Figure S1

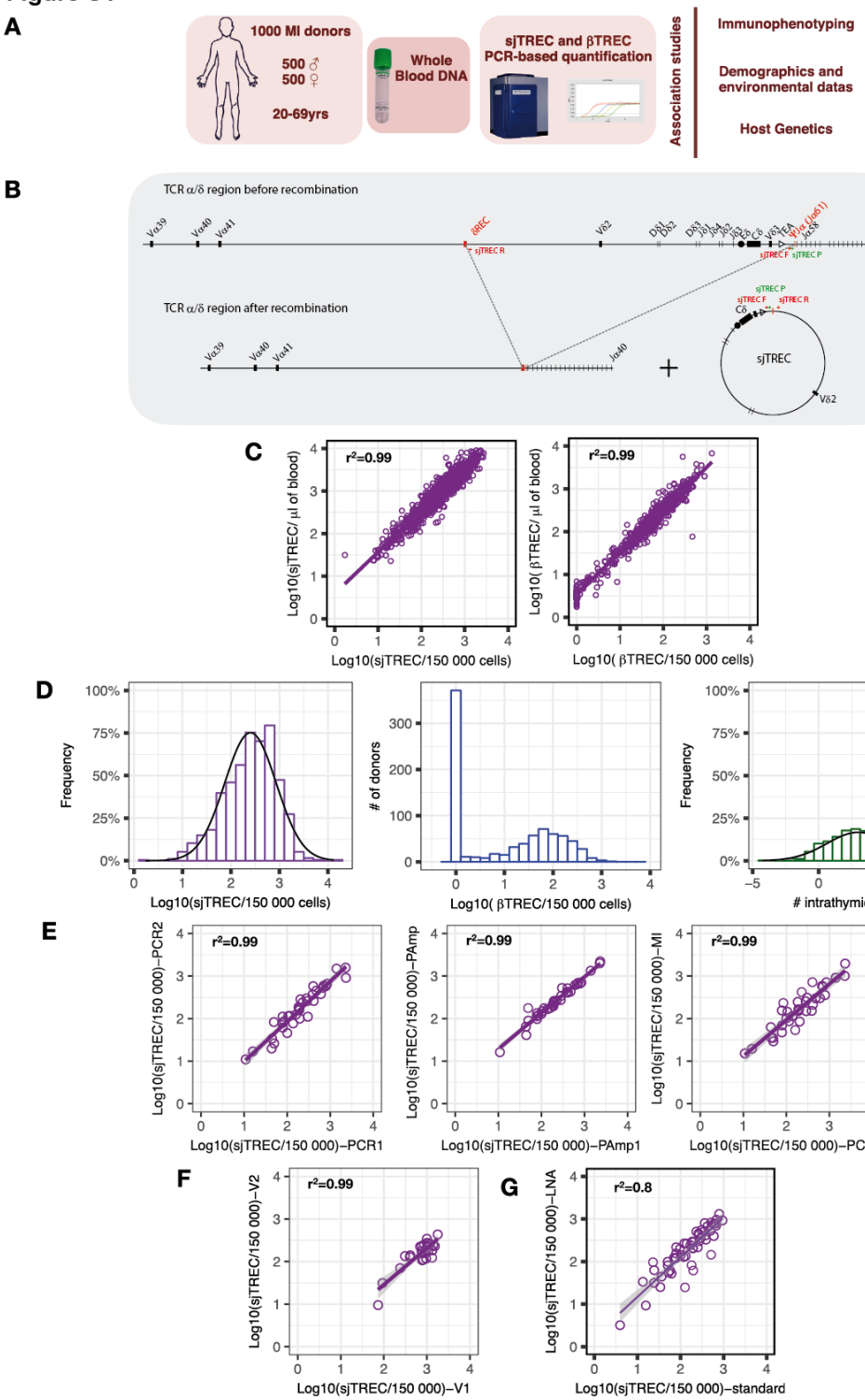
A
1000 MI donors
500 ♂
500 ♀
20-69yrs
Whole Blood DNA
sjTREC and βTREC PCR-based quantification
Association studies
Immunophenotyping
Demographics and environmental datas
Host Genetics
B
TCR α/δ region before recombination
TCR α/δ region after recombination
sjTREC
C
r²=0.99
Log10(sjTREC/ µl of blood)
Log10(sjTREC/150 000 cells)
Log10(βTREC/ µl of blood)
Log10(βTREC/150 000 cells)
D
Frequency
Log10(sjTREC/150 000 cells)
# of donors
Log10(βTREC/150 000 cells)
# intrathymic divisions
E
Log10(sjTREC/150 000)–PCR2
Log10(sjTREC/150 000)–PCR1
Log10(sjTREC/150 000)–PAmp
Log10(sjTREC/150 000)–PAmp1
Log10(sjTREC/150 000)–MI
F
Log10(sjTREC/150 000)–V2
Log10(sjTREC/150 000)–V1
G
r²=0.8
Log10(sjTREC/150 000)–LNA
Log10(sjTREC/150 000)–standard

Fig. S1. Technical workflow of the study and validation of the TREC assay in the MI cohort.

(A) 500 healthy men and 500 healthy women, stratified over 5 decades of life, were recruited in the *MI* study. For each donor, the circulating levels of sjTRECs and βTRECs were quantified by RT-PCR on DNA extracted from whole blood. In parallel, hundreds of demographic, medical and lifestyle variables were collected and extensive immunophenotyping and genome-wide genotyping were performed. **(B)** Genomic location in the *TCRA-TCRD* locus of primers (sjTREC-F and sjTREC-R, in red) and probe (sjTREC-P, in green) used for the sjTREC assay. Sequences of primers and probe are provided in table S1. **(C)** TRECs were normalized "per 150,000 cells", an estimate of 1μg DNA content (n=969). Numbers of cells were evaluated using an RT-PCR targeting the Albumin gene, obtained simultaneously as TREC quantification. TRECs can also be normalized per μl of blood, using the absolute number of CD45+ cells obtained by flow cytometry (n=892). We confirmed in our cohort that the two normalized measures are tightly correlated for both sjTRECs and βTRECs (r^2=0.99), and used the "number of TRECs per 150,000 cells" measurement in all subsequent analyses. **(D)** Distributions of $\log_{10}$-transformed sjTRECs (purple, n=969), $\log_{10}$-transformed βTRECs (blue, n=969), and the number of intrathymic divisions (green, n=506 because the ratio could not be calculated on the 463 donors negative for βTRECs) in the *MI* cohort. Over the 969 healthy donors, $\log_{10}$-transformed sjTRECs were approximately normally distributed. In contrast, βTRECs showed a bimodal distribution, with 463 donors under the threshold of detection of the assay (βTREC=0). Among donors with measurable whole-blood βTRECs, the $\log_{10}$-transformed values approximately followed a normal distribution. The number of intrathymic divisions was estimated as the $\log_2$-transformation of the ratio between sjTRECs and βTRECs, and was approximately normally distributed in the cohort. **(E)** To assess the reproducibility of the TREC assay on Biomark HD, technical replicates we made using 43 random donors from the cohort: first, two independent Biomark quantifications were done on the same preamplification plate (PCR1 and PCR2, r^2=0.99) and secondly, two independent preamplifications were done on the same donors (PAmp 1 and PAmp 2; r^2=0.99). Finally, the PCR1 technical replicates were compared to the values obtained for the same 43 donors tested on different plates during the course of

the study (PCR1 and MI; r^2=0.99). **(F)** To assess the biological reproducibility of the TREC assay, we quantified sjTREC levels on biological replicates collected for 29 same individuals at 2 to 3 weeks interval. A high correlation was found between the two measures (r^2=0.99). **(G)** Impact of the rs76132819 SNP on probe binding and assay efficiency. This SNP was found to locate inside the sjTREC probe (position 26 C/A), and was in moderate linkage disequilibrium (*D'*=1.0, r^2=0.13) with the SNPs identified by genome wide association study (table S3). Although the A allele has a frequency below 8% and does not influence the binding of the probe at 60°C (as predicted *in silico*), we designed a shorter probe that contains only the 26 first nucleotides, thus excluding the rs76132819 SNP (table S1). Same Tm as original probe was obtained by using 4 Locked Nucleic Acids (LNA). We quantified sjTRECs on 58 *MI* donors. A high correlation was found between the two measures (r^2=0.8).

Figure S2

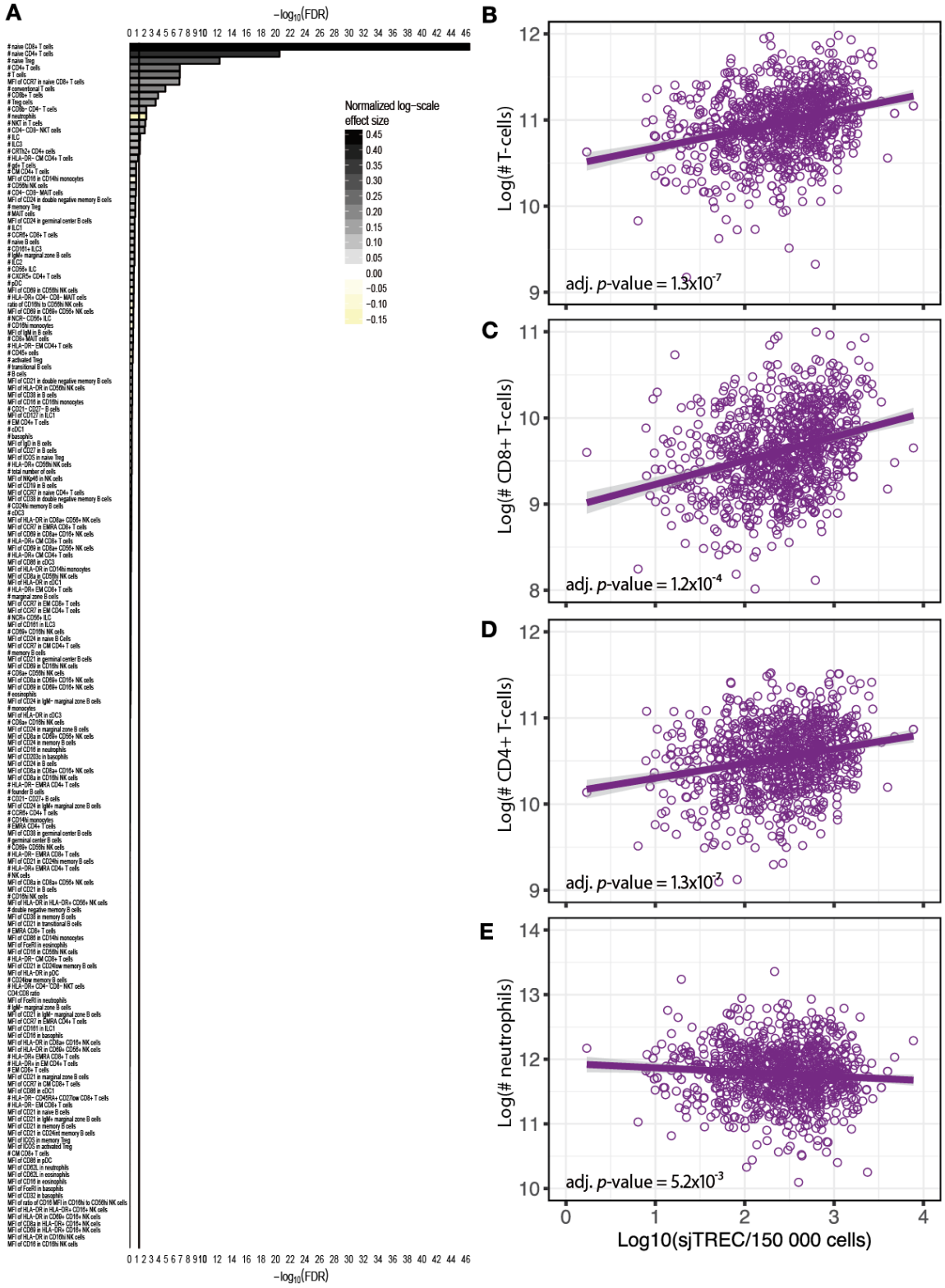
A
−log10(FDR)
0 1 2 3 4 5 6 7 8 9 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46
Normalized log−scale effect size
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
B
Log(# T-cells)
adj. p-value = 1.3x10-7
C
Log(# CD8+ T-cells)
adj. p-value = 1.2x10-4
D
Log(# CD4+ T-cells)
adj. p-value = 1.3x10-2
E
Log(# neutrophils)
adj. p-value = 5.2x10-3
Log10(sjTREC/150 000 cells)

Fig. S2. Association of sjTRECs with immune cell counts and parameters.

(**A**) Significance of the effect of sjTRECs on immunophenotypes, as estimated by the $-\log_{10}$(adjusted *P*-values) of the Kenward-Rogers approximated F tests from mixed models including age, sex, CMV serostatus and smoking status as fixed effect covariates together with day of blood draw as a random effect. Normalized log-scale effect sizes were color-coded. The vertical black line indicates the $-\log_{10}$ of chosen threshold for statistical significance ($-\log_{10}$(adj. P=0.05)=1.30103). (**B-E**) Significant associations between sjTRECs and the numbers of circulating T-cells, $CD4^+$ and $CD8^+$ T-cells, and neutrophils (n=969). Regression lines were fitted using simple linear regression. Adjusted *P*-values were obtained using the mixed model described in (**A**).

Figure S3

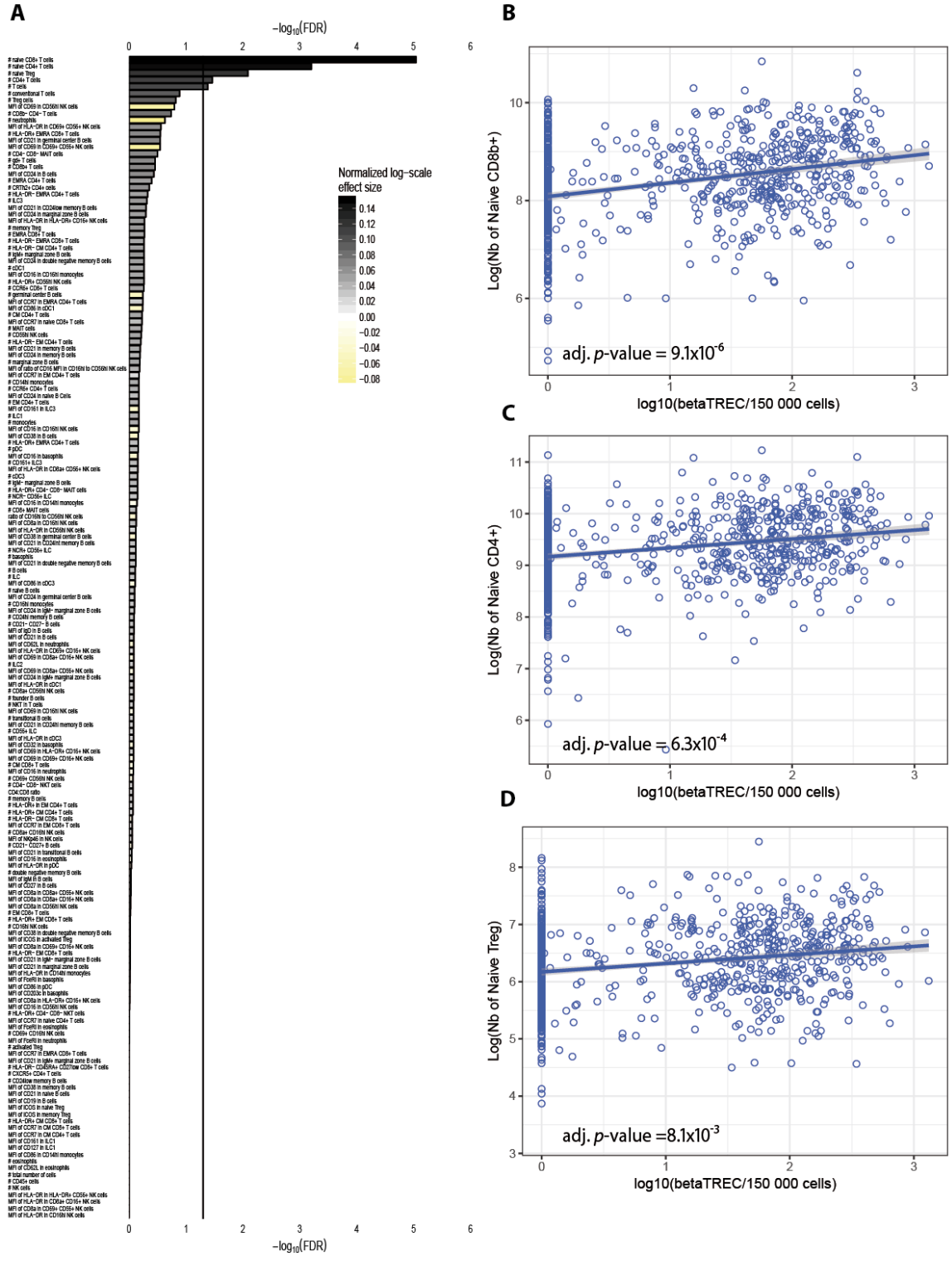
A
−log10(FDR)
Normalized log−scale effect size
B
Log(Nb of Naive CD8b+)
adj. p-value = 9.1x10-6
log10(betaTREC/150 000 cells)
C
Log(Nb of Naive CD4+)
adj. p-value = 6.3x10-4
log10(betaTREC/150 000 cells)
D
Log(Nb of Naive Treg)
adj. p-value =8.1x10-3
log10(betaTREC/150 000 cells)

Fig. S3. Association of βTRECs with immune cell counts and parameters.

(A) Significance of the effect of βTRECs on immunophenotypes, as estimated by the $-\log_{10}$(adjusted *P*-values) of the Kenward-Rogers approximated F tests from mixed models including age, sex, CMV serostatus and smoking status as fixed effect covariates together with day of blood draw as a random effect. Log-scale normalized effect sizes were color-coded. The vertical black line indicates the $-\log_{10}$ of chosen threshold for statistical significance ($-\log_{10}$(adj. P=0.05)=1.30103). Significant association between βTRECs and **(B)** the number of circulating naive $CD8^+$ T-cells, **(C)** the number of circulating naïve $CD4^+$ T-cells and **(D)** the number of circulating naïve Treg cells (n=832). Regression lines were fitted using simple linear regression. Adjusted *P*-values were obtained using the mixed model described in (**A**).

Figure S4

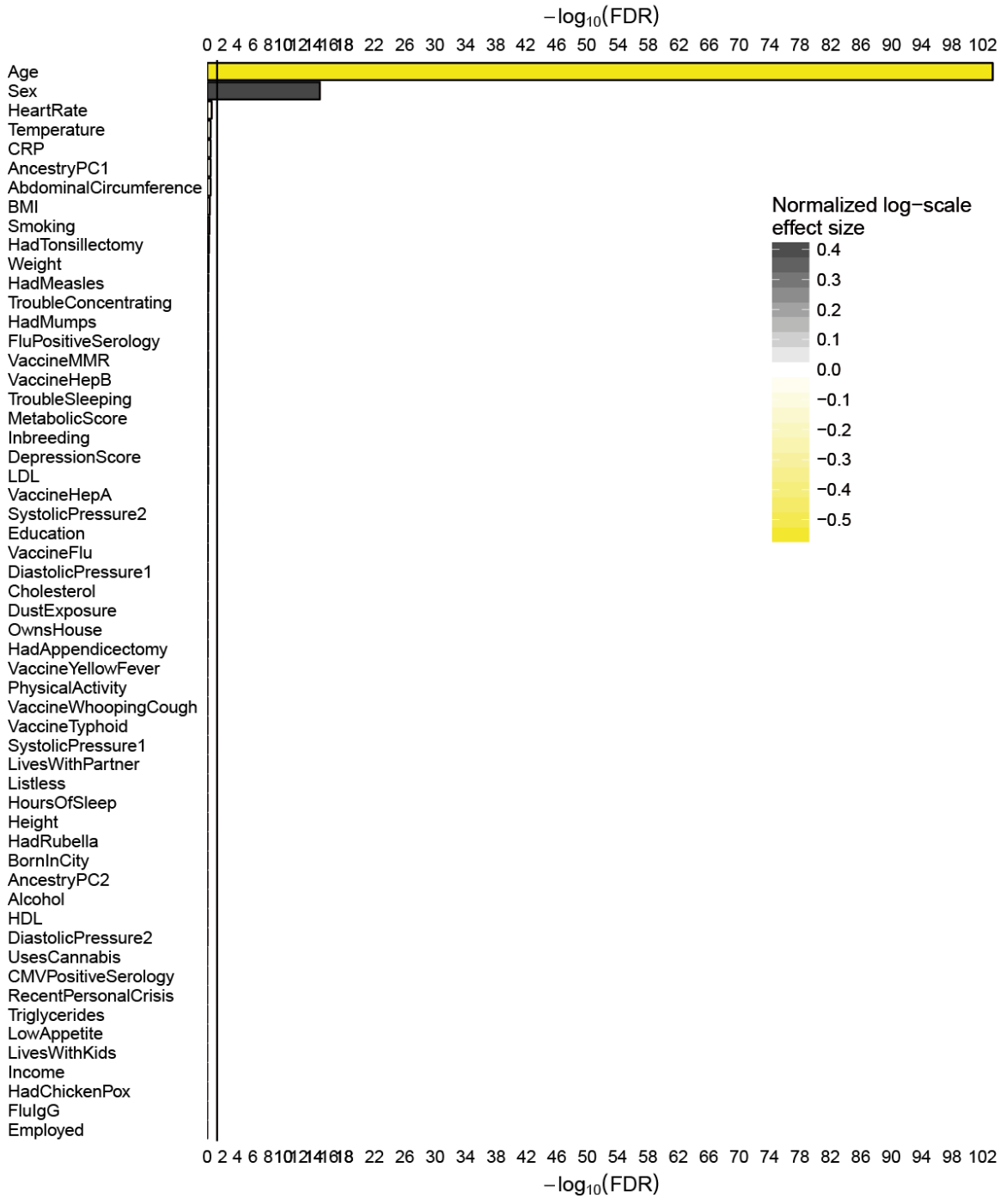

Fig. S4. Association of sjTRECs with nonheritable factors.

Significance of the effect of demographic, medical and lifestyle variables on sjTRECs, as estimated by the $-\log_{10}$(adjusted *P*-values) of the Kenward-Rogers approximated F tests

from mixed models including the tested treatment variable, age, sex, (unless these were the treatment) as fixed effects covariates and TREC processing plates as random effects (n=969). Normalized log-scale effect sizes were color-coded. The vertical black line indicates the $-\log_{10}$ of chosen threshold for statistical significance ($-\log_{10}$(adj. P=0.05)=1.30103).

Figure S5

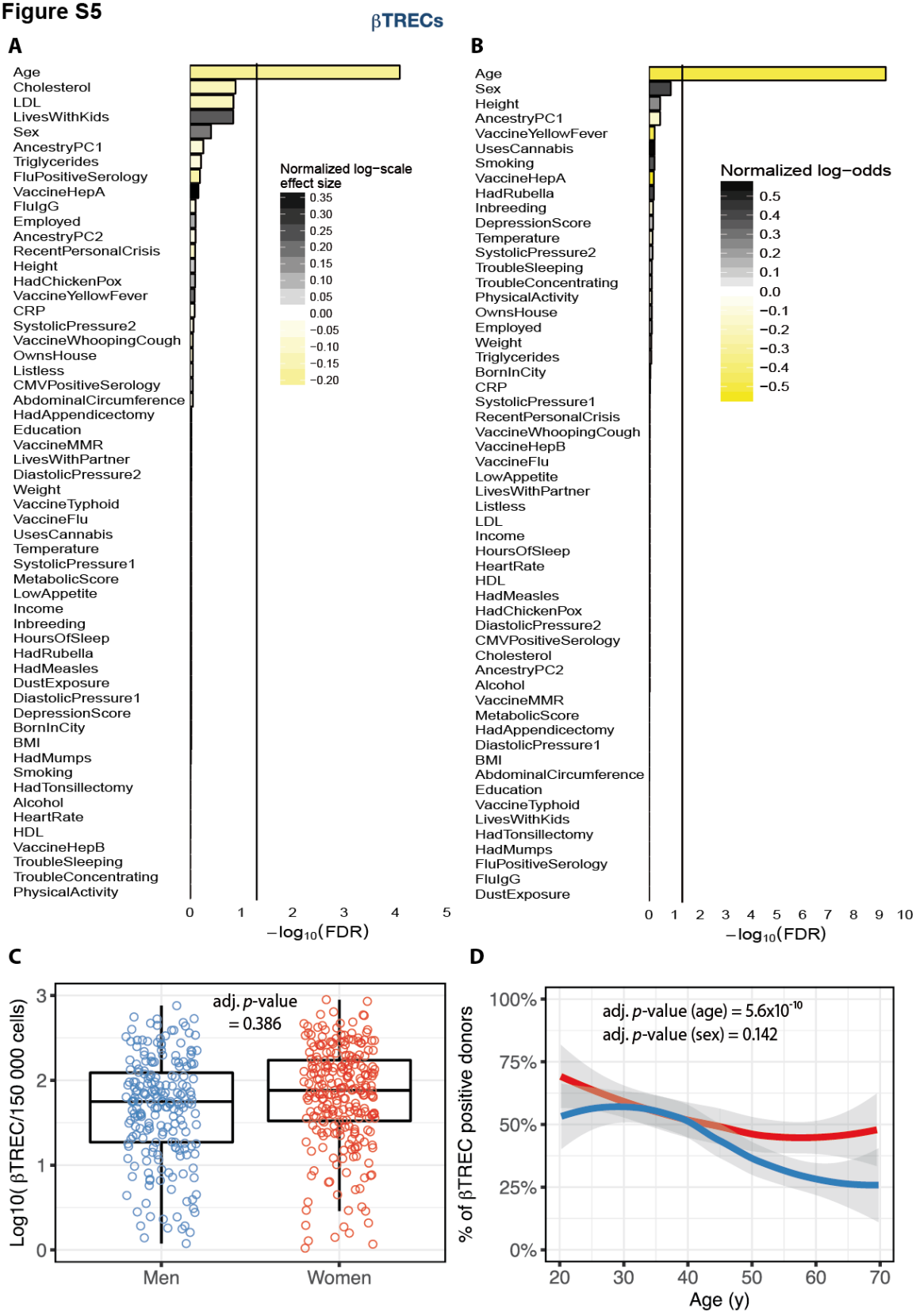
βTRECs
A
Age
Cholesterol
LDL
LivesWithKids
Sex
AncestryPC1
Triglycerides
FluPositiveSerology
VaccineHepA
FluIgG
Employed
AncestryPC2
RecentPersonalCrisis
Height
HadChickenPox
VaccineYellowFever
CRP
SystolicPressure2
VaccineWhoopingCough
OwnsHouse
Listless
CMVPositiveSerology
AbdominalCircumference
HadAppendicectomy
Education
VaccineMMR
LivesWithPartner
DiastolicPressure2
Weight
VaccineTyphoid
VaccineFlu
UsesCannabis
Temperature
SystolicPressure1
MetabolicScore
LowAppetite
Income
Inbreeding
HoursOfSleep
HadRubella
HadMeasles
DustExposure
DiastolicPressure1
DepressionScore
BornInCity
BMI
HadMumps
Smoking
HadTonsillectomy
Alcohol
HeartRate
HDL
VaccineHepB
TroubleSleeping
TroubleConcentrating
PhysicalActivity
Normalized log−scale effect size
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
−0.20
0 1 2 3 4 5
−log10(FDR)
B
Age
Sex
Height
AncestryPC1
VaccineYellowFever
UsesCannabis
Smoking
VaccineHepA
HadRubella
Inbreeding
DepressionScore
Temperature
SystolicPressure2
TroubleSleeping
TroubleConcentrating
PhysicalActivity
OwnsHouse
Employed
Weight
Triglycerides
BornInCity
CRP
SystolicPressure1
RecentPersonalCrisis
VaccineWhoopingCough
VaccineHepB
VaccineFlu
LowAppetite
LivesWithPartner
Listless
LDL
Income
HoursOfSleep
HeartRate
HDL
HadMeasles
HadChickenPox
DiastolicPressure2
CMVPositiveSerology
Cholesterol
AncestryPC2
Alcohol
VaccineMMR
MetabolicScore
HadAppendicectomy
DiastolicPressure1
BMI
AbdominalCircumference
Education
VaccineTyphoid
LivesWithKids
HadTonsillectomy
HadMumps
FluPositiveSerology
FluIgG
DustExposure
Normalized log−odds
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
0 1 2 3 4 5 6 7 8 9 10
−log10(FDR)
C
Log10(βTREC/150 000 cells)
adj. p-value = 0.386
Men
Women
D
% of βTREC positive donors
adj. p-value (age) = 5.6x10-10
adj. p-value (sex) = 0.142
100%
75%
50%
25%
0%
20 30 40 50 60 70
Age (y)

Fig. S5. Association of βTRECs with nonheritable factors.

(**A**) Significance of the effect of demographic, medical and lifestyle variables (table S2) on βTREC levels, as estimated by the $-\log_{10}$(adjusted *P*-values) of the Kenward-Rogers approximated F tests from mixed models including the tested treatment variable, age, sex, (unless these were the treatment) as fixed effects covariates and TREC processing plates as random effects. Normalized log-scale effect sizes were color-coded. The vertical black line indicates the $-\log_{10}$ of chosen threshold for statistical significance ($-\log_{10}$(adj. *P*=0.05)=1.30103). **(B)** Significance of the effect of demographic, medical and lifestyle variables (table S2) on the probability of having detectable βTREC levels, as estimated by the $-\log_{10}$(adjusted *P*-values) of the likelihood ratio tests from logistic regression models including the tested treatment variable, age and sex (unless these were the treatment). Normalized log-scale effect sizes were color-coded. The vertical black line indicates the $-\log_{10}$ of chosen threshold for statistical significance ($-\log_{10}$(adj. *P*=0.05)=1.30103). **(C)** βTREC levels as a function of sex (n=506). Adjusted *P*-values were obtained using the mixed model described in (A). (**D**) Percentages of having detectable βTREC values as a function of age in men (blue) and women (red). *P*-values were obtained using a logistic regression with a likelihood ratio chi-square test, with the βTREC binary variable (i.e., 0 and 1 indicating undetectable and detectable βTREC levels) as the response, and age and sex as treatments (n=463).

Figure S6

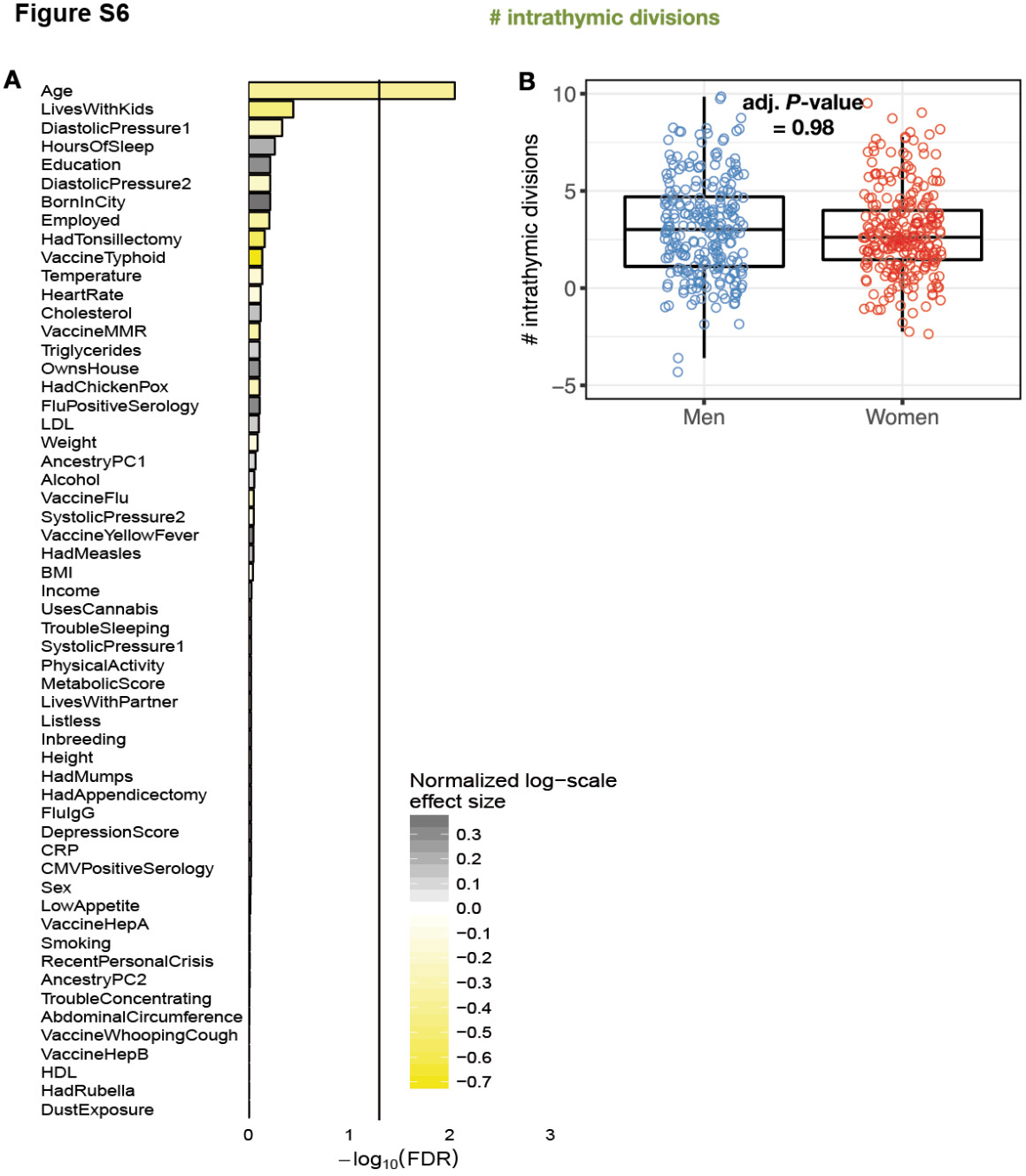

Fig. S6. Association of intrathymic division number with nonheritable factors.

(A) Significance of the effect of demographic, medical and lifestyle variables (table S2) on the number of intrathymic divisions, as estimated by the $-\log_{10}$(adjusted *P*-values) of the Kenward-Rogers approximated F tests from mixed models including the tested treatment variable, age, sex, (unless these were the treatment) as fixed effects covariates and TREC processing plates as random effects. Log-scale normalized effect sizes were

color-coded. Vertical black line indicates the $-\log_{10}$ of chosen threshold for statistical significance ($-\log_{10}$(adj. P=0.05)=1.30103). **(B)** Number of intrathymic divisions in women and men of the *MI* cohort (n=506). Adjusted *P*-values were obtained using the mixed model described in (**A**).

Figure S7

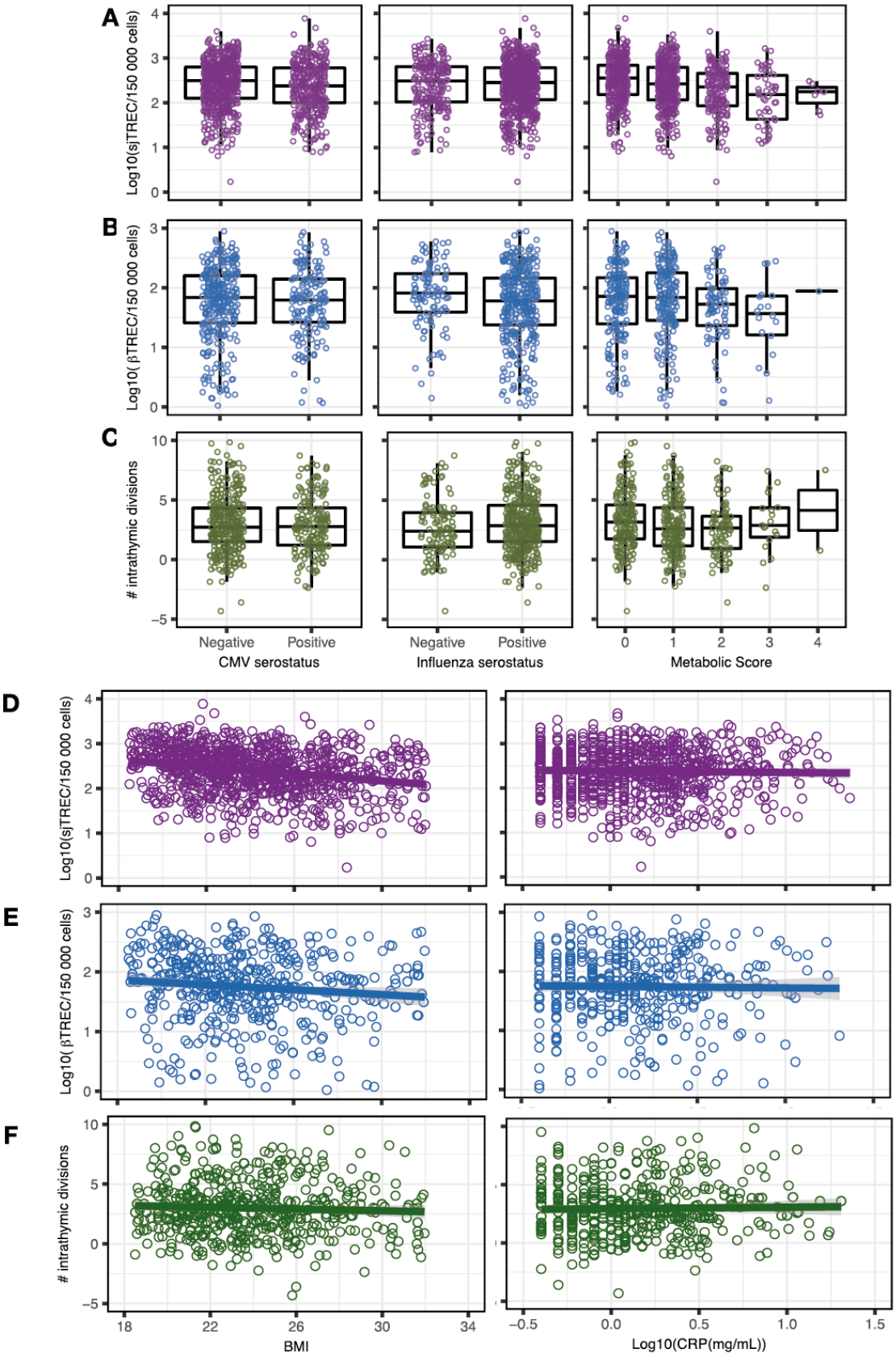
A
Log10(sjTREC/150 000 cells)
B
Log10(βTREC/150 000 cells)
C
intrathymic divisions
Negative
Positive
CMV serostatus
Negative
Positive
Influenza serostatus
0
1
2
3
4
Metabolic Score
D
Log10(sjTREC/150 000 cells)
E
Log10(βTREC/150 000 cells)
F
intrathymic divisions
18
22
26
30
34
BMI
−0.5
0.0
0.5
1.0
1.5
Log10(CRP(mg/mL))

Fig. S7. Association of thymic function parameters with specific nonheritable factors.

Effects of cytomegalovirus (CMV) serostatus, influenza A virus (IVA) serostatus and metabolic score (*10*) on **(A)** sjTRECs (n=969), **(B)** βTRECs (n=506), and **(C)** number of intrathymic divisions (n=506). Effects of body mass index (BMI) and $\log_{10}$ of C reactive protein (CRP) on **(D)** sjTRECs (n=969), **(E)** βTRECs (n=506) and **(F)** number of intrathymic divisions (n=506).

Figure S8

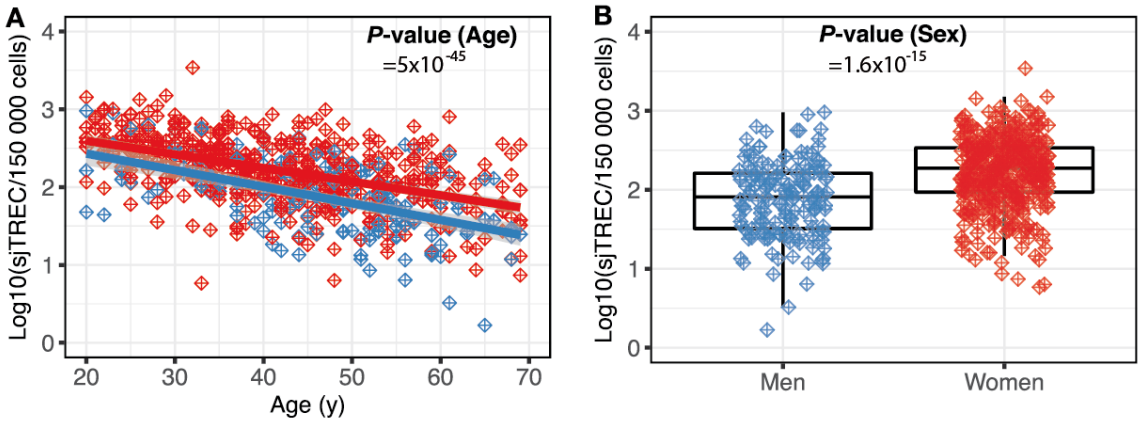

Fig. S8. Age and sex impact on sjTRECs in the MARTHA cohort.

Effects of (**A**) age and (**B**) sex on sjTRECs in the independent MARTHA cohort (n=612). Adjusted *P*-values are the FDR-adjusted *P*-values for the Kenward-Rogers approximated F test of a age (or sex) effect obtained using a mixed model for the response variable $\log_{10}$ sjTRECs including sex (or age) and TREC processing plate as fixed effects, and additional batch variables as random effects. Regression lines were fitted using linear regression. Blue indicates men (n=160) and red indicates women (n=452).

Figure S9

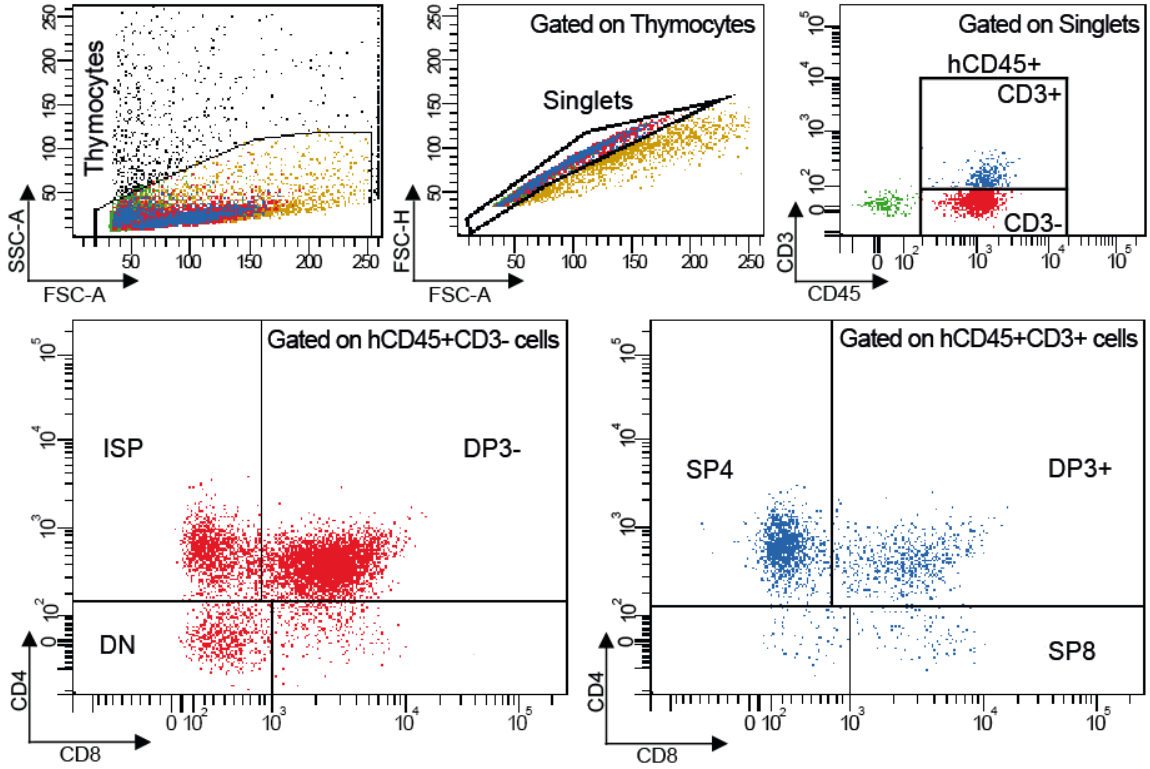

Fig. S9. **Human immune system mice flow cytometry gating strategy.**

Identification of the human thymocyte subsets (hCD45+, CD3+, DN, ISP, DPCD3- and CD3+, SP4 and SP8) using flow cytometry. Human $CD45^{pos}$ was used to establish the optimal human T lymphocyte gate for immunophenotyping and exclude the murine thymic cells. Data are representative of individual samples.

Figure S10

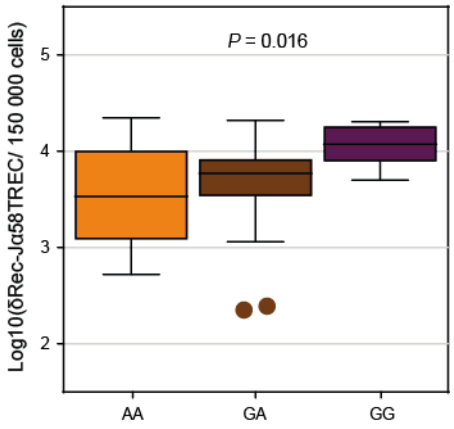

Fig. S10. Effect of the SNP rs2204985 polymorphism on the *TCRD* locus deletion alternative rearrangement δRec-Jα58. δRec-Jα58 rearrangement excision circles (fig. S1B) were evaluated in reconstituted thymi from 8-29 weeks old immunodeficient Balb/c Rag2$^{-/-}$Il2rg$^{-/-}$Sirpa^{NOD} (BRGS) mice reconstituted with human CD34+ hematopoietic progenitors harvested from fetal livers with rs2204985 genotype AA (orange, n=10), GA (brown, n=21) or GG (purple, n=12). Values were log_{10}-transformed and *P*-value of a Kruskal-Wallis test is indicated.

Figure S11

		AA	AA	AA	AA	AA	AA	AA	AA	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG	GG
		45605T3R	45609T3R	45615T3R	45742T3R	47139T5T	47140T5T	47143T5T	47145T5T	47692T13	47693T13	47694T13	47695T13	47697T13	47702T13	47707T13	62T14TH	64T14TH	65T14TH	68T14TH	69T14TH
AA	45605T3RD	NA	0,08219	0,15453	0,36750	0,01368	0,08937	0,06151	0,04373	0,10516	0,01729	0,12829	0,12062	0,18010	0,18677	0,16889	0,53319	0,29839	0,05639	0,29951	0,21629
AA	45609T3RD	0,08219	NA	0,23644	0,11159	0,02027	0,15318	0,12258	0,11364	0,04756	0,02959	0,04026	0,03973	0,04860	0,05057	0,04956	0,05777	0,05498	0,05058	0,05634	0,05615
AA	45615T3RD	0,15453	0,23644	NA	0,22060	0,02318	0,22536	0,15961	0,09764	0,07293	0,03917	0,06848	0,06901	0,08736	0,09608	0,08187	0,11386	0,10061	0,07474	0,10395	0,08604
AA	45742T3RD	0,36750	0,11159	0,22060	NA	0,01442	0,11476	0,07355	0,05082	0,09241	0,02055	0,10748	0,10364	0,15118	0,15541	0,14029	0,32693	0,22199	0,05852	0,22348	0,16821
AA	47139T5TH	0,01368	0,02027	0,02318	0,01442	NA	0,03319	0,04447	0,03760	0,00872	0,00943	0,00911	0,00939	0,01075	0,00915	0,01082	0,01063	0,01039	0,01442	0,01049	0,00924
AA	47140T5TH	0,08937	0,15318	0,22536	0,11476	0,03319	NA	0,17450	0,09200	0,05342	0,03984	0,04770	0,04828	0,05994	0,05559	0,05701	0,06368	0,06117	0,06319	0,06011	0,05406
AA	47143T5TH	0,06151	0,12258	0,15961	0,07355	0,04447	0,17450	NA	0,13742	0,03905	0,03740	0,03367	0,03521	0,03969	0,04183	0,04297	0,04409	0,04242	0,06483	0,04370	0,03953
AA	47145T5TH	0,04373	0,11364	0,09764	0,05082	0,03760	0,09200	0,13742	NA	0,02634	0,02309	0,02362	0,02348	0,02282	0,03156	0,02865	0,02874	0,03022	0,03715	0,02782	0,03331
GG	47692T13TH	0,10516	0,04756	0,07293	0,09241	0,00872	0,05342	0,03905	0,02634	NA	0,01371	0,05920	0,05683	0,07165	0,06878	0,06953	0,09824	0,08057	0,03476	0,08470	0,07089
GG	47693T13TH	0,01729	0,02959	0,03917	0,02055	0,00943	0,03984	0,03740	0,02309	0,01371	NA	0,01226	0,01134	0,01416	0,01331	0,01294	0,01391	0,01471	0,01605	0,01364	0,01186
GG	47694T13TH	0,12829	0,04026	0,06848	0,10748	0,00911	0,04770	0,03367	0,02362	0,05920	0,01226	NA	0,06004	0,07926	0,07993	0,07588	0,12268	0,10124	0,03493	0,10182	0,08531
GG	47695T13TH	0,12062	0,03973	0,06901	0,10364	0,00939	0,04828	0,03521	0,02348	0,05683	0,01134	0,06004	NA	0,07736	0,07690	0,07532	0,11481	0,09434	0,03038	0,09573	0,07809
GG	47697T13TH	0,18010	0,04860	0,08736	0,15118	0,01075	0,05994	0,03969	0,02282	0,07165	0,01416	0,07926	0,07736	NA	0,10082	0,10107	0,17285	0,12986	0,03901	0,13740	0,10532
GG	47702T13TH	0,18677	0,05057	0,09608	0,15541	0,00915	0,05559	0,04183	0,03156	0,06878	0,01331	0,07993	0,07690	0,10082	NA	0,09734	0,18036	0,13906	0,03834	0,14226	0,11139
GG	47707T13TH	0,16889	0,04956	0,08187	0,14029	0,01082	0,05701	0,04297	0,02865	0,06953	0,01294	0,07588	0,07532	0,10107	0,09734	NA	0,16065	0,12769	0,04071	0,13078	0,10470
GG	62T14TH	0,53319	0,05777	0,11386	0,32693	0,01063	0,06368	0,04409	0,02874	0,09824	0,01391	0,12268	0,11481	0,17285	0,18036	0,16065	NA	0,30206	0,04945	0,30945	0,21547
GG	64T14TH	0,29839	0,05498	0,10061	0,22199	0,01039	0,06117	0,04242	0,03022	0,08057	0,01471	0,10124	0,09434	0,12986	0,13906	0,12769	0,30206	NA	0,04811	0,21370	0,16154
GG	65T14TH	0,05639	0,05058	0,07474	0,05852	0,01442	0,06319	0,06483	0,03715	0,03476	0,01605	0,03493	0,03038	0,03901	0,03834	0,04071	0,04945	0,04811	NA	0,04622	0,04376
GG	68T14TH	0,29951	0,05634	0,10395	0,22348	0,01049	0,06011	0,04370	0,02782	0,08470	0,01364	0,10182	0,09573	0,13740	0,14226	0,13078	0,30945	0,21370	0,04622	NA	0,16237
GG	69T14TH	0,21629	0,05615	0,08604	0,16821	0,00924	0,05406	0,03953	0,03331	0,07089	0,01186	0,08531	0,07809	0,10532	0,11139	0,10470	0,21547	0,16154	0,04376	0,16237	NA

Fig. S11. Assessment of TCR repertoire overlap.

Repertoire overlap was calculated using Morisita's index on AA CDR3 sequences (0 No overlap, 1 total Overlap). First lane indicates genotype of the fetal liver donor and dotted lines mice grafted with the same donor (see tab.S5 supplementary Excel file).

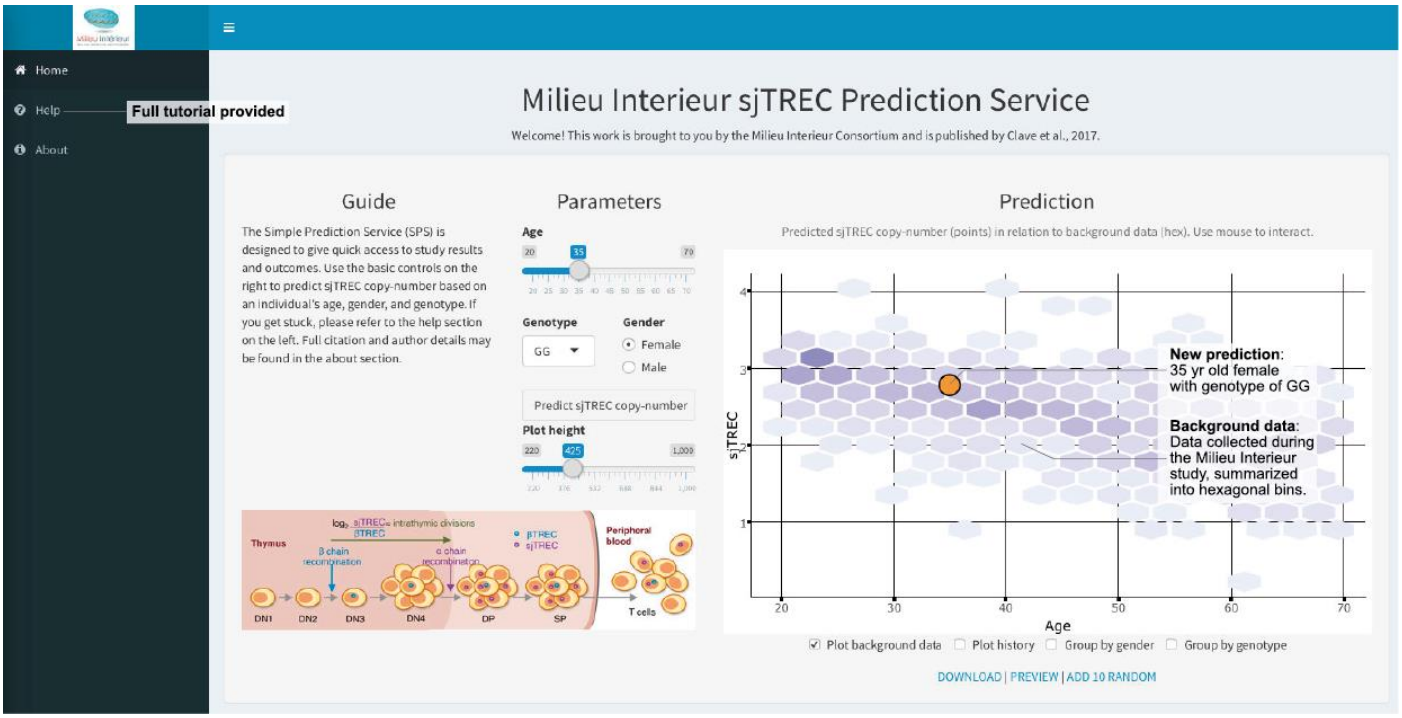

Fig. S12. Annotated screenshot of Shiny Web application showing interface for visualization and prediction.

Table S1. Demographic, medical, and lifestyle variables included in the MI study.

Variable	Category	Description
SUBJID	ID	ID
Age	Demographics	Age
OwnsHouse	Demographics	Does the subject own his/her housing?
PhysicalActivity	Demographics	Hours per week of physical activity during leisure
Sex	Demographics	Sex
LivesWithPartner	Demographics	Subject shares housing with...
LivesWithKids	Demographics	Subject shares housing with...
AbdominalCircumference	Basic physiological measurements	Abdominal circumference
BMI	Basic physiological measurements	BMI
DiastolicPressure1	Basic physiological measurements	Diastolic measure 1
DiastolicPressure2	Basic physiological measurements	Diastolic measure 2
HeartRate	Basic physiological measurements	Heart rate
Height	Basic physiological measurements	Height
SystolicPressure1	Basic physiological measurements	Systolic measure 1
SystolicPressure2	Basic physiological measurements	Systolic measure 2
Temperature	Basic physiological measurements	Ear temperature
Temperature	Basic physiological measurements	Ear temperature
Weight	Basic physiological measurements	Weight
Alcohol	Food and nutrition	Alcohol
BornInCity	Geographic origin	Born in a city with population larger than 20,000
Cholesterol	Laboratory measure	Biochemistry: Total cholesterol
CMVPositiveSerology	Laboratory measure	Serology: CMV
CRP	Laboratory measure	Biochemistry: CRP
HDL	Laboratory measure	Biochemistry: HDL
FluIgG	Laboratory measure	Serology: Log10-transformed ratio of IVA.ABS and the assay threshold value
LDL	Laboratory measure	Biochemistry: LDL
MetabolicScore	Laboratory measure	Metabolic score, estimated as described in Thomas et al., Clin Immunol 2015
Triglycerides	Laboratory measure	Biochemistry: Triglycerides
FluPositiveSerology	Laboratory measure	Serology: influenza
HadMeasles	Medical history	Childhood disease: Measles
HadRubella	Medical history	Childhood disease: Rubella
HadChickenPox	Medical history	Childhood disease: Chicken pox
HadMumps	Medical history	Childhood disease: Mumps
HadTonsillectomy	Medical history	Tonsillectomy
HadAppendicectomy	Medical history	Appendicectomy
AncestryPC1	OmniExpress SNP genotyping array	First component (5.42% of the variance explained) of a PCA of the OmniExpress array, together with reference populations (Behar et al., Nature 2010)
AncestryPC2	OmniExpress SNP genotyping array	Second component (1.63%) of a PCA of the OmniExpress array, together with reference populations (Behar et al., Nature 2010)
Inbreeding	OmniExpress SNP genotyping array	Cumulative length of Runs of Homozygosity
LowAppetite	Sleep habits, drug habits, and psychological problems	Little or too much appetite, last 2 weeks
TroubleConcentrating	Sleep habits, drug habits, and psychological problems	Difficulty concentrating on things like reading the newspaper or watching the television, last 2 weeks
TroubleSleeping	Sleep habits, drug habits, and psychological problems	Does the subject often find it difficult to fall asleep or to remain asleep?
HoursOfSleep	Sleep habits, drug habits, and psychological problems	Hours of sleep in decimal
Listless	Sleep habits, drug habits, and psychological problems	Feeling tired or having little energy, last 2 weeks
UsesCannabis	Sleep habits, drug habits, and psychological problems	Haschich
RecentBadSelfesteem	Sleep habits, drug habits, and psychological problems	Poor self-image, or you think that you are a loser or have not achieved your own expectations or those of your family., last 2 weeks
RecentlyApathetic	Sleep habits, drug habits, and psychological problems	Lack of interest or pleasure in doing things, last 2 weeks
RecentPersonalCrisis	Sleep habits, drug habits, and psychological problems	Major Negative Life event, loss of loved one etc., last 12 month
RecentlyDepressed	Sleep habits, drug habits, and psychological problems	Feeling sad, depressed or despairing, last 2 weeks
Smoking	Smoking habits	Smoking tobacco?
Employed	Socio-professional information	Steady job
Education	Socio-professional information	Level of education
DustExposure	Socio-professional information	Exposure to dust
Income	Socio-professional information	Net monthly income of the household (EUR)
VaccineHepA	Vaccination history	Vaccination against Hepatitis A
VaccineMMR	Vaccination history	Vaccination against rubella
VaccineTyphoid	Vaccination history	Vaccination against typhoid
VaccineWhoopingCough	Vaccination history	Vaccination against whooping cough
VaccineYellowFever	Vaccination history	Vaccination against yellow fever
VaccineHepB	Vaccination history	Vaccination against hepatitis B
VaccineFlu	Vaccination history	Vaccination against flu
HourOfSampling	Batch effect	Hour at which 25ml blood (FACS) sample was taken
DayOfSampling	Batch effect	Date at which V1 was done, in days since 09-01-2012
HourOfSampling	Batch effect	Hour at which 25ml blood (FACS) sample was taken
DayOfSampling	Batch effect	Date at which V2 was done, in days since 09-01-2012

Table S2. Statistics of association with sjTRECs for suggestive loci and replication at the *TCRA-TCRAD* locus.

						MI cohort					Re-genotyped candidate SNPs‡					MARTHA cohort					Meta-analysis		
Chr.	Position	Closest gene	SNP	Effect allele	Other allele	Imputed	EAF*	Effect size (β)	SE†	*P*-value	Assay ID	EAF*	Effect size (β)	SE†	*P*-value	Imputed	EAF*	Effect size (β)	SE†	*P*-value	Effect size (β)	SE†	*P*-value
2	99403933	*KIAA1211L*	rs13389386	T	C	1	0,417	0,0755	0,0183	2,9E-05	-	-	-	-	-	0	0,413	-0,0077	0,0199	7,0E-01	-	-	-
2	99543105	*KIAA1211L*	rs1607255	A	C	0	0,173	0,0887	0,0231	9,4E-05	-	-	-	-	-	0	0,187	0,0046	0,0262	8,6E-01	-	-	-
2	99965464	*EIF5B*	rs11902994	G	T	1	0,436	0,0669	0,0175	7,1E-05	-	-	-	-	-	0	0,458	-0,0104	0,0197	6,0E-01	-	-	-
4	3451960	*HGFAC*	rs1730794	T	C	1	0,312	0,0979	0,0191	3,5E-07	-	-	-	-	-	0	0,322	-0,0334	0,0241	1,7E-01	-	-	-
5	166699738	*TENM2*	rs10074548	G	A	1	0,084	-0,1322	0,0326	5,3E-05	-	-	-	-	-	0	0,069	0,0063	0,0393	8,7E-01	-	-	-
10	97325729	*SORBS1*	rs7918947	A	C	0	0,469	0,0728	0,0179	4,3E-05	-	-	-	-	-	0	0,453	0,0146	0,0197	4,6E-01	-	-	-
10	97353584	*SORBS1*	rs1270505	A	G	0	0,471	0,0753	0,0180	2,3E-05	-	-	-	-	-	0	0,455	0,0142	0,0197	4,7E-01	-	-	-
11	24642587	*LUZP2*	rs906882	A	G	0	0,347	-0,0846	0,0188	3,5E-06	-	-	-	-	-	0	0,365	0,0051	0,0209	8,1E-01	-	-	-
11	24649362	*LUZP2*	rs7124697	T	C	0	0,418	-0,0797	0,0183	1,0E-05	-	-	-	-	-	0	0,410	0,0247	0,0207	2,3E-01	-	-	-
17	64779154	*PRKCA*	rs4790904	C	T	0	0,229	0,0959	0,0212	6,2E-06	-	-	-	-	-	0	0,228	0,0093	0,0237	6,9E-01	-	-	-
14	22886319	*TCRD*	rs3811236	C	T	0	0,505	0,0737	0,0181	5,2E-05	C___2609758_10	0,506	0,075	0,018	4,1E-05	0	0,529	0,0726	0,0195	2,1E-04	0,0732	0,0133	3,5E-08
14	22886777	*TCRD*	rs2301199	G	C	1	0,505	0,0727	0,0182	6,9E-05	-	-	-	-	-	1	0,529	0,0726	0,0195	2,1E-04	0,0727	0,0133	4,7E-08
14	22886992	*TCRD*	rs2301200	C	T	1	0,505	0,0727	0,0182	6,9E-05	-	-	-	-	-	1	0,529	0,0726	0,0195	2,1E-04	0,0727	0,0133	4,7E-08
14	22887459	*TCRD*	rs3216790	CG	C	1	0,505	0,0721	0,0182	8,1E-05	-	-	-	-	-	1	0,529	0,0726	0,0195	2,1E-04	0,0723	0,0133	5,5E-08
14	22888814	*TCRD*	rs62762262	TAGTC	T	1	0,505	0,0740	0,0182	5,2E-05	-	-	-	-	-	1	0,527	0,0701	0,0195	3,4E-04	0,0722	0,0133	5,8E-08
14	22890411	*TCRD*	rs2331618	A	G	1	0,504	0,0736	0,0183	6,0E-05	-	-	-	-	-	1	0,526	0,0734	0,0195	1,8E-04	0,0735	0,0133	3,4E-08
14	22890959	*TCRD*	rs1109130	G	A	1	0,505	0,0728	0,0182	7,0E-05	-	-	-	-	-	1	0,529	0,0726	0,0195	2,1E-04	0,0727	0,0133	4,7E-08
14	22891883	*TCRD*	rs2072616	G	C	1	0,505	0,0728	0,0182	7,0E-05	-	-	-	-	-	1	0,529	0,0726	0,0195	2,1E-04	0,0727	0,0133	4,7E-08
14	22892461	*TCRD*	rs8012481	G	A	1	0,505	0,0728	0,0182	7,0E-05	-	-	-	-	-	1	0,529	0,0726	0,0195	2,1E-04	0,0727	0,0133	4,7E-08
14	22899847	*TCRD*	rs8014525	A	G	1	0,456	0,0700	0,0180	1,1E-04	-	-	-	-	-	1	0,473	0,0493	0,0195	1,2E-02	0,0605	0,0132	4,9E-06
14	22900719	*TCRD*	rs6572448	T	A	1	0,506	0,0727	0,0181	6,1E-05	-	-	-	-	-	1	0,527	0,0680	0,0195	5,3E-04	0,0706	0,0133	1,0E-07
14	22902506	*TCRD*	rs741712	A	G	1	0,455	0,0720	0,0179	6,1E-05	-	-	-	-	-	1	0,474	0,0499	0,0195	1,1E-02	0,0619	0,0132	2,7E-06
14	22902753	*TCRD*	rs2222940	C	T	1	0,455	0,0708	0,0179	8,2E-05	-	-	-	-	-	1	0,471	0,0499	0,0196	1,1E-02	0,0613	0,0132	3,5E-06
14	22903462	*TCRD*	rs2141988	A	G	0	0,506	0,0731	0,0180	5,1E-05	C__15828563_10	0,505	0,074	0,018	4,3E-05	0	0,525	0,0661	0,0195	7,6E-04	0,0699	0,0132	1,3E-07
14	22904492	*TCRD*	rs916052	A	G	1	0,508	0,0729	0,0180	5,4E-05	-	-	-	-	-	1	0,525	0,0661	0,0195	7,6E-04	0,0698	0,0132	1,3E-07
14	22904588	*TCRD*	rs916053	C	A	1	0,458	0,0709	0,0179	8,3E-05	-	-	-	-	-	1	0,471	0,0479	0,0196	1,5E-02	0,0604	0,0132	4,9E-06
14	22905065	*TCRD*	rs8021297	T	C	1	0,507	0,0750	0,0180	3,3E-05	-	-	-	-	-	0	0,523	0,0649	0,0195	9,4E-04	0,0703	0,0132	1,0E-07
14	22906945	*TCRD*	rs7150030	T	C	1	0,458	0,0718	0,0179	6,7E-05	-	-	-	-	-	1	0,471	0,0479	0,0196	1,5E-02	0,0609	0,0132	4,1E-06
14	22907068	*TCRD*	rs11157596	G	A	1	0,458	0,0718	0,0179	6,7E-05	-	-	-	-	-	0	0,471	0,0479	0,0196	1,5E-02	0,0609	0,0132	4,1E-06
14	22907318	*TCRD*	rs7492759	G	A	1	0,507	0,0735	0,0180	4,8E-05	-	-	-	-	-	1	0,526	0,0657	0,0195	8,0E-04	0,0699	0,0132	1,2E-07
14	22909221	*TCRD*	rs2204984	A	G	1	0,505	0,0765	0,0181	2,6E-05	C__16160754_10	0,504	0,075	0,018	3,6E-05	1	0,522	0,0654	0,0196	8,7E-04	0,0714	0,0133	7,7E-08
14	22909825	*TCRD*	rs71115550	C	CATAAA	1	0,505	0,0782	0,0182	1,8E-05	-	-	-	-	-	1	0,523	0,0646	0,0195	9,9E-04	0,0719	0,0133	6,3E-08
14	22909827	*TCRD*	rs201497432	T	TA	1	0,521	0,0726	0,0184	8,8E-05	-	-	-	-	-	1	0,548	0,0748	0,0196	1,5E-04	0,0737	0,0134	4,3E-08
14	22910504	*TCRD*	rs2331620	C	T	1	0,457	0,0733	0,0180	5,1E-05	-	-	-	-	-	1	0,473	0,0464	0,0196	1,8E-02	0,0610	0,0134	5,5E-06
14	22910826	*TCRD*	rs11623632	A	G	1	0,456	0,0733	0,0180	5,3E-05	-	-	-	-	-	1	0,471	0,0467	0,0196	1,7E-02	0,0610	0,0133	4,2E-06
14	22911702	*TCRD*	rs11851663	G	A	1	0,455	0,0726	0,0180	6,2E-05	-	-	-	-	-	1	0,471	0,0474	0,0196	1,6E-02	0,0610	0,0133	4,3E-06
14	22913575	*TCRD*	rs8013419	A	G	1	0,490	0,0811	0,0177	5,1E-06	C___2609804_10	0,490	0,081	0,018	5,1E-06	1	0,515	0,0655	0,0196	8,7E-04	0,0741	0,0131	1,6E-08
14	22913994	*TCRD*	rs10873018	G	A	0	0,488	0,0803	0,0177	6,7E-06	C___2609805_20	0,488	0,080	0,018	6,7E-06	1	0,514	0,0668	0,0196	8,6E-05	0,0742	0,0131	1,6E-08
14	22914747	*TCRD*	rs10137305	G	C	1	0,328	0,0781	0,0191	4,7E-05	-	-	-	-	-	1	0,345	0,0503	0,0209	1,6E-02	0,0654	0,0141	3,4E-06
14	22915864	*TCRD*	rs12147006	A	C	1	0,490	0,0811	0,0177	5,1E-06	C__11721207_10	0,490	0,080	0,018	6,1E-06	0	0,518	0,0670	0,0196	6,5E-04	0,0748	0,0131	1,2E-08
14	22917633	*TCRD*	rs2204985	G	A	1	0,489	0,0794	0,0177	8,5E-06	C___2609811_10	0,488	0,078	0,018	1,2E-05	0	0,515	0,0669	0,0195	9,7E-05	0,0737	0,0131	1,9E-08
14	22919581	*TCRD*	rs11844715	C	T	1	0,485	0,0774	0,0179	1,6E-05	C___2609812_10	0,482	0,070	0,018	8,6E-05	1	0,513	0,0641	0,0196	1,1E-03	0,0714	0,0132	6,3E-08
14	22921202	*TCRD*	rs7144549	C	G	1	0,323	0,0765	0,0193	8,3E-05	-	-	-	-	-	1	0,345	0,0516	0,0208	1,3E-02	0,0649	0,0142	4,6E-06
14	22925352	*TCRD*	rs1882704	G	A	0	0,260	0,0843	0,0207	5,2E-05	-	-	-	-	-	0	0,275	0,0520	0,0219	1,8E-02	0,0689	0,0161	2,0E-05

*Effect allele frequency.

†Standard error of the effect size.

‡A selection of eight SNPs were re-genotyped on Biomark in the Milieu Intérieur primary cohort, to confirm association at imputed SNPs.

Table S3. *TCRA-TCRD* next-generation sequencing data.

Sample Name	Mouse Recipient Sex	Human Donor Sex	rs2204985	locus	Age (weeks)	total rearrangements	productive rearrangements	productive clonality	max productive frequency	Shanon Equitability	Inverse Simpson
45605T3RD	Female	Men	AA	TCRAD	19	2 945 927	897 631	0.0417	0.038182	0.98428115051	369 487
45609T3RD	Female	Men	AA	TCRAD	19	216 585	76 022	0.0664	0.401470	0.970865080466	14 895
45615T3RD	Male	Men	AA	TCRAD	19	270 333	88 760	0.0558	0.266205	0.982650106856	30 808
45742T3RD	Male	Men	AA	TCRAD	19	872 887	272 706	0.0362	0.058768	0.98924927183	90 236
47139T5TH	Female	Men	AA	TCRAD	9	5 959	2 118	0.0595	0.471729	0.995952765477	1 941
47140T5TH	Female	Men	AA	TCRAD	9	68 065	21 518	0.0585	0.336460	0.986729670136	11 682
47143T5TH	Male	Men	AA	TCRAD	9	243 539	75 292	0.1143	0.361357	0.932140757822	6 149
47145T5TH	Female	Men	AA	TCRAD	9	49 935	22 221	0.0665	0.682569	0.97468615593	6 829
47692T13TH	Female	Women	GG	TCRAD	14	149 212	45 726	0.0430	0.052720	0.997330235266	42 327
47693T13TH	Female	Women	GG	TCRAD	14	18 656	6 186	0.0533	0.421263	0.989667839494	4 727
47694T13TH	Female	Women	GG	TCRAD	14	224 288	68 310	0.0407	0.020894	0.995161945481	59 381
47695T13TH	Female	Women	GG	TCRAD	14	211 684	64 792	0.0419	0.030033	0.994933552869	55 786
47697T13TH	Female	Women	GG	TCRAD	14	327 052	99 201	0.0395	0.036163	0.995302652077	85 880
47702T13TH	Male	Women	GG	TCRAD	14	372 808	114 460	0.0396	0.058301	0.993467458253	91 406
47707T13TH	Male	Women	GG	TCRAD	14	340 604	104 208	0.0422	0.112624	0.990646547876	71 032
62T14TH	Female	Men	GG	TCRAD	16	3 548 800	1 081 567	0.0394	0.025733	0.989872263019	601 867
64T14TH	Female	Men	GG	TCRAD	16	616 551	191 175	0.0366	0.024099	0.994079147466	154 894
65T14TH	Female	Men	GG	TCRAD	16	66 190	20 664	0.0493	0.272425	0.99531082054	17 642
68T14TH	Male	Men	GG	TCRAD	16	647 721	200 035	0.0365	0.027283	0.994379025079	163 730
69T14TH	Male	Men	GG	TCRAD	16	435 543	135 036	0.0412	0.030474	0.993517174734	109 246
Median AA					14	230 062	75 657	0.0590	0.348908	0.983465628683	13 289
Median GG					14	333 828	101 705	0.0409	0.033318	0.9942290862725	78 456
*p (*Mann-Whitney)					*1.000*	*0.521*	*0.624*	*0.016*	*0.005*	*0.003*	*0.069*

Productive clonality is calculated by normalizing productive entropy using the total number of unique productive rearrangements and subtracting the result from 1. The max productive frequency is the maximum productive frequency value found within a sample. Shannon Equitability (or Pielou Evenness) is a normalized Shannon's entropy.

Table S4. Primers and probes used for sex determination and TREC quantification.

			Sequence	Preamp (nM)	2x Assay Biomark (µM)	PCR [final] (nM)
TREC Assay	Primers	**sjTREC-F**	CACATCCCTTTCAACCATGCT	100	8	400
		sjTREC-R	GCCAGCTGCAGGGTTTAGG	100	8	400
		Jα58TREC-F	GCTTCGTTAAGCTGATGCCAC			400
		Jα58TREC-R	GTGCCAGCTGCAGGGTTTA			200
		ALB-F	GCTGTCATCTCTTGTGGGCTGT	100	4	20
		ALB-R	ACTCATGGGAGCTGCTGGTTC	100	4	20
		1.1	TGTCCTCCATCCTAGCCAGG	200	4	
		1.2	TCCGTCACAGGGAAAAGTGG	200	4	
		1.3	TGTCCCTGTGAGGGAAGAGTT	200	4	
		1.4	TGGACTTGGGGAGGCAGGA	200	4	
		1.5	CTCATAAAATGTGGGTCAGTGGA	200	4	
		1.6	TGAATCCAGGCAGAGAAAGG	200	4	
		2.1	CCAGCTAACTCGAGACAGGAA	200	4	
		2.2	GAACCCTGTTCTTAGGGGAGT	200	4	
		2.3	TGAGAGGGGCTGTGCTGAGA	200	4	
		2.4	AAGCGGGGGCTCCCGCTGAA	200	4	
		Db1	TGTGACCCAGGAGGAAAGAAG	400	4	
		Db2	GGACCAGCCCCAGAGAA	400	4	
	Probes	**sjTREC**	*FAM*-ACACCTCTGGTTTTTGTAAAGGTGCCCACT-*TAMRA*		2	
		sjTREC-LNA	*FAM*-ACACCTCTGGTTTTTGTAAAGGTG-*EclipseDQ*		2	200
		Jα58TREC	*FAM*-TGCAAAGCCCCTCAGCACAG-*NFQ-MGB*			100
		Alb	*FAM*-CCTGTCATGCCCACACAAATCTCTCC-*TAMRA*		2	
		Alb-VIC	*VIC*-CCTGTCATGCCCACACAAATCTCTCC-*TAMRA*			100
		Db1	*FAM*-CAAAAACCTCTCTGGCGGTCCCAAC-*TAMRA*		8	
		Db2	*FAM*-CCCACCCAGCTCAGGGAATGCA-*TAMRA*		4	
Sexing	H.	**hSex2-F**	AAGTGCCCTCTTGCACATA			200
		hSex2-R	CTCGACTTAAACTTCTTCCC			200
	Mouse	**MSRY-F**	TGGGACTGGTGACAATTGTC			200
		MSRY-R	TCTCGACGTGTGGACATGAG			200
		mIL3-F	GGGACTCCAAGCTTCAATCA			200
		mIL3-R	AACGAAAAGAAGGAGGAGGT			200

All primers and probes are from Eurogentec (Paris, France) except Alb-VIC probe (ThermoFisher Scientific). Underlined nucleotides represent Locked Nucleic Acids (LNA). All primers are together in preamplification. H. = Human.