{side}(n) = [I_{G_1 \neq \varnothing} L_{G_1}(n) + I_{G_3 \neq \varnothing} L_{G_3}(n) + (1 - I_{G_1 \neq \varnothing} - I_{G_3 \neq \varnothing})] L_{G_2}(n) L_{G_4}(n), \quad (5)$$

where $I_{G_\alpha \neq \varnothing}$ equals 0 if group α is empty and 1 if it is not. The final likelihood, $L(n)$, is the product of the two side likelihoods. Point estimation of generation number n is obtained at $L(n)$ maximum.

Equations (1)–(5) hold without ambiguity if the presence or absence of a recombination event within intervals is known. However, they do not account for mutation, resulting in a false positive recombination signal, or low allele diversity, resulting in a false negative recombination signal. The adjustments introduced to deal with this ambiguity are described in the Appendix. Further calculation for 1623_1624delinsTT included both corrections.

Computation of 95% confidence intervals (CI) for an estimated generation number was based on Bayesian principles. Briefly, 95% CI is given by all n which satisfies:

$$0.025 \le \frac{\sum_{u=1}^{n} L(u)}{\sum_{v=1}^{V} L(v)} \le 0.975,$$

where V is a large generation number (e.g. 10^4).

Algorithms were programmed on R version 2.6.0, a free software environment. Scripts are available upon request. Likelihood corrections for mutation and allele frequency calculation are described in the Appendix.

2.4. Extragenic markers recombination fraction estimation

The preceding theory is based on *a priori* knowledge on recombination fractions $\theta_1,\ldots,\theta_k$. Thereafter, recombination rates for the intervals among the selected extragenic markers at both sides of 1623_1624 delGCinsTT were estimated by means of a robust linear regression, carried out with published data on 37 neighboring polymorphic sites incorporated into the Marshfield linkage map (Fig. 2). Linkage (Kosambi centiMorgans, cM) was inferred for D19S1037, CRLF1, D19S895, SFRS14, TSSK6 and D19S215 based on sequence position on chromosome 19 (mega base pairs, Mb). Calculated fractions are presented in Table 1. The result represents an average of 1.16 cM per Mb across the entire region.

3. Results

3.1. Mutation 1623_1624delinsTT haplotyping

All patients carrying mutation 1623_1624delinsTT in homozygous or heterozygous state shared a common haplotype on the mutated allele of chromosome 19, which included the complete *IL12RB1* gene (Table 1). Thirty-four polymorphic sites on chromosome 19 were analyzed: 21 intragenic *IL12RB1* (15 exonic, 6 intronic) and 13 extrageneic. Of the 34 markers analyzed, 32 were included on the mutated haplotype (from C72A to TSSK6:T989C; distance between markers, 1.45 Mb). Markers D19S1037 and D19S215, which showed variations in homozygous and heterozygous patients, flanked and defined the mutated haplotype (distance between markers, 3.51 Mb).

None of the 100 control chromosomes analyzed carried the mutated haplotype above described. Control chromosomes were screened for SNPs G451C and A705G (variants 451C and 705G on the mutated allele) by restriction fragments length polymorphism (RFLP) analysis (Fig. 1). Eighty-eight percent of the tested samples were excluded as carrying the mutated haplotype by this method; the remaining 12% were ruled out by *IL12RB1* gDNA direct sequencing. Since the mutated haplotype was not found in any of the 100 control chromosomes tested, its frequency remains ≤0.01% in the general population. Therefore, the probability that mutation 1623_1624delinsTT have occurred by chance sharing the same set of polymorphic markers in the 9 affected alleles is not larger than the probability of randomly sampling nine times in a row the mutated haplotype, namely $(0.01)^9 = 10^{-18}$. Together these data strongly suggested that the 1623_1624delinsTT mutated allele originated from a single founder.

3.2. Mutation 1623_1624delinsTT age calculation

The most likely age for mutation 1623_1624delinsTT was 19 generations, with a 95% Confidence Interval (CI) of 7–51. Considering 25 years for each generation, this represents 475 years (95% CI, 175–1275) (Fig. 2). Estimation of the mutation age was calculated from 6 extragenic recombination markers applying a modified likelihood-method based on that originally described by Genin et al., 2004. Variations on microsatellites D19S1037, D19S895, and D19S215, and 10 SNPs distributed along genes CRLF1 (3 SNPs), SFRS14 (5 SNPs) and TSSK6 (2 SNPs) were tested. The 10 SNPs were considered as 3 independent loci—one for each gene—assuming null recombination frequencies within each gene. Mutations rates and allele frequencies were taken into account according to previous reports (Weber and Wong, 1993; Nachman and Crowell, 2000). Mutation rates were set to 5.6×10^{-4}, 2.1×10^{-3} and 2.55×10^{-8} for dinucleotide tandem repeats, tetranucleotide tandem repeats, and SNPs, respectively, as shown in Section 2.

To assess the contribution of the assumed mutation frequencies to the estimated time of the founder mutation, the same calculations were carried out for varying mutation rates. Assumption that the rate of mutations was one order (10 times) higher resulted in an estimation of 12 generations (95% CI, 4–30), while if the rate was assumed to be one order lower, the estimated mutation age would be 21 generations (95% CI, 7–55). Of note, these modifications yielded a projected outcome encompassed by the initial confidence

Table 1
Mutation 1623_1624delinsTT haplotyping.

Marker	rs/UniSTS	Location	PD (Mb)	AF	RF	P1	P2	P3	P4	P5	P6
D19S1037	UniSTS:38487	Extragenic	0,094	0.143	0.138	1/2	2/3	ND	2/4	ND	1/4
C 72 A	rs17884651	*IL12RB1* ex.		1	im	**C**	**C**	**C**	**C**	**C**	**C**
C 202 T	rs17887176	*IL12RB1* ex.		1	im	**C**	**C**	**C**	**C**	**C**	**C**
G 451 C	rs11086087	*IL12RB1* ex.		0.087	im	**C**	**C**	**C**	**C**	**C**/G	**C**/G
G 531 A	rs11575926	*IL12BB1* ex.		0.804	im	**G**	**G**	**G**	**G**	**G**/A	**G**
A 705 G	rs11575934	*IL12RB1* ex.		0.304	im	**G**	**G**	**G**	**G**/A	**G**/A	**G**
C 748 T	rs17852635	*IL12RB1* ex.		0.295	im	**T**	**T**	**T**	**T**/C	**T**/C	**T**
C 1081 A	rs17884957	*IL12RB1* ex.		1	im	**C**	**C**	**C**	**C**	**C**	**C**
T 1158 C	rs375947	*IL12RB1* ex.		0.304	im	**C**	**C**	**C**	**C**/T	**C**/G	**C**
G 1196 C	rs401502	*IL12RB1* ex.		0.696	im	**C**	**C**	**C**	**C**/G	**C**/G	**C**/G
C 1213 T	rs17882216	*IL12RB1* ex.		1	im	**C**	**C**	**C**	**C**	**C**	**C**
C 1306 T	rs376271	*IL12RB1* ex.		ND	im	**C**	**C**	**C**	**C**	**C**	**C**
C 1376 T		*IL12RB1* ex.		ND	im	**C**	**C**	**C**	**C**	**C**	**C**
1623_1624delinsTT		**Mutation**				**Homoz**	**Homoz**	**Homoz**	**Heteroz**	**Heteroz**	**Heteroz**
C 108530 T		*IL12RB1* int.		ND	im	**C**	**C**	**C**	**C**	C	**C**
C 108571 T		*IL12RB1* int.		ND	im	**T**	**T**	**T**	**T**	**T**/C	**T**/C
G 108616 A		*IL12RB1* int.		ND	im	**A**	**A**	**A**	**A**	**A**/G	**A**/G
T 108661 C		*IL12RB1* int.		ND	im	**C**	**C**	**C**	**C**	**C**/T	**C**
G 1637 A	rs11575935	*IL12RB1* ex.		ND	im	**G**	**G**	**G**	**G**	**G**	**G**
C 1783 T	rs17885102	*IL12RB1* ex.		1	im	**C**	**C**	**C**	**C**	**C**	**C**
G 1845 C		*IL12RB1* ex.		ND	im	**G**	**G**	**G**	**G**	**G**	**G**
G rh66781 A	UniSTS:47762	*IL12RB1* int.		ND	im	**G**	**G**	**G**	**G**	**G**	**G**
5C SHGC156161 6C	UniSTS:185287	*IL12RB1* int.		ND	im	**6C**	**6C**	**6C**	**6C**	**6C**/5C	**6C**
CRLF1:G6489A	rs2238647	Extragenic	0,506	0.783	0.596	**G**	**G**	**G**	**G**	**G**	**G**
CRLF1: G6560-	rs34603196	Extragenic	0,506	ND	0.596	**G**	**G**	**G**	**G**	**G**	**G**
CRLF1:C6579T	rs7259478	Extragenic	0,506	1	0.596	**C**	**C**	**C**	**C**	**C**	**C**
D19S895	UniSTS:79811	Extragenic	0,555	0.376	0.645	**1/1**	**1/1**	**1/1**	**1**/2	**1/1**	**1**/3
SFRS14:C274T	rs35401849	Extragenic	0,904	0.987	1.088	**C**	**C**	**C**	**C**	**C**	**C**
SFRS14:A305G	rs10404860	Extragenic	0,904	1	1.088	**A**	**A**	**A**	**A**	**A**	**A**
SFRS14:G606T	rs12459416	Extragenic	0,904	1	1.088	**G**	**G**	**G**	**G**	**G**	**G**
SFRS14:T1102C	rs34471092	Extragenic	0,904	0.986	1.088	**T**	**T**	**T**	**T**	**T**	**T**
SFRS14:C1344T	rs4808907	Extragenic	0,904	ND	1.088	**C**	**C**	**C**	**C**	**C**	**C**
TSSK6:A350G	rs7250893	Extragenic	1,427	0.562	1.653	**A**	**A**	**A**	**A**	**A**	**A**
TSSK6:T989C	rs17851212	Extragenic	1,427	ND	1.653	**T**	**T**	**T**	**T**	**T**	**T**
D19S215	UniSTS:28209	Extragenic	3,416	0.379	3.941	1/4	1/1	1/1	1/1	1/2	3/4

Single allele nomenclature is used for homozygous SNPs, while both alleles are shown for heterozygous SNPs. Single repeat numbers denote homozygous microsatellite repeats, while both repeats are shown for heterozygous ones. Black-shaded: mutation 1623_1624delinsTT. Gray-shaded: distal markers defining and flanking the mutated haplotype. PD, Physical Distance to *IL12RB1* expressed in mega-bases (Mb); AF, Allelic Frequency; RF, Recombination Frequency; P1–P6, Patients 1–6; ND, Not Done; im, *IL12RB1* intragenic marker (null recombination was assumed); Homoz., Homozygous for mutation 1623_1624delinsTT; Heteroz., Heterozygous for mutation 1623_1624delinsTT; *IL12RB1* ex., *IL12RB1* exonic; *IL12RB1* int., *IL12RB1* intronic.

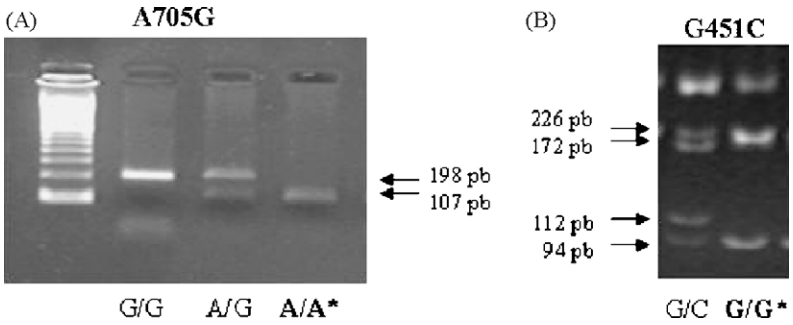

Fig. 1. Mutated haplotype screening on control chromosomes. (A) Gel electrophoresis of PvuII-digested exon 7 *IL12RB1*. The 3 digestion patterns found in the Control population are shown. A705 homozygosity (A/A*) ruled out the mutated haplotype. (B) Gel electrophoresis of HphI-digested exon 4 *IL12RB1*. The 2 possible digestion patterns found in the Control population are shown. G451 homozygosity (G/G*) ruled out the mutated haplotype.

interval. Variations on the ancestral allele frequency for D19S1037, arbitrarily set to 1/7 (1/n° of described alleles), only had a slight influence on inference. Specifically, for an allele frequency of 0.5, the estimation would correspond to 20 generations (95% CI, 7–52), while if D19S1037 allele frequency contribution were set to 0, the inferred number of generations was 19, same as initial calculations (95% CI, 7–50). A basic model assumption for the derived estimations was a phylogeny with star topology (i.e., equal distance to a common ancestor) and so haplotype relatedness would bias computed generation numbers. In an attempt to assess this assumption, age to origin was evaluated for 9 sub-samples, sequentially dropping one chromosome at the time. All estimations rounded the initial calculation of 19 generations with a range of 13 generations (95% CI, 3–40) to 22 generations (95% CI, 8–60). This relatively small variation suggests that this assumption is valid, though relatedness among some inherited haplotypes cannot be ruled out. In conclusion, the estimation of 19 generations (7–51 with a 95% confidence) to a common origin for the 1623_1624delinsTT mutation appeared robust.

4. Discussion

In the *IL12RB1* gene mutation database (http://153.1.44.13/IL12RB1base/), 24 different mutations affecting 50 individuals from 38 unrelated families are summarized. Mutation 1623_1624delinsTT is described in 7 individuals (5 homozygous and 2 heterozygous) from 4 different families previously reported by Fieschi et al. (2003). This same mutation is even more prevalent among Argentinean IL-12Rβ1 deficient patients: 8-out-of-12 affected alleles (almost 67%) bear mutation 1623_1624delinsTT. Noteworthy, although all the Argentinean IL-12Rβ1 deficient patients, as well as the Belgian patient studied, belonged to families with European ancestry, their families did not appear to be related.

When identical mutations are detected in apparently unrelated patients, it is important to discriminate whether they have aroused by recurrent mutations (*mutational hot spot*) or they are identical by descent (*mutational founder effect*). In this case, the absence of direct repeats, palindromes/quasi-palindromes, polypurine runs, polypyrimidine runs or consensus motifs typically associated with different types of mutational hotspots in the surrounding area argued against a putative mutational hotspot (Krawczak and Cooper, 1993). On the other hand, the detection of a common haplotype spanning 1.45 (distance between C72A and TSSK:T989C, distal markers common to all affected alleles) to 3.51 Mb (distance between D19S1037 and D19S215, first markers to show differences on the affected alleles), strongly supports the mutational founder effect hypothesis.

Many Mendelian disorders demonstrate mutations that can be traced to a founder whose existence can be inferred from the unique chromosomal background on which the mutation occurred (Zeegers et al., 2004). Founder mutations have been described for several primary immunodeficiencies, including those conferring MSMD. Mutation g.482 + 82_856–854 (a large loss-of-function deletion) and g.315_316insA (a frameshift insertion) in IL12B (IL-12p40) are founder mutations associated to increased susceptibility to mycobacteria and salmonella infections arising 700 and 1100 years ago, respectively (Picard et al., 2002; Li et al., 2002; Campbell et al., 2003; Sanchez et al., 2007). In the case of IL-12Rβ1 deficiency due to mutation 1623_1624delinsTT, the founder effect was estimated to occur 475 years ago (95% CI, 175–1275). To determine so, a slightly modified likelihood method from that originally described by Genin et al. (2004) was applied. The introduced modifications used different parameters to better adjust the intrinsic bias on mutation age-estimation inherent to this type of calculations.

It is noteworthy, that the city of Buenos Aires was funded by Pedro de Mendoza in 1536, approximately at the same time as the estimated arousal of mutation 1623_1624delinsTT. It is tempting to speculate that mutation 1623_1624delinsTT was introduced into the country in the mid-16th century by the Spanish colonization. The current Argentinean population has a genetically heterogeneous ethnic background mainly conferred by waves of European immigrants and their integration to a sparse native population. Following the colonization by the Spaniards in the mid-16th century there was a significant decrease (“bottlenecks”) on the native population associated with massive killings and transmitted

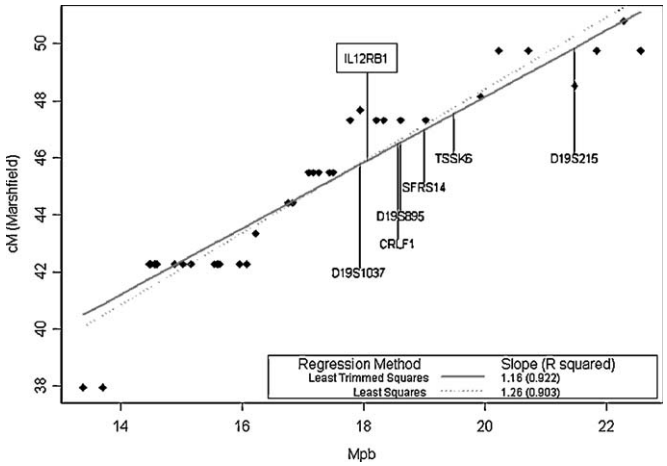

Fig. 2. Extragenic markers recombination fraction estimation. Known distances and linkage for polymorphic sites surrounding IL12RB1 locus (filled rhombi) are shown. Linkage estimation for the extragenic markers indicated on the graph was based on a robust regression method. The slope derived from the least squares method is also despicted, although it shows a poorer fit and a slight bias due to outlying observations.

diseases (Sanchez et al., 2007; Martínez Marignac et al., 2004; Mulligan et al., 2004; Alfaro et al., 2005; Resano et al., 2007). While mating between the European colonizers and the natives helped to spread the mutation, the reduction on the local population contributed to the mutation enrichment of the genetic pool.

In summary, herein we present strong evidence to support that mutation *IL12RB1* 1623_1624delinsTT, a rare genomic change, is inherited as a haplotype block with a common founder arousing 475 years ago (95% CI, 175–1275). This finding represents the first founder effect described on the IL-12/23 receptor complex, critical for controlling mycobacterial as well as salmonella infections, and affecting patients with European ancestry. The reason(s) behind the persistency of this mutation across multiple generations, or whether it confers any kind of selective advantage, has yet to be established.

Acknowledgments

JY and SDR work is supported by the National Institutes of Health Fogarty International Center, and by the Fogarty International Research Collaboration Award Grant R01TW006644. The authors have no conflict of interests to declare.

Appendix A. Appendix

A.1. Likelihood correction for mutation

Beside recombination, haplotype differences can also be explained by mutation. Some formulations from Genin et al. (2004) were applied to account for this ambiguity. Assuming known mutation rates, $\mu_1,\ldots,\mu_x,\ldots,\mu_k$ (for $M_1,\ldots,M_x,\ldots,M_k$, respectively), the probability that during n generations no marker mutation has occurred from Δ up to M_x, denoted $U(x)$, is

$$U(x) = \prod_{w=1}^{x} (1-\mu_h)^n$$

and the probability of mutation at marker M_x over n generations, $u(x)$, has the expression

$$u(x) = 1-(1-\mu_x)^n.$$

Thus, the likelihood contribution for a chromosome i which shares an ancestral haplotype until marker M_{x-1} and has different allele at marker M_x is

$$h(x) = U(x_i-1)[f(x_i)+\mu_{x_i}S(x_i)],$$

which contemplates the possibilities that either a recombination occurred between markers M_{x-1} and M_x or no recombination took place up to marker M_x but a mutation event did happen at M_x.

Lastly, to account for mutation in the likelihood formulations, the following corrections were applied to equations (1)–(4) in the main text:

$$L^{\mathrm{mut}}_{G_1}(n) = h(k)^y + yU(k)S(k)h(k)^{y-1}$$

$$L^{\mathrm{mut}}_{G_2}(n) = \prod_{i\in G_2} h(x_i)$$

$$L^{\mathrm{mut}}_{G_3}(n) = [U(k)S(k)]^z$$

$$L^{\mathrm{mut}}_{G_4}(n) = \prod_{j\in G_4} U(x_j)S(x_j)$$

A.2. Likelihood correction for allele frequency

When polymorphisms of markers are low, an event of recombination with the same allele becomes more plausible. Thereafter, to account for hindered crossing-over, a modification to the second order approximation suggested in Genin et al. (2004) was implemented.

Let p_x be the frequency of the allele present on the ancestral haplotype at the M_x locus. Considering only the possibility of two unnoticed events of recombination, expressions (1)–(4) become:

$$\begin{aligned} L^{allele}(n) = {} & f(k)^y + y\,p_{k-1}\,f(k-1)\,f(k)^{y-1} \\ & + \binom{y}{2} p^2\,f(k-1)^2\,f(k)^{y-2} + y\,p_{k-1}\,p_{k-2}\,f(k-2)\,f(k)^{y-1} \\ & + y(y-1)\,p^2\,p_{k-2}\,f(k-2)\,f(k-1)\,f(k) + yS(k)\,f(k)^{y-1} \\ & + y(y-1)S(k)\,p_{k-1}\,f(k-1)\,f(k)^{y-2} \\ & + y\binom{y-1}{2} p^2 S(k)\,f(k-1)^2\,f(k)^{y-3} + y(y \\ & - 1)\,p_{k-1}\,p_{k-2}S(k)\,f(k-2)\,f(k)^{y-2} + y(y-1)(y \\ & - 2)\,p^2\,p_{k-2}S(k)\,f(k-2)\,f(k-1)\,f(k) \end{aligned}$$

- The term with coefficient $\binom{y-1}{2}$ is only computed if $y>2$.

$$L^{\mathrm{allele}}(n) = \prod_{i\in G_2} [f(x_i) + p_{x_i-1}\,f(x_i-1) + p_{x_i-1}\,p_{x_i-2}\,f(x_i-2)],$$

$$L^{\mathrm{allele}}(n) = [S(k) + p_k\,f(k) + p_k\,p_{k-1}\,f(k-1)]^z,$$

$$L^{\mathrm{allele}}(n) = \prod_{j\in G_4} [S(x_j) + p_{x_j}\,f(x_j) + p_{x_j-1}\,p_{x_j}\,f(x_j-1)].$$

Table A1
Primers.

Intragenic markers		
SNP		Sequence (5′ → 3′)
C72A (exon 1)	Forward	ctc agc ttc aat gtg ttc cgg
	Reverse	aca cat aca cgt gcc tcc ac
C202T (exon 3)	Forward	ttg tct tag gca cga aga cc
	Reverse	gat cac gca tcc gag agt ag
G451C (exon 4)	Forward	atg cac aaa gtc ggc tgt gg
	Reverse	gga atc ctc tct agc tcc ag
G531A (exon 5)	Forward	agt tgc tta ccc tgg aca gg
	Reverse	gga cag atg cag aga tgg tg
A705G (exon 7)	Forward	aga gtg aga ccg tgt ctc ag
C748T (exon 7)	Reverse	gct gga act aca ggt gtg ca
C1081A (exon 9)	Forward	gtc tgt ctt cat agg ctg tg
	Reverse	cac ccc taa atc agg cac tc
T1158C (exon 10)	Forward	gga gag aga tgg caa ctg tc
G1196C (exon 10)	Reverse	acc cac cag gac cta aaa gg
C1213T (exon 10)		
C1306T (exon 11)	Forward	agt agg agt ttt ctg ggg gt
C1376T (exon 11)	Reverse	aga cag ggt ttc gcc aca tt
G1637A (exon 13)	Forward	gat cat tgt ggg aag cgc ag
	Reverse	ggc ctc tga gga gta aag ag
C1783T (exon 15)	Forward	aca gag caa gac tcc gtc tc
G1845C (exon 15)	Reverse	atg cgt aac cct tgt cca gc
Extragenic markers		
Amplicon		Sequence (5′ → 3′)
D19S1037	Forward	ctg cggagt cag aaa aca gt
	Reverse	atg cag cta tcc ctc att ca
CRLF1 (exon 2)	Forward	cca act tac ggc caa cat ag
	Reverse	aca atc att aac agc gtc ttt t
D19S895	Forward	ttg cag taa cca tgc cac
	Reverse	cca gga cac ttt tgc tag tca c
SFRS14 (exon 3)	Forward	aac ttt ggc tga ttt cct tac c
	Reverse	tcc agg ctg att gag aaa gag t
TSSK6 (exon 1)	Forward	cct ggc gga cgg gct gaa ctg
	Reverse	gcg cca tgg ctt aga ccc gag att
D19S215	Forward	cat gca tta aaa atg aca act gt
	Reverse	gct ctg can tcc att act ca

Note that the only difference to previous formulations from Genin et al. (2004) are the terms for $L_{G1}(n)$ that contemplate recombination on M_{k-1} and M_{k-2} on two independent events, namely

$$y(y-1)\,p^2\,p_{k-2}\,f(k-2)\,f(k-1)\,f(k)^{y-2}$$

and

$$y(y-1)(y-2)\,p^2\,p_{k-2}S(k)\,f(y-2)\,f(y-1)\,f(k)^{y-3}.$$

Following the same line of thought that derives in the original expression, the extension includes the probabilities of observing ancestral alleles at M_{k-1} and M_{k-2} after one recombination event on the $(k-2)$th interval, $p_{k-1}p_{k-2}(k-2)$, and one ancestral allele at M_{k-1} after one recombination event on the $(k-1)$th interval, $p_{k-1}(k-1)$. The additional terms make a more significant contribution to the final likelihood as $(y-1)p_{k-1}$ increases and $f(k)/f(k-1)$ decreases.

References

Alfaro, E.L., Pierini, J.E., Gutierrez, N.I., Vullo, C.M., 2005. Genetic structure and admixture in urban populations of the Argentine North-West. Ann. Hum. Biol. 32, 724–737.

Altare, F., Durandy, A., Lammas, D., Emile, J.F., Lamhamedi, S., Le Deist, F., Drysdale, P., Jouanguy, E., Döffinger, R., Bernaudin, F., Jeppsson, O., Gollob, J.A., Meinl, E., Segal, A.W., Fischer, A., Kumararatne, D., Casanova, J.L., 1998. Impairment of mycobacterial immunity in human interleukin-12 receptor deficiency. Science 280, 1432–1435.

Campbell, C., Mitui, M., Eng, L., Coutinho, G., Thorstenson, Y., Gatti, R.A., 2003. ATM mutations on distinct SNP and STR haplotypes in ataxia-telangiectasia patients of differing ethnicities reveal ancestral founder effects. Hum. Mutat. 21, 80–85.

de Jong, R., Altare, F., Haagen, I.A., Elferink, D.G., Boer, T., van Breda Vriesman, P.J., Kabel, P.J., Draaisma, J.M., van Dissel, J.T., Kroon, F.P., Casanova, J.L., Ottenhoff, T.H., 1998. Severe mycobacterial and *Salmonella* infections in interleukin-12 receptor–deficient patients. Science 280, 1435–1438.

Fieschi, C., Dupuis, S., Catherinot, E., Feinberg, J., Bustamante, J., Breiman, A., Altare, F., Baretto, R., Le Deist, F., Kayal, S., Koch, H., Richter, D., Brezina, M., Aksu, G., Wood, P., Al-Jumaah, S., Raspall, M., Da Silva Duarte, A.J., Tuerlinckx, D., Virelizier, J.L., Fischer, A., Enright, A., Bernhöft, J., Cleary, A.M., Vermylen, C., Rodriguez-Gallego, C., Davies, G., Blütters-Sawatzki, R., Siegrist, C.A., Ehlayel, M.S., Novelli, V., Haas, W.H., Levy, J., Freihorst, J., Al-Hajjar, S., Nadal, D., De Moraes Vasconcelos, D., Jeppsson, O., Kutukculer, N., Frecerova, K., Caragol, I., Lammas, D., Kumararatne, D.S., Abel, L., Casanova, J.L., 2003. Low penetrance, broad resistance, and favorable outcome of interleukin 12 receptor β1 deficiency, medical and immunological implications. J. Exp. Med. 197, 527–535.

Fieschi, C., Bosticardo, M., de Beaucoudrey, L., Boisson-Dupuis, S., Feinberg, J., Santos, O.F., Bustamante, J., Levy, J., Candotti, F., Casanova, J.L., 2004. A novel form of complete IL-12/IL-23 receptor beta1 deficiency with cell surface–expressed nonfunctional receptors. Blood 104, 2095–2101.

Filipe-Santos, O., Bustamante, J., Chapgier, A., Vogt, G., de Beaucoudrey, L., Feinberg, J., Jouanguy, E., Boisson-Dupuis, S., Fieschi, C., Picard, C., Casanova, J.L., 2006. Inborn errors of IL-12/23- and IFN-gamma-mediated immunity: molecular, cellular, and clinical features. Semin. Immunol. 18, 347–361.

Genin, E., Tullio-Pelet, A., Begeot, F., Lyonnet, S., Abel, L., 2004. Estimating the age of rare disease mutations: the example of Triple-A syndrome. J. Med. Genet. 41, 445–449.

Haerynck, F., Holland, S.M., Rosenzweig, S.D., Casanova, J.L., Van daele, S., Schelstraete, P., De Baets, F., 2008. Disseminated Mycobacterium avium infection in a patient with a novel mutation in the interleukin-12 receptor β1 chain. J. Pediatr. 153, 721–722.

Hoeve, M.A., de Boer, T., Langenberg, D.M.L., Sanal, O., Verreck, F.A.W., Ottenhoff, T.H.M., 2003. IL-12 receptor deficiency revisited: IL-23-mediated signaling is also impaired in human genetic IL-12 receptor b1 deficiency. Eur. J. Immunol. 33, 3393–3397.

Krawczak, M., Cooper, D.N., 1993. Human Gene Mutations. Bios Scientific Publishers, Oxford.

Li, L., Moshous, D., Zhou, Y., Wang, J., Xie, G., Salido, E., Hu, D., de Villartay, J.P., Cowan, M.J., 2002. A founder mutation in Artemis, an SNM1-like protein, causes SCID in Athabascan-speaking Native Americans. J. Immunol. 168, 6323–6329.

Martínez Marignac, V.L., Bertoni, B., Parra, E.J., Bianchi, N.O., 2004. Characterization of admixture in an urban sample from Buenos Aires, Argentina, using uniparentally and biparentally inherited genetic markers. Hum. Biol. 76, 543–557.

Mulligan, C.J., Hunley, K., Cole, S., Long, J.C., 2004. Population genetics, history, and health patterns in Native Americans. Annu. Rev. Genomics Hum. Genet. 5, 295–315.

Nachman, M.W., Crowell, S.L., 2000. Estimate of the mutation rate per nucleotide in humans. Genetics 156, 297–304.

Picard, C., Fieschi, C., Altare, F., Al-Jumaah, S., Al-Hajjar, S., Feinberg, J., Dupuis, S., Soudais, C., Al-Mohsen, I.Z., Génin, E., Lammas, D., Kumararatne, D.S., Leclerc, T., Rafii, A., Frayha, H., Murugasu, B., Wah, L.B., Sinniah, R., Loubser, M., Okamoto, E., Al-Ghonaium, A., Tufenkeji, H., Abel, L., Casanova, J.L., 2002. Inherited interleukin-12 deficiency: IL12B genotype and clinical phenotype of 13 patients from six kindreds. Am. J. Hum. Genet. 70, 336–348.

Resano, M., Esteban, E., González-Pérez, E., Vía, M., Athanasiadis, G., Avena, S., Goicoechea, A., Bartomioli, M., Fernández, V., Cabrera, A., Dejean, C., Carnese, F., Moral, P., 2007. How many populations set foot through the Patagonian door? Genetic composition of the current population of Bahía Blanca (Argentina) based on data from 19 Alu polymorphisms. Am. J. Hum. Biol. 19, 835–877.

Rosenzweig, S.D., Yancoski, J., Bernasconi, A., Krasovec, S., Marciano, B.E., Casimir, L., Berberian, G., Simboli, N., Rousseau, M., Calle, G., 2006. Thirteen years of culture-positive M. bovis-BCG infection in an IL-12Rbeta1 deficient patient: treatment and outcome. J. Infect. 52, e69–e72.

Sanchez, J.J., Monaghan, G., Børsting, C., Norbury, G., Morling, N., Gaspar, H.B., 2007. Carrier frequency of a nonsense mutation in the adenosine deaminase (ADA) gene implies a high incidence of ADA-deficient severe combined immunodeficiency (SCID) in Somalia and a single, common haplotype indicates common ancestry. Ann. Hum. Genet. 71, 336–347.

van de Vosse, E., Ottenhoff, T.H., 2006. Human host genetic factors in mycobacterial and Salmonella infection: lessons from single gene disorders in IL-12/IL-23-dependent signaling that affect innate and adaptive immunity. Microbes Infect. 8, 1167–1173.

Weber, J.L., Wong, C., 1993. Mutation of human short tandem repeats. Hum. Mol. Genet. 2, 1123–1128.

Zeegers, M.P., van Poppel, F., Vlietinck, R., Spruijt, L., Ostrer, H., 2004. Founder mutations among the Dutch. Eur. J. Hum. Genet. 12, 591–600.