Supplementary material

Table S1. Genotypes and alleles distribution.

IRF5 SNPs		Number of carriers (%)				
rs4728142	**Allele/Genotype**	**RA RF-positive, n=82**	**RA RF-negative, n=40**	**RA overall, n=122**	**Controls, n=24**	**Subjects in a methylation study, n=76**
	MAF (A allele)	46.3%	58.8%	50.4%	45.8%	48%
	GG	26; 31.7%	6; 15%	32; 26.2%	7; 29.2%	24; 31.6%
	GA	36; 43.9%	21; 52.5%	57; 46.7%	12; 50%	31; 40.8%
	AA	20; 24.4%	13; 32.5%	33; 27%	5; 20.8%	21; 27.6%
rs10488631	**Allele/Genotype**					
	MAF (C allele)	14.6%	23.4%	17.6%	12.5%	14.5%
	TT	59; 72%	23; 57.5%	82; 67.2%	18; 75%	54; 71.1%
	TC	22; 26.8%	15; 37.5%	37; 30.3%	6; 25%	22; 28.9%
	CC	1; 1.2%	2; 5%	3; 2.5%	0	0

Abbreviations: *IRF5*; interferon regulatory factor V; MAF, minor allele frequency; RA, rheumatoid arthritis patients; RF, rheumatoid factor; SNP, single nucleotide polymorphism.

Table S2. Distribution of genotypes in a different genetic models between patients and controls.

| SNP	Model	RA n=122	Controls, n=24	OR (95% CI)	P value	Fisher's exact P value
rs4728142	**Codominant 1** GG vs.GA	32; 26.2% vs. 57; 46.7%	7; 29.2% vs. 12; 50%	1.13 (0.4-3.25)	0.81	0.57
	Codominant 2 GG vs. AA	32; 26.2% vs. 33; 27%	7 (29.2) vs. 5; 20.8%	1.2 (0.36-4.05	0.76	0.40
	Dominant GA+AA vs. GG	90; 73.8% vs. 32; 26.2%	17 (70.8) vs. 7; 29.2%	0.86 (0.33-2.29)	0.77	0.47
	Overdominant GA vs. GG+AA	57; 46.7% vs. 65; 53.3%	12 (50) vs. 12; 50%	0.96 (0.4-2.34)	0.94	0.47
	Recessive AA vs. GA+GG	33; 27% vs. 89; 73%	5 (20.8) vs. 19; 79.2%	0.9 (0.33-2.48)	0.85	0.36
rs10488631						
	Codominant 1 TT vs.TC	82; 67.1% vs. 37; 30.3%	18; 75% vs. 6; 25%	1.35 (0.49-3.72)	0.55	0.37
	Codominant 2 TT vs. CC	82; 67.1% vs. 3; 2.5%	18; 75% vs. 0	1.2 (0.36-4.05)	0.76	0.56
	Dominant TC+CC vs. TT	40; 32.8% vs. 82; 67.1%	6; 25% vs. 18; 75%	0.68 (0.25-1.85)	0.43	0.31
	Overdominant TC vs. TT+CC	37; 30.3% vs. 85; 69.7%	6; 25% vs. 18; 75%	0.77 (0.28-2.1)	0.60	0.40
	Recessive CC vs. TT+TC	3; 2.5% vs. 119; 97.5%	0 vs. 24; 100%	N/A		0.58

Data are presented as number; %. Abbreviations: CI, confidence interval; OR, odds ratio. For other please refer to Table 2 in the manuscript or Table S1 in supplementary file.

Table S3. Distribution of genotypes in a different genetic models between seropositive and seronegative patients.

SNP	Model	RF positive, n=82	RF negative, n=40	OR (95% CI)	P value	Fisher's exact P value
rs4728142	**Codominant 1** GG vs.GA	26; 31.7% vs. 36; 43.9%	6; 15% vs. 21; 52.5%	2.53 (0.88-7.24)	0.068	0.06
	Codominant 2 GG vs. AA	26; 31.7% vs. 20; 24.4%	6; 15% vs. 13; 32.5%	2.81 (0.89-8.91)	0.065	0.59
	Dominant GA+AA vs. GG	56; 68.3% vs. 26; 31.7%	34; 85% vs. 6; 15%	0.38 (0.14-1.03)	**0.0416**	**0.037**
	Overdominant GA vs. GG+AA	36; 43.9% vs. 46; 56.1%	21; 52.5% vs. 19; 47.5%	0.71 (0.33-1.52)	0.37	0.24
	Recessive AA vs. GA+GG	20; 24.4% vs. 62; 75.6%	13; 32.5% vs. 27; 67.5%	0.67(0.29-1.55)	0.35	0.23
rs10488631						
	Codominant 1 TT vs.TC	59; 72% vs. 22; 26.8%	23; 57.5% vs. 15; 37.5%	1.74 (0.77-3.98)	0.18	0.13
	Codominant 2 TT vs. CC	59; 72% vs. 1; 1.2%	23; 57.5% vs. 2; 5%	5.13 (0.42-61.5)	0.17	0.21
	Dominant TC+CC vs. TT	23; 28% vs. 59; 72%	17; 42.5% vs. 23; 57.5%	0.53 (0.24-1.17)	0.11	0.08
	Overdominant TC vs. TT+CC	22; 26.8% vs. 60; 73.2%	15; 37.5% vs. 25; 62.5%	0.61 (0.27-1.38)	0.23	0.16
	Recessive CC vs. TT+TC	1; 1.2% vs. 81; 98.8%	2; 5% vs. 38; 95%	0.23 (0.02-2.73)	0.22	0.25

For abbreviations and data presentation please refer to Table S2 in the supplementary file. A P value <0.05 is in bold.

Figure S1. Graphical representation of the studied locus.
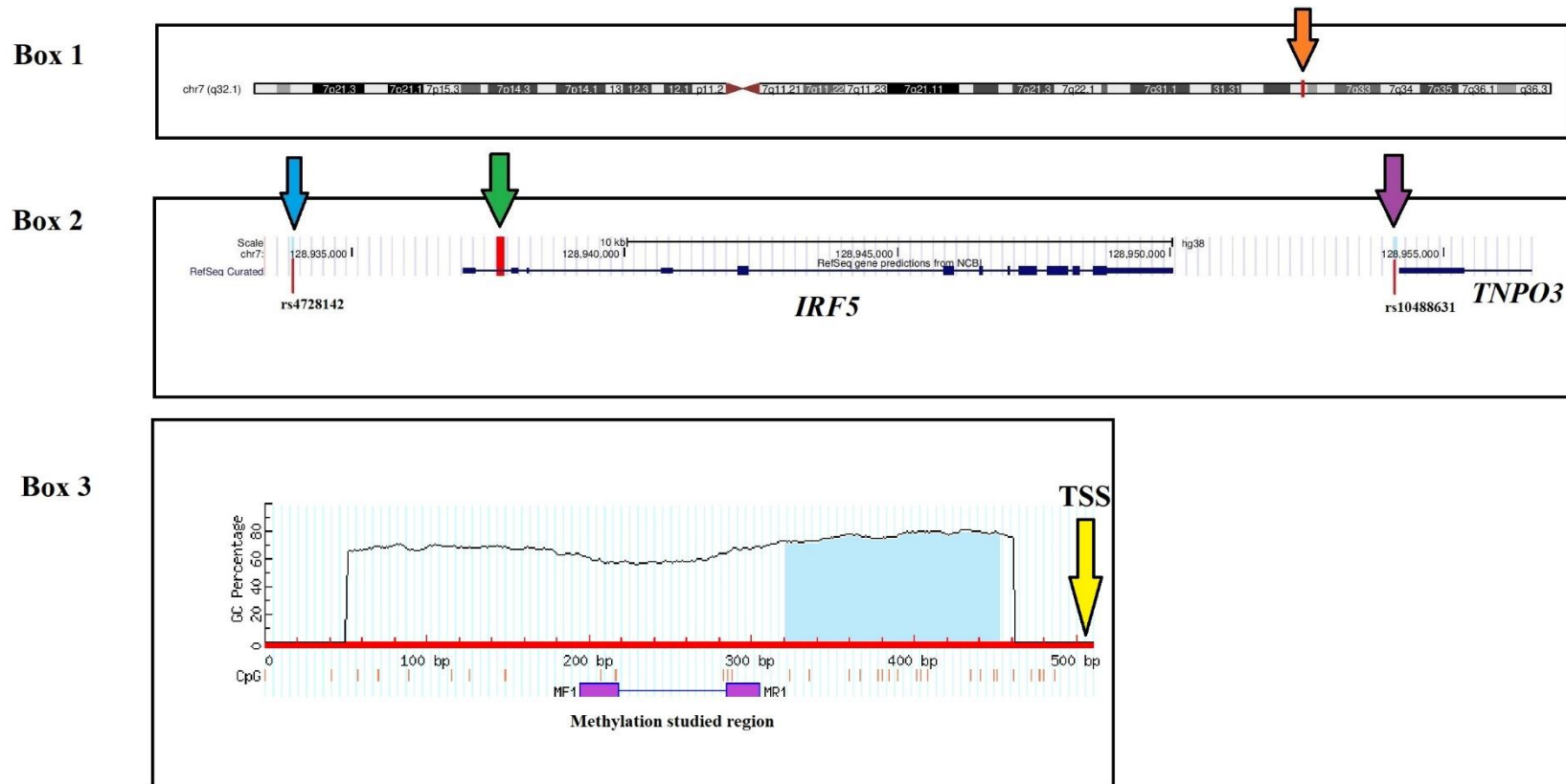
Box 1: The orange arrow shows the location of *IRF5* gene on the chromosome 7. **Box 2**: The blue arrow shows the position of *IRF5* rs4728142, the green arrow shows the methylation studied region and the purple shows *IRF5* rs10488631. **Box 3** represents the methylation studied region in details. The purple boxes represent the primers. Images in Boxes 1 and 2 were adopted from UCSC Genome Browser (available from: *http://genome.ucsc.edu* ; for more details, please refer Kent et al[1]. Abbreviations: chr7, chromosome 7; kb, kilobases; MF1- IRF5_ methyl_sense forward; MF2, IRF5_ methyl_antisense; NCBI, National Center for Biotechnology Information; TNPO3, transportin 3 gene; TSS, transcription start site. For other please refer to supplementary Table S1.

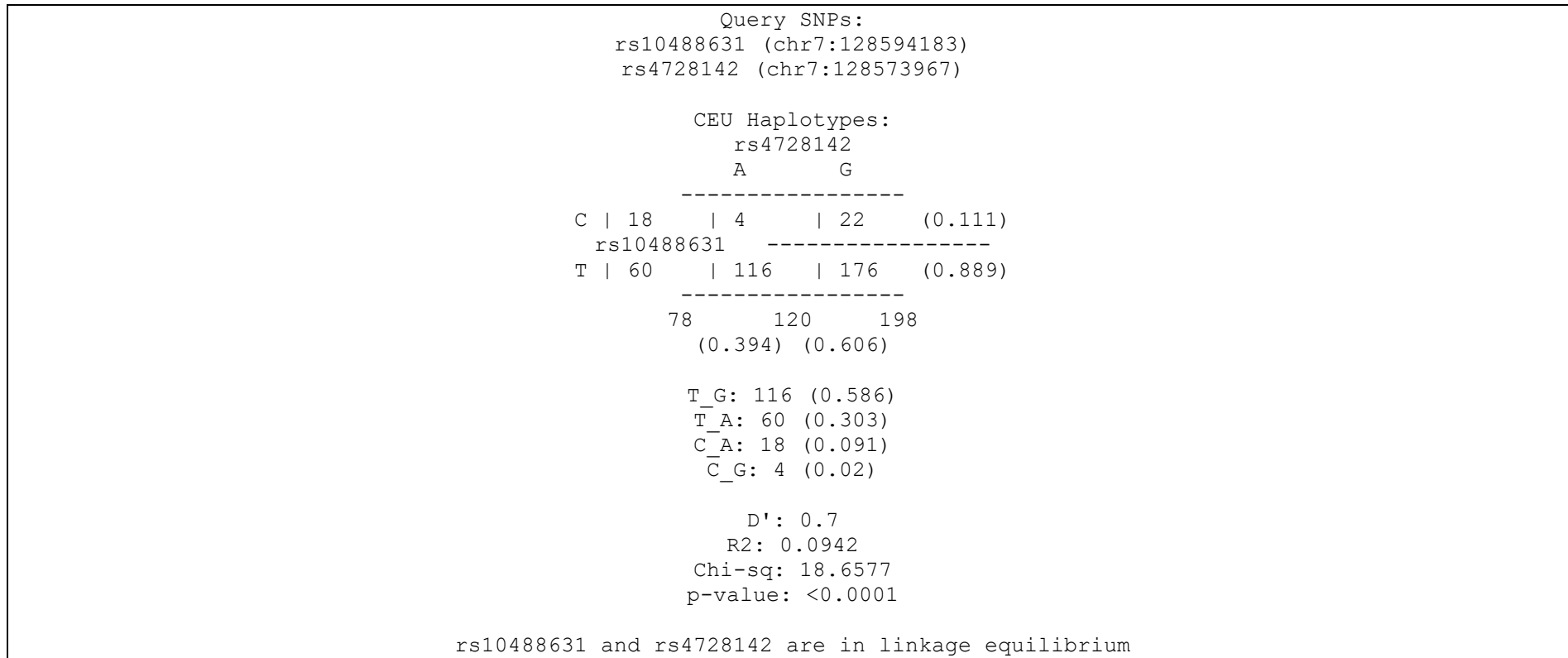Figure S2. Linkage equilibrium between rs4728142 and rs10488631.

```
                         Query SNPs:
                   rs10488631 (chr7:128594183)
                    rs4728142 (chr7:128573967)


                        CEU Haplotypes:
                          rs4728142
                          A       G
                       -----------------
           C | 18     | 4     | 22     (0.111)
             rs10488631    -----------------
           T | 60     | 116   | 176    (0.889)
                       -----------------
                        78      120     198
                       (0.394) (0.606)


                       T_G: 116 (0.586)
                       T_A: 60 (0.303)
                       C_A: 18 (0.091)
                        C_G: 4 (0.02)


                          D': 0.7
                         R2: 0.0942
                       Chi-sq: 18.6577
                       p-value: <0.0001

        rs10488631 and rs4728142 are in linkage equilibrium
```

Figure S3. Methylation status between rheumatoid arthritis patients and controls.
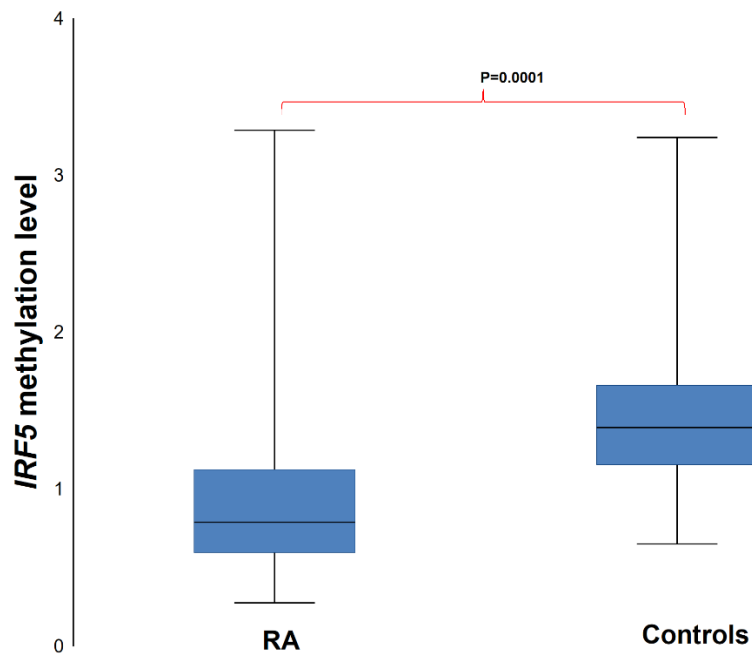
Box chart represents median with interquartile range. Lower and upper extreme were marked by whiskers. Abbreviations: please refer to supplementary Table S1.

Figure S4. Correlation between disease activity and methylation status between rheumatoid arthritis patients.
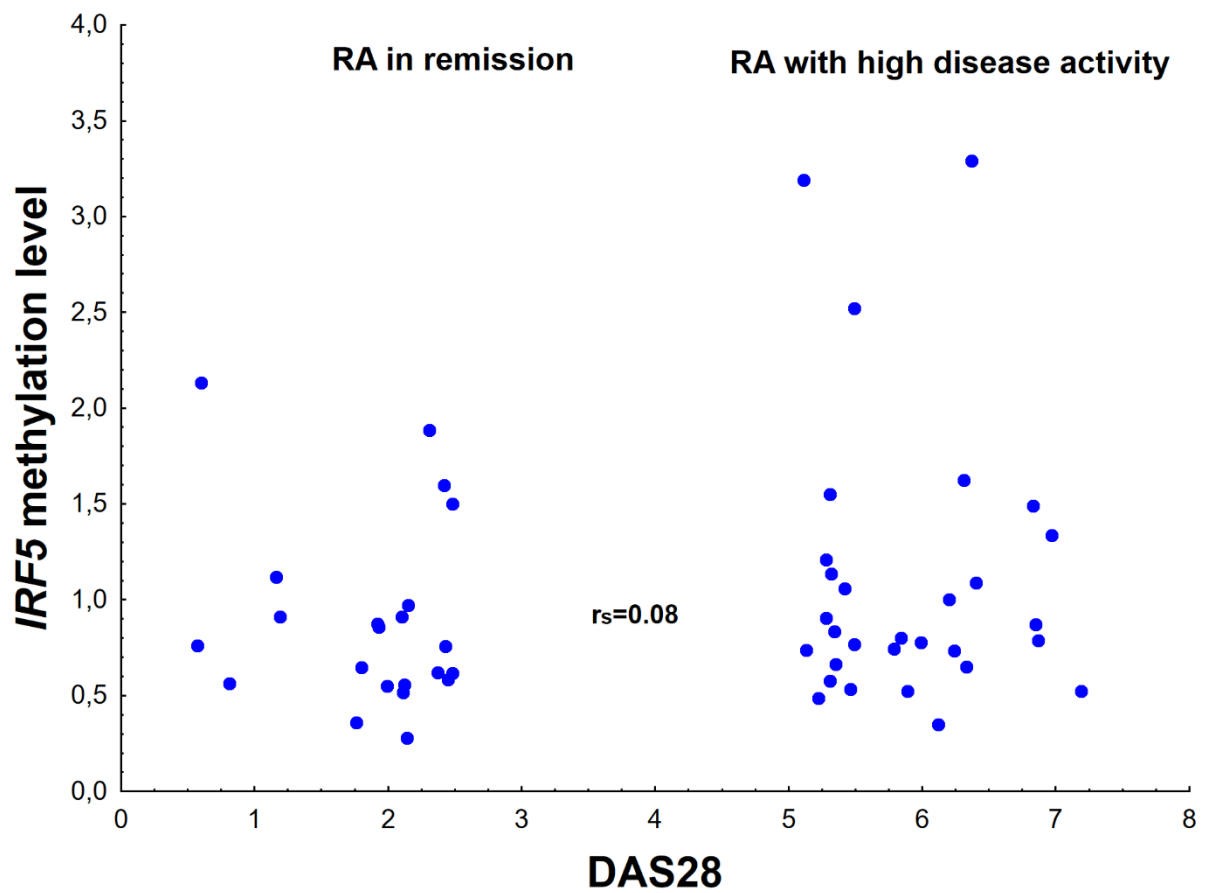
Abbreviations: DAS28, disease activity score. For other please refer to supplementary Table S1.

References

1   Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. Genome Res. 2002; 12: 996-1006.