Table S1. Genes within the IL-2 and IL-7 Signalling Pathways, Related to Table 1

Chr	Gene	Chr	Gene	Chr	Gene
1	*LCK*	6	*BACH2*	13	*FOXO1*
1	*JUN*	6	*FYN*	14	*FOS*
1	*JAK1*	6	*FOXO3A*	14	*BCL11B*
1	*ILF2*	7	*IKZF1*	14	*AKT1*
1	*SHC1*	7	*PIK3CG*	15	*GABPB1*
1	*POU2F1*	7	*TWIST1*	15	*MAP2K1*
1	*TNFSF4*	8	*EGR3*	15	*FURIN*
1	*PTPRC*	8	*PTK2B*	16	*SOCS1*
1	*IL12Rb2*	8	*IL7*	16	*MAPK3*
1	*TRAF5*	8	*LYN*	17	*STAT5A*
2	*SOS1*	9	*SYK*	17	*GRB2*
2	*REL*	9	*TNFSF8*	17	*SOCS3*
2	*XRCC5*	10	*IL2RA*	17	*FOXK2*
2	*CD28*	10	*MAPK8*	18	*BCL2*
2	*CD8a/b*	10	*EGR2*	18	*NFATC1*
3	*RAF1*	10	*NFKB2*	19	*ILF3*
3	*RHOA*	10	*ZEB1*	19	*JAK3*
3	*EVI1*	10	*IL15RA*	19	*RELB*
3	*CBLB*	11	*HRAS*	19	*ELSPBP1*
4	*NFKB1*	11	*SPI1*	19	*KLF2*
4	*IL2*	11	*RELA*	19	*IL12RB1*
4	*IL15*	11	*POU2AF1*	19	*TNFSF9*
4	*TLR2*	11	*CBL*	20	*NFATC2*
4	*TLR3*	12	*NAB2*	21	*GABPA*
5	*IL7R*	12	*CD27*	21	*ICOSLG*
5	*EGR1*	12	*CD69*	22	*IL2RB*
5	*ITK*	12	*KLRK1*	22	*XRCC6*

Genes involved in the IL-2 or IL-7 signalling pathways were identified using published literature and publically available pathway libraries (BioCarta, www.biocarta.com; KEGG, Kyoto encyclopedia of genes and genomes)

HUGO gene symbols shown

Chr, Chromosome

Table S3. rs12212067 Genotype Counts for Each Disease Cohort, Related to Tables 1 and 2

	Major Allele Homozygotes (T/T)	Heterozygotes (T/G)	Minor Allele Homozygotes (G/G)	MAF
CD primary cohort (GWAS) - indolent disease	294	88	7	0.131
CD primary cohort (GWAS) - aggressive disease	555	109	2	0.085
CD replication cohort 1 - indolent disease	275	72	10	0.129
CD replication cohort 1 - aggressive disease	285	53	6	0.094
CD replication cohort 2 - indolent disease	69	19	1	0.118
CD replication cohort 2 - aggressive disease	118	17	1	0.070
CD replication cohort 3 - indolent disease	249	76	6	0.133
CD replication cohort 3 - aggressive disease	562	105	3	0.083
Norfolk Arthritis Registry cohort	875	185	11	0.097
Early Rheumatoid Arthritis Study cohort	307	75	0	0.098
Kenyan Severe Malaria	765	494	81	0.245
Kenyan Controls	1481	824	131	0.223
Vietnamese Severe Malaria	126	83	14	0.249
Vietnamese Controls	424	200	29	0.198

Genotype counts at rs12212067 for each of the CD, RA and malaria cohorts.

MAF, Minor Allele Frequency

Table S4. Linkage Disequilibrium around rs12212067, Related to Figure 1

RefSNP	Genome Coordinates	Genomic Location	Location (Relative to rs12212067)	Distance from rs12212067 (bp)	r^2
rs12200544	chr6:108971057	Promoter	Upstream	116832	0.84
rs768024	chr6:108982801	Promoter	Upstream	105088	0.84
rs7746906	chr6:108985506	Promoter	Upstream	102383	0.84
Unannotated	chr6:108986522	Promoter	Upstream	101367	0.57
Unannotated	chr6:108992697	Intronic	Upstream	95192	0.73
Unannotated	chr6:108996111	Intronic	Upstream	91778	0.73
rs12200646	chr6:109004354	Intronic	Upstream	83535	0.54
rs58157973	chr6:109027850	Intronic	Upstream	60039	0.84
rs12196602	chr6:109029025	Intronic	Upstream	58864	0.84
rs12192569	chr6:109040350	Intronic	Upstream	47539	1
rs12192758	chr6:109040501	Intronic	Upstream	47388	1
rs72942519	chr6:109041658	Intronic	Upstream	46231	1
rs12202234	chr6:109045776	Intronic	Upstream	42113	1
rs7341233	chr6:109046973	Intronic	Upstream	40916	0.92
rs17598747	chr6:109048086	Intronic	Upstream	39803	1
rs17069665	chr6:109048161	Intronic	Upstream	39728	0.82
rs12213895	chr6:109058661	Intronic	Upstream	29228	1
rs73763154	chr6:109062554	Intronic	Upstream	25335	1
rs72942595	chr6:109064125	Intronic	Upstream	23764	1
rs17310529	chr6:109064223	Intronic	Upstream	23666	1
rs12202209	chr6:109066718	Intronic	Upstream	21171	1
rs12207868	chr6:109071578	Intronic	Upstream	16311	1
rs12202049	chr6:109075370	Intronic	Upstream	12519	1
rs12209092	chr6:109077373	Intronic	Upstream	10516	1
rs12197634	chr6:109078349	Intronic	Upstream	9540	1
rs11153120	chr6:109081146	Intronic	Upstream	6743	0.91
rs12154031	chr6:109081271	Intronic	Upstream	6618	1
rs12203787	chr6:109082157	Intronic	Upstream	5732	1
rs12203834	chr6:109082255	Intronic	Upstream	5634	1
rs7772662	chr6:109083800	Intronic	Upstream	4089	1
rs11153121	chr6:109088975	Intronic	Downstream	1086	1
Unannotated	chr6:109089855	Intronic	Downstream	1966	0.82
rs12196996	chr6:109090760	Intronic	Downstream	2871	0.82
rs3778586	chr6:109093299	Intronic	Downstream	5410	1
rs3800226	chr6:109095288	Intronic	Downstream	7399	1
Unannotated	chr6:109096900	Intronic	Downstream	9011	0.57
rs73763159	chr6:109098379	Intronic	Downstream	10490	1
rs61192764	chr6:109101880	Intronic	Downstream	13991	0.70
rs3800230	chr6:109104821	Intronic	Downstream	16932	0.70
rs3800232	chr6:109105646	Intronic	Downstream	17757	0.70

Publically available sequencing data from Pilot 1 of the 1000 Genomes Project was used to identify all SNPs in LD (r^2>0.5) with rs12212067.

Genome co-ordinates were based upon hg18 build.

bp, base pairs.

None of the 45 coding SNPs that have been described within *FOXO3A* (source; dbSNP) were in any detectable LD with rs12212067 (r^2 > 0.001)

Table S6. Predicted FOXO3 Binding Sites within TGFβ1 Promoter, Related to Figure 4

Start Site (Position)	Stop Site (Position)	Log-Odds Score	p Value	Matched Sequence
392	399	6.01	0.001	TGTAAATT
130	137	1.74	0.004	TGCACACA
959	966	0.74	0.006	TCAAAACC

Sites at which FOXO3 would be predicted to bind within the 1000bp upstream of the *TGFβ1* start site were identified using the "Find Individual Motif Occurrences" (FIMO) tool in the MEME suite. The start and stop positions, log-odds score, p value and matched sequence are shown for sites at which FOXO3 would be predicted to bind.

Table S7. rs12212067: T>G associates with a Milder Course of RA, Related to Table 2

Cohort		Time 0	Year 1	Year 2	Year 3	Year 4	Year 5	Larsen Score Coefficient (95% CI)	p
NOAR cohort (1071 patients, 1696 x-rays)	n	306	722	242	0	0	423		
	MAF mild	0.090	0.101	0.140	-	-	0.138	-1.6 (-2.8; -0.4)	0.0041
	MAF aggressive	0.090	0.097	0.010	-	-	0.093	-	-
ERAS cohort (382 patients, 1915 x-rays)	n	366	339	349	346	210	305		
	MAF mild	0.091	0.104	0.115	0.130	0.130	0.098	-1.5 (-2.7; -0.4)	0.0043
	MAF aggressive	0.105	0.092	0.094	0.086	0.096	0.096	-	-

NOAR; Norfolk Arthritis Register

ERAS; Early Rheumatoid Arthritis Study

Mild RA; complete absence of radiological damage in hands and feet, including joint space narrowing (Larsen score = 0)

Aggressive RA; presence of radiological damage in hands or feet (Larsen score > 0)

One-tailed p values calculated using Generalised Linear Latent and Mixed Modeling to consider both disease duration and Larsen score as continuous variables to enable inclusion of multiple records per patient over time. As Larsen score was considered a continuous variable and analysed as part of a longitudinal regression analysis, the mild and aggressive groups shown here are illustrative of the differences in MAF but were not directly compared for the statistical analysis.

Larsen score coefficient; change in Larsen score per copy of minor (G) allele carried

MAF, minor allele frequency