Extended Data Table 1 | Characteristics of the study cohorts

a

Study stage	Cohort	Ethnicity	Geographical origin	No. subjects			RA case sero-positivity
				Cases	Controls	Total	
GWAS meta-analysis (Stage 1)	BRASS	European	North America	483	1,631	2,114	100% CCP+
	CANADA		Canada	589	1,554	2,143	100% CCP+
	EIRA		Sweden	1,097	1,044	2,141	100% CCP+
	NARAC1		North America	863	1,191	2,054	100% CCP+
	NARAC2		North America	896	6,603	7,499	100% CCP+
	WTCCC		United Kingdom	1,520	10,507	12,027	100% CCP+ or RF+
	RACI-UK		United Kingdom	1,645	6,082	7,727	100% CCP+
	RACI-US		North America	997	2,132	3,129	100% CCP+
	RACI-SE-E		Sweden	740	1,117	1,857	100% CCP+
	RACI-SE-U		Sweden	522	962	1,484	100% CCP+
	RACI-NL		Netherland	303	2,001	2,304	100% CCP+
	RACI-ES		Spain	397	399	796	100% CCP+
	RACI-i2b2		North America	882	1,863	2,745	100% CCP+
	ReAct		France	275	804	1,079	70% CCP+ or RF+
	Dutch (AMC, BeSt, LUMC, DREAM)		Netherland	1,172	1,684	2,856	80% CCP+ or RF+
	ACR-REF (BRAGGSS, BRAGGSS2, ERA, KI, TEAR)		North America & Europe	347	264	611	85% CCP+ or RF+
	CORRONA		North America	894	1,838	2,732	61% CCP+ or RF+, 32% unknown
	Vanderbilt		North America	739	2,247	2,986	31% CCP+ or RF+, 56% unknown
	GARNET (BioBank Japan Project; BBJ)	Asian	Japan	2,414	14,245	16,659	79% CCP+, 76% RF+
	GARNET (Kyoto University)		Japan	1,237	2,087	3,324	85% CCP+, 86% RF+
	GARNET (IORRA)		Japan	423	559	982	87% CCP+, 88% RF+
	Korea		Korea	799	751	1,550	100% CCP+
	European	-	-	14,361	43,923	58,284	-
	Asian	-	-	4,873	17,642	22,515	-
	Trans-ethnic	-	-	19,234	61,565	80,799	-
In-silico replication study (Stage 2)	Genentech	European	North America	2,780	4,700	7,480	44% CCP+, 52% unknown 81% RF+, 1.7% unknown
	China	Asian	China	928	835	1,763	48% CCP+
	Total	-	-	3,708	5,535	9,243	-
De-novo replication study (Stage 3)	CANADAII	European	Canada	995	1,101	2,096	100% CCP+
	GARNET	Asian	Japan	5,943	5,557	11,500	81% CCP+, 86% RF+
	Total	-	-	6,938	6,658	13,596	-
Total	European	-	-	18,136	49,724	67,860	-
	Asian	-	-	11,744	24,034	35,778	-
	Trans-ethnic	-	-	29,880	73,758	103,638	-

b

Study stage	Cohort	Genotyping platform	GWAS QC criteria				Imputation method			No.SNPs after QC		Inflation factor		Covariates	X chrom. data
			Sample call rate	SNP call rate	MAF	HWE P-value	Reference panel	MAF	Quality score	Genotyped	Imputed	λ_{GC}	$\lambda_{GC,1000}$		
GWAS meta-analysis (Stage 1)	BRASS	Affymetrix Genome-wide Human SNP Array 6.0	>0.95	>0.95	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	649,178	8,201,244	1.015	1.008	Top 5 PCs	Available
	CANADA	Illumina HumanCNV370-Duo BeadChip	>0.95	>0.95	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	295,430	7,933,623	1.002	1.001	Top 5 PCs	Available
	EIRA	HumanHap300 BeadChip	>0.95	>0.95	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	298,193	8,163,538	0.991	0.994	Top 5 PCs	N.A.
	NARAC1	Illumina HumanHap550 BeadChip	>0.95	>0.95	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	479,671	8,254,787	1.017	1.012	Top 5 PCs	N.A.
	NARAC2	HumanHap300 BeadChip	>0.95	>0.95	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	261,974	7,733,592	1.023	1.003	Top 5 PCs	N.A.
	WTCCC	Affymetrix Genome-wide Human SNP Array 5.0	>0.99	>0.99	>0.01	$>10^{-5}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	339,790	7,385,370	1.043	1.004	Top 5 PCs	N.A.
	RACI-UK	Illumina Immunochip custom array	>0.99	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.7	126,740	873,840	1.058	1.008	Top 10 PCs	Available
	RACI-US	Illumina Immunochip custom array	>0.99	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.7	120,589	843,395	1.031	1.012	Top 10 PCs	Available
	RACI-SE-E	Illumina Immunochip custom array	>0.99	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.7	124,801	870,585	1.003	1.002	Top 10 PCs	Available
	RACI-SE-U	Illumina Immunochip custom array	>0.99	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.7	123,998	870,797	0.986	0.988	Top 10 PCs	Available
	RACI-NL	Illumina Immunochip custom array	>0.99	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.7	124,480	862,815	1.109	1.051	Top 10 PCs	Available
	RACI-ES	Illumina Immunochip custom array	>0.99	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.7	124,459	859,540	1.081	1.152	Top 10 PCs	Available
	RACI-i2b2	Illumina Immunochip custom array	>0.99	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.7	118,731	829,507	1.003	1.001	Top 10 PCs	Available
	ReAct	Illumina OmniExpress BeadChip Illumina Human 660W-Quad BeadChip	>0.98	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	257,299	7,588,538	0.992	0.991	Top 5 PCs	Available
	Dutch	Illumina Human 660W-Quad BeadChip Illumina HumanHap550 BeadChip Illumina HumanCNV370-Duo BeadChip	>0.95	>0.95	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	284,884	7,956,686	1.023	1.011	Top 5 PCs	Available
	ACR-REF	Illumina OmniExpress BeadChip Illumina Human 660W-Quad BeadChip	>0.95	>0.95	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	234,075	7,593,678	1.026	1.070	Top 5 PCs	Available
	CORRONA	Illumina OmniExpress BeadChip	>0.98	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	552,896	8,400,238	1.001	1.000	Top 5 PCs	Available
	Vanderbilt	Illumina OmniExpress BeadChip	>0.98	>0.99	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	541,143	8,372,666	0.987	0.995	Top 5 PCs	Available
	BBJ	Illumina HumanHap610-Quad BeadChip	>0.98	>0.99	>0.01	$>10^{-7}$	1000 Genomes Phase I (α) Asians	>0.005	>0.5	477,784	6,874,738	1.038	1.002	-	Available
	Kyoto	Illumina HumaHap610-Quad BeadChip Illumina HumanHap550 BeadChip Illumina HumanCNV370-Duo BeadChip	>0.90	>0.95	>0.05	$>10^{-7}$	1000 Genomes Phase I (α) Asians	>0.005	>0.5	227,348	6,254,431	1.099	1.038	-	N.A.
	IORRA	Affymetrix Genome-wide Human SNP Array 6.0	>0.95	>0.98	>0.05	$>10^{-6}$	1000 Genomes Phase I (α) Asians	>0.005	>0.5	465,832	6,567,923	0.992	0.989	-	Available
	Korea	Illumina Human 660W-Quad BeadChip Illumina HumanHap550 BeadChip	>0.90	>0.90	>0.01	$>10^{-6}$	1000 Genomes Phase I (α) Asians	>0.005	>0.5	418,837	6,424,378	1.007	1.007	-	Available
	European	-	-	-	-	-	-	-	-	-	8,747,962	1.073	1.003	-	-
	Asian	-	-	-	-	-	-	-	-	-	6,619,871	1.041	1.005	-	-
	Trans-ethnic	-	-	-	-	-	-	-	-	-	9,739,303	1.072	1.002	-	-
In-silico replication study (Stage 2)	Genentech	Illumina HumanOmni1-Quad_v1-0_B Illumina Humanhap550K	>0.95	>0.95	>0.10	$>10^{-4}$	1000 Genomes Phase I (α) Europeans	>0.005	>0.5	-	-	-	-	Top 5 PCs	N.A.
	China	Affymetrix Genome-wide Human SNP Array 6.0	>0.95	>0.95	>0.05	$>10^{-3}$	1000 Genomes Phase I (α) Asians	>0.005	>0.5	-	-	-	-	Top 5 PCs	N.A.
De-novo replication study (Stage 3)	CANADAII	iPlex genotyping system	-	-	-	-	-	-	-	-	-	-	-	-	Available
	GARNET	Taqman genotyping system	-	-	-	-	-	-	-	-	-	-	-	-	Available

a, Characteristics of the cohorts and subjects enrolled in the study. **b**, Genotype and imputation methods of the studies. CCP, anti-citrullinated peptide antibody; chrom, chromosome; N.A., not available; PC, principal component; QC, quality control; RF, rheumatoid factor.

Extended Data Table 2 | *cis*-eQTL of RA risk SNPs

a

RA risk SNP	Chr.	Position (bp)	eQTL gene	Cis-eQTL effect of best proxy SNP				Cis-eQTL effect of top eQTL SNP			
				Proxy SNP	Position (bp)	eQTL *P*	r^2	eQTL SNP	Position (bp)	eQTL *P*	r^2
chr1:2523811	1	2,523,811	*PLCH2*	rs10910099	2,533,552	2.2E-18	0.87	rs2494435	2,359,280	2.6E-45	<0.25
			TNFRSF14	rs2843401	2,528,133	1.1E-28	0.87	rs734999	2,513,216	2.1E-90	0.43
rs227163	1	7,961,206	*PARK7*	rs227163	7,961,206	4.6E-10	1.00	rs3766606	8,022,197	1.0E-53	<0.25
rs28411352	1	38,278,579	*MANEAL, YRDC*	rs2306627	38,260,503	3.9E-09	0.84	rs2306426	36,451,618	7.7E-10	<0.25
			INPP5B	rs2306627	38,260,503	7.5E-23	0.84	rs4072980	38,456,106	1.2E-113	<0.25
			SF3A3	rs2306627	38,260,503	3.3E-17	0.84	rs4072980	38,456,106	1.1E-190	<0.25
			FHL3	rs2306627	38,260,503	1.1E-11	0.84	rs4634868	38,465,315	9.8E-198	<0.25
rs2476601	1	114,377,568	*PTPN22*	rs2476601	114,377,568	3.4E-10	1.00	rs7555834	114,367,116	5.3E-43	<0.25
rs2228145	1	154,426,970	*AQP10*	rs6684439	154,395,839	3.3E-06	0.89	rs6668968	154,293,675	3.8E-40	<0.25
			IL6R	rs4129267	154,426,264	3.2E-27	1.00	rs4537545	154,418,879	2.0E-29	0.96
			UBE2Q1	rs4129267	154,426,264	9.7E-08	1.00	rs6660775	154,538,554	3.9E-21	<0.25
rs2317230	1	157,674,997	*FCRL5*	rs3761959	157,669,278	1.7E-09	0.87	rs6427386	157,530,097	9.8E-198	<0.25
			FCRL3	rs7528684	157,670,816	9.8E-198	0.87	rs2210913	157,668,993	9.8E-198	0.87
rs4656942	1	160,831,048	*LY9*	rs4656942	160,831,048	2.7E-96	1.00	rs576334	160,797,514	5.8E-195	<0.25
rs72717009	1	161,405,053	*SDHC*	rs12731669	161,410,458	5.5E-05	0.97	rs16832871	161,335,758	1.4E-142	<0.25
			FCGR2B	rs12731669	161,410,458	4.2E-83	0.97	rs6674499	161,618,151	9.8E-198	<0.25
rs17668708	1	198,640,488	*PTPRC*	rs17669032	198,653,174	5.2E-05	0.97	rs2296618	198,666,232	2.1E-05	0.78
rs1980422	2	204,610,396	*CD28*	rs1980421	204,610,004	7.3E-18	1.00	rs2140148	204,572,140	8.1E-21	0.47
rs10028001	4	79,502,972	*ANXA3*	rs10028001	79,502,972	1.1E-04	1.00	rs4975144	79,474,040	1.4E-09	<0.25
rs2561477	5	102,608,924	*PAM*	rs411648	102,602,902	2.2E-113	1.00	rs2431321	102,118,794	9.8E-198	<0.25
			GIN1	rs2288786	102,600,754	1.3E-06	1.00	rs42431	102,400,063	2.6E-13	0.43
rs657075	5	131,430,118	*ACSL6*	rs657075	131,430,118	3.8E-12	1.00	rs253946	131,330,461	9.2E-26	0.32
chr6:14103212	6	14,103,212	*CD83*	rs12530098	14,107,197	2.6E-24	1.00	rs16874672	14,087,484	2.2E-26	0.90
rs2234067	6	36,355,654	*KCTD20*	rs4713969	36,349,008	8.2E-05	0.99	rs4711453	36,439,391	3.1E-32	<0.25
			STK38	rs4713969	36,349,008	1.4E-06	0.99	rs1812018	36,557,976	6.8E-15	<0.25
			-	rs4713969	36,349,008	2.1E-26	0.99	rs10947614	36,573,822	1.1E-146	<0.25
			SFRS3	rs4713969	36,349,008	2.6E-11	0.99	rs7743396	36,579,252	1.5E-52	<0.25
rs9373594	6	149,834,574	*C6orf72*	rs9377224	149,853,707	4.0E-06	1.00	rs9322189	149,909,933	1.8E-15	0.30
			NUP43	rs9377224	149,853,707	4.1E-64	1.00	rs9688350	150,052,113	9.8E-198	0.27
rs2451258	6	159,506,600	*RSPH3*	rs2485363	159,506,121	5.0E-05	1.00	rs12216499	159,368,524	2.0E-119	<0.25
rs1571878	6	167,540,842	*RNASET2*	rs1571878	167,540,842	9.8E-198	1.00	rs429083	167,383,972	9.8E-198	0.39
chr7:128580042	7	128,580,042	*TNPO3*	rs3807306	128,580,680	1.4E-150	0.81	rs3807306	128,580,680	1.4E-150	0.81
			-	rs3807306	128,580,680	2.4E-32	0.81	rs10229001	128,599,397	4.5E-49	0.48
			IRF5	rs3807306	128,580,680	9.8E-198	0.81	rs7807018	128,640,188	9.8E-198	0.48
rs2736337	8	11,341,880	*C8orf13, C8orf12*	rs2736340	11,343,973	1.6E-174	0.99	rs4840568	11,351,019	3.8E-175	0.92
			BLK	rs1478901	11,347,833	1.8E-120	0.99	rs998683	11,353,000	1.5E-120	0.95
rs10985070	9	123,636,121	*TRAF1*	rs10985070	123,636,121	3.9E-72	1.00	rs2416804	123,676,396	3.8E-73	0.96
			PHF19	rs10985070	123,636,121	2.9E-10	1.00	rs10760129	123,700,183	2.2E-10	0.95
			C5	rs10985070	123,636,121	4.9E-68	1.00	rs2416811	123,789,634	2.0E-146	0.31
rs947474	10	6,390,450	-	rs947474	6,390,450	6.5E-06	1.00	rs12416248	6,391,031	1.1E-43	<0.25
rs2671692	10	50,097,819	*WDFY4*	rs2671692	50,097,819	3.0E-09	1.00	rs7072606	49,933,974	1.1E-50	<0.25
rs968567	11	61,595,564	*C11orf10*	rs968567	61,595,564	3.1E-39	1.00	rs174538	61,560,081	2.5E-67	0.40
			FADS1	rs968567	61,595,564	8.1E-62	1.00	rs968567	61,595,564	8.1E-62	1.00
			FADS2	rs968567	61,595,564	4.8E-34	1.00	rs968567	61,595,564	4.8E-34	1.00
rs10774624	12	111,833,788	*SH2B3*	rs653178	112,007,756	1.7E-19	0.86	rs2239195	111,881,309	1.0E-33	<0.25
			ALDH2	rs653178	112,007,756	8.7E-07	0.86	rs16941669	112,245,637	4.4E-50	<0.25
rs4780401	16	11,839,326	*TXNDC11*	rs11075010	11,826,013	8.3E-09	0.93	rs12919035	11,821,508	4.4E-12	0.49
rs72634030	17	5,272,580	*ZNF594*	rs8080217	5,164,761	8.7E-11	0.88	rs2071456	5,031,946	1.5E-12	0.65
			C17orf87	rs8080217	5,164,761	3.3E-05	0.88	rs2641232	5,087,602	1.4E-53	<0.25
			-	rs8080217	5,164,761	3.6E-70	0.88	rs7426	5,288,983	9.8E-198	<0.25
			NUP88	rs8080217	5,164,761	3.3E-27	0.88	rs1989946	5,313,152	8.9E-96	<0.25
			MIS12	rs8080217	5,164,761	8.5E-10	0.88	rs1805448	5,384,327	2.2E-35	<0.25
rs1877030	17	37,740,161	*FBXL20*	rs12937013	37,665,571	3.4E-15	1.00	rs8076462	37,400,025	3.1E-42	<0.25
			PPP1R1B	rs1877030	37,740,161	1.8E-10	1.00	rs879606	37,781,849	8.0E-18	0.41
			-	rs11657058	37,699,378	3.9E-05	1.00	rs7219814	37,478,801	2.1E-111	<0.25
			IKZF3	rs4795385	37,733,148	8.8E-24	1.00	rs2517955	37,843,681	5.2E-82	0.33
chr17:38031857	17	38,031,857	-	rs907092	37,922,259	6.6E-11	0.90	rs7219814	37,478,801	2.1E-111	<0.25
			IKZF3	rs11557467	38,028,634	3.3E-05	0.84	rs9896940	37,895,975	3.1E-25	<0.25
			GSDMB	rs10852936	38,031,714	9.8E-198	0.98	rs9901146	38,043,343	9.8E-198	0.84
			ORMDL3	rs10852936	38,031,714	9.8E-198	0.98	rs8076131	38,080,912	9.8E-198	0.86
rs2469434	18	67,544,046	*CD226*	rs1610555	67,543,147	2.3E-33	0.99	rs763361	67,531,642	2.4E-50	0.66
rs4239702	20	44,749,251	*CD40*	rs4239702	44,749,251	1.3E-34	1.00	rs745307	44,747,086	1.5E-72	<0.25
rs73194058	21	34,764,288	*IL10RB*	rs11702844	34,759,876	1.3E-11	0.97	rs1058867	34,669,381	3.0E-69	<0.25
			IFNAR1	rs11702844	34,759,876	8.0E-12	0.97	rs2257167	34,715,699	4.2E-73	<0.25
			TMEM50B	rs11702844	34,759,876	3.1E-11	0.97	rs1059293	34,809,693	2.2E-103	<0.25
			-	rs11702844	34,759,876	2.8E-34	0.97	rs2834217	34,822,150	9.8E-198	<0.25
rs1893592	21	43,855,067	*UBASH3A*	rs1893592	43,855,067	6.4E-92	1.00	rs1893592	43,855,067	6.4E-92	1.00
rs2236668	21	45,650,009	*ICOSLG*	rs7278940	45,648,992	3.7E-06	1.00	rs3788111	45,668,171	8.4E-16	<0.25
rs11089637	22	21,979,096	-	rs11089637	21,979,096	9.8E-198	1.00	rs5754217	21,939,675	9.8E-198	0.87
rs909685	22	39,747,671	*SYNGR1*	rs909685	39,747,671	1.0E-140	1.00	rs909685	39,747,671	1.0E-140	1.00
			MAP3K7IP1	rs909685	39,747,671	1.3E-05	1.00	rs5750824	39,830,123	5.9E-07	0.28

b

SNP	Chr.	Position (bp)	eQTL gene	Nominal *P* for cis-eQTL	
				$CD4^+$ T-cell	$CD14^+16^-$ Monocyte
rs28411352	1	38,278,579	*INPP5B*	0.022	**3.6E-16**
			FHL3	0.081	**8.9E-13**
rs2317230	1	157,674,997	*FCRL3*	**3.6E-06**	0.87
rs9653442	2	100,825,367	*AFF3*	**5.2E-08**	0.18
rs7731626	5	55,444,683	*IL6ST*	**2.3E-07**	0.0087
			ANKRD55	**4.1E-14**	0.43
rs2234067	6	36,355,654	*ETV7*	2.9E-04	**1.1E-10**
rs9373594	6	149,834,574	*NUP43*	5.4E-04	**1.5E-05**
rs1571878	6	167,540,842	*RNASET2*	**6.9E-20**	**1.3E-05**
rs67250450	7	28,174,986	*JAZF1*	**3.6E-17**	2.0E-04
chr7:128580042*	7	128,580,042	*TNPO3*	1.0E-04	**3.0E-07**
rs10985070	9	123,636,121	*MEGF9*	**3.3E-06**	0.10
			PSMD5	0.017	**1.8E-05**
			PHF19	0.0016	**5.6E-06**
rs968567	11	61,595,564	*FADS2*	**1.4E-31**	**8.9E-35**
			FADS1	**2.1E-32**	0.094
rs11605042	11	72,411,684	*STARD10*	0.82	**1.0E-07**
rs4409785	11	95,311,422	*SESN3*	**1.5E-11**	0.43
rs773125	12	56,394,954	*SUOX*	0.27	**1.1E-09**
rs1633360	12	58,108,052	*TSPAN31*	0.13	**1.0E-05**
			METTL21B	**1.4E-09**	**4.0E-10**
rs9603616	13	40,368,069	*COG6*	0.0011	**1.2E-06**
rs4780401	16	11,839,326	*TXNDC11*	**1.3E-05**	0.62
rs72634030	17	5,272,580	*MIS12*	0.0039	**1.3E-05**
rs1877030	17	37,740,161	*STARD3*	0.048	**4.5E-05**
chr17:38031857†	17	38,031,857	*GSDMA*	**2.1E-06**	0.63
			GSDMB	**4.3E-11**	0.19
			ORMDL3	**6.8E-09**	0.0098
rs4239702	20	44,749,251	*CD40*	0.31	**1.7E-08**
rs73194058	21	34,764,288	*IFNGR2*	0.096	**1.9E-06**
			TMEM50B	**7.5E-07**	0.013
rs1893592	21	43,855,067	*UBASH3A*	**3.8E-14**	0.92

a, *cis*-eQTL of PBMCs in the RA risk SNPs. Significant *cis*-eQTLs of RA risk SNPs is indicated (FDR $q < 0.05$). SNPs and positions are based on the positive strand of NCBI build 37. Linkage disequilibrium of the proxy SNPs evaluated in the eQTL study and the best *cis*-eQTL SNP in the region with the RA risk SNPs is indicated as r^2 values. When the expression probe was not assigned to any genes, the eQTL gene is labelled with a dash. **b,** *cis*-eQTL of T cells and monocytes in the RA risk SNPs. Significant *cis*-eQTLs of RA risk SNPs are indicated in bold (gene-based permutation $P < 0.05$).

* *cis*-eQTL of the proxy SNP (rs3807307, $r^2 = 0.96$) was evaluated.

† *cis*-eQTL of the proxy SNP (rs11557466, $r^2 = 0.98$) was evaluated.

Biochemical and Biophysical Research Communications 452 (2014) 254–262

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

ELSEVIER

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

Review

Genetic basis of rheumatoid arthritis: A current review

CrossMark

Yuta Kochi [a,*], Akari Suzuki [a], Kazuhiko Yamamoto [a,b]

[a] Laboratory for Autoimmune Diseases, Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan
[b] Department of Allergy and Rheumatology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

ARTICLE INFO

Article history:
Received 13 June 2014
Available online 29 July 2014

Keywords:
Rheumatoid arthritis
Genome-wide association study
HLA-DRB1
Anti-citrullinated protein antibody

ABSTRACT

Rheumatoid arthritis (RA) is one of the most common autoimmune diseases. As with other complex traits, genome-wide association studies (GWASs) have tremendously enhanced our understanding of the complex etiology of RA. In this review, we describe the genetic architecture of RA as determined through GWASs and meta-analyses. In addition, we discuss the pathologic mechanism of the disease by examining the combined findings of genetic and functional studies of individual RA-associated genes, including *HLA-DRB1*, *PADI4*, *PTPN22*, *TNFAIP3*, *STAT4*, and *CCR6*. Moreover, we briefly examine the potential use of genetic data in clinical practice in RA treatment, which represents a challenge in medical genetics in the post-GWAS era.

Contents

1. Genetic aspects of rheumatoid arthritis

Rheumatoid arthritis (RA) is one of the most common forms of autoimmune arthritis, affecting approximately 0.5–1.0% of the world's population. The serum of most RA patients contains autoantibodies, such as rheumatoid factor (RF) or anti-citrullinated protein antibodies (ACPAs), the presence of which constitutes one of the new classification criteria for RA revised in 2010 [1]. Although RF is also present in other autoimmune diseases and immunological conditions, such as chronic infection and inflammation, ACPAs have a higher specificity, suggesting a central role for citrulline as an antigenic determinant in this disease [2] (Fig. 1). This suggests that autoimmunity to citrullinated proteins may be the hallmark of RA pathogenesis. However, the rest of RA patients lack these autoantibodies, suggesting a heterogeneity in the disease etiology. In clinical practice, the appearance of biologics that target inflammatory cytokines has dramatically improved the outcome of RA, although some patients still suffer destructive arthritis that leads to disability. The limitations of current RA therapies underscore the need for further investigation of disease pathogenesis and identification of new therapeutic targets.

* Corresponding author. Address: Laboratory for Autoimmune Diseases, Center for Integrative Medical Sciences, RIKEN, Yokohama 230-0045, Japan.
E-mail address: ykochi@src.riken.jp (Y. Kochi).

http://dx.doi.org/10.1016/j.bbrc.2014.07.085
0006-291X/

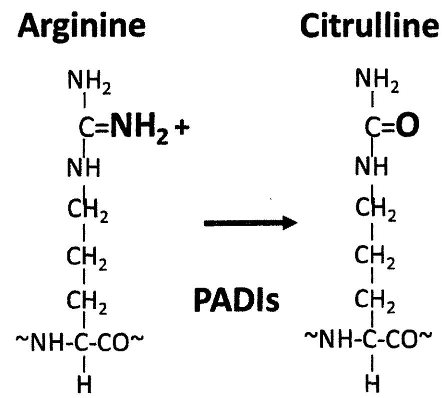

Fig. 1. Citrullination of arginine residues by PADI enzymes.

As with other complex traits, evidence from familial studies suggests that RA is caused by a combination of genetic and environmental factors. For instance, a recent population-based epidemiologic study in Sweden demonstrated that the familial odds ratio for RA is approximately 3 in first-degree relatives of RA patients and 2 in second-degree relatives [3]. In addition, higher concordance rates in monozygotic twins over dizygotic twins suggest the involvement of genetic factors [4–6]. The heritability of a disease, which is defined as the contribution of genetic variation to variation in the liability of that disease, has been estimated at around 60% in RA by the above-mentioned studies.

The establishment of a comprehensive catalog of common genetic variants in human populations by the HapMap project [7], as well as significant recent advances in genotyping technology, now enable searching of the entire genome at once for risk loci for complex diseases. This methodologic approach, now broadly known as the "genome-wide association study (GWAS)," has greatly advanced understanding of the genetic background of complex traits such as RA. In contrast, with a few exceptions, such as cigarette smoking, infections, and diet, little is known about the role of environmental factors in the development of RA. Although environmental aspects of RA are beyond the topic of this review, readers are referred to other excellent articles [8,9].

2. *HLA-DRB1* gene

Since the first evidence suggesting the involvement of human leukocyte antigens (HLAs) in RA was reported in 1969 [10], polymorphisms in the *HLA* region have been at the center of genetic studies of RA. That study demonstrated reduced lymphocyte responses in autologous mixed cultures of cells from RA patients, suggesting that polymorphisms in *HLA* genes (which encode the major histocompatibility complex [MHC] molecules that present antigens to T cells) are shared among patients [10]. Subsequently, serologic studies showed that the frequency of the HLA-DR4 serotype is higher in RA patients compared with control subjects [11,12]. Other serotypes, such as DR1, are also associated with increased risk for RA, although the increase in risk is moderate compared with that of DR4 [13]. Sequencing of *HLA-DRB1*, which encodes the polymorphic β-chain of the DR molecule, revealed that the prominent subtypes of DR4 differ between populations. For example, Europeans harbor the *04:01 and *04:04 *DRB1* alleles and East Asians harbor the *04:05 allele. In addition, several subtypes of the *DR4* allele, such as *04:02 and *04:03, were shown to protect against the disease. These observations led to the hypothesis that a conserved epitope (i.e., QKRAA/QRRAA/RRRAA) spanning amino acid residues 70–74 in the third hypervariable region of the β chain (which is now referred to as a "shared epitope [SE]") is associated with RA susceptibility [14]. Although this SE hypothesis is generally accepted, there have been several attempts to reclassify or refine it. Recently, two studies demonstrated that the amino acids at residues 11 and 13 are also independently associated with the disease, which may explain the higher risk associated with *DR4* (*04:01/*04:04/*04:05) compared with *DR1* (*01:01) [15,16].

As the importance of ACPAs in RA has become apparent over the last decade, the strong association between SE alleles and the appearance of ACPAs in RA patients has been demonstrated in multiple populations [17–20], suggesting that DR molecules encoded by SE alleles are involved in the presentation of citrullinated peptides to T cells (Fig. 2). This hypothesis is supported by the observation in human *DR4*-transgenic mice that the conversion of arginine (positively charged) to citrulline (neutral) leads to a substantial increase in HLA-peptide affinity and subsequent activation of CD4 T cells [21]. The molecular basis of these observations was determined in a recent crystal structure analysis showing that citrulline residues of peptides are accommodated within the electropositive P4 pocket of *DRB1**04:01/04, whereas the electronegative P4 pocket of the non-risk allele product *04:02 are not accommodated [22]. As the amino acid residues at positions 13 and 71 comprise the P4 pocket and directly contact the citrulline residue, the nature of these residues may be crucial in the presentation of citrullinated peptides and may explain the genetic association between *HLA-DRB1* and RA.

The primary citrullinated autoantigens that directly cause RA are poorly defined because clinical laboratory testing of serum samples from RA patients typically involves an artificial cyclic-citrullinated peptide that reacts with multiple citrullinated self proteins. However, fibrinogen, α-enolase, vimentin, immunoglobulin binding protein (BiP), and type II collagen, all of which are expressed in the synovial joint tissues, are potential candidates [23]. The primary autoantigen may differ between individuals, as a study examining patient serum samples showed that epitope spreading with an increase in the recognition of citrullinated antigens occurs before the onset of RA [24]. Differences in antibody profile between patients could depend on other genetic and environmental factors. Cigarette smoking is an environmental factor that substantially increases the risk of ACPA appearance. Intriguingly, gene-environment interactions (defined as a departure from a multiplicative model) between the *HLA-DRB1* SE allele and smoking have been reported [25–27]. Another study demonstrated that the combination of smoking and genetic factors, including *HLA-DRB1*, may determine the specificity of ACPAs in RA patients [28,29].

The association between *HLA-DRB1* and ACPA-negative RA has been relatively understudied due to the higher prevalence of APCA-positive RA. In Europeans, *HLA-DR3* (*DRB1**03:01) is associated with ACPA-negative RA [30,31]. A study of Japanese populations (in which the *DRB1**03:01 allele is rare) indicated that both ACPA- and RF-negative RA are associated with *DR14* and the *HLA-DR8* homozygote [32]. These observations suggest that the contribution of *HLA-DRB1* alleles is distinctly different in ACPA-negative RA. However, the lack of a specific serologic test for ACPA-negative RA could result in heterogeneity in studies of different cohorts. To overcome this problem, a recent study of ACPA-negative patients that statistically adjusted for the clinical heterogeneity of ACPA-negative RA identified two independent association signals in the *HLA-DRB1* and *HLA-B* gene products: serine 11 (encoded by *DRB1**03) and aspartate 9 (encoded by *HLA-B**08), respectively, providing additional evidence that ACPA-positive and -negative RA are genetically distinct [33].

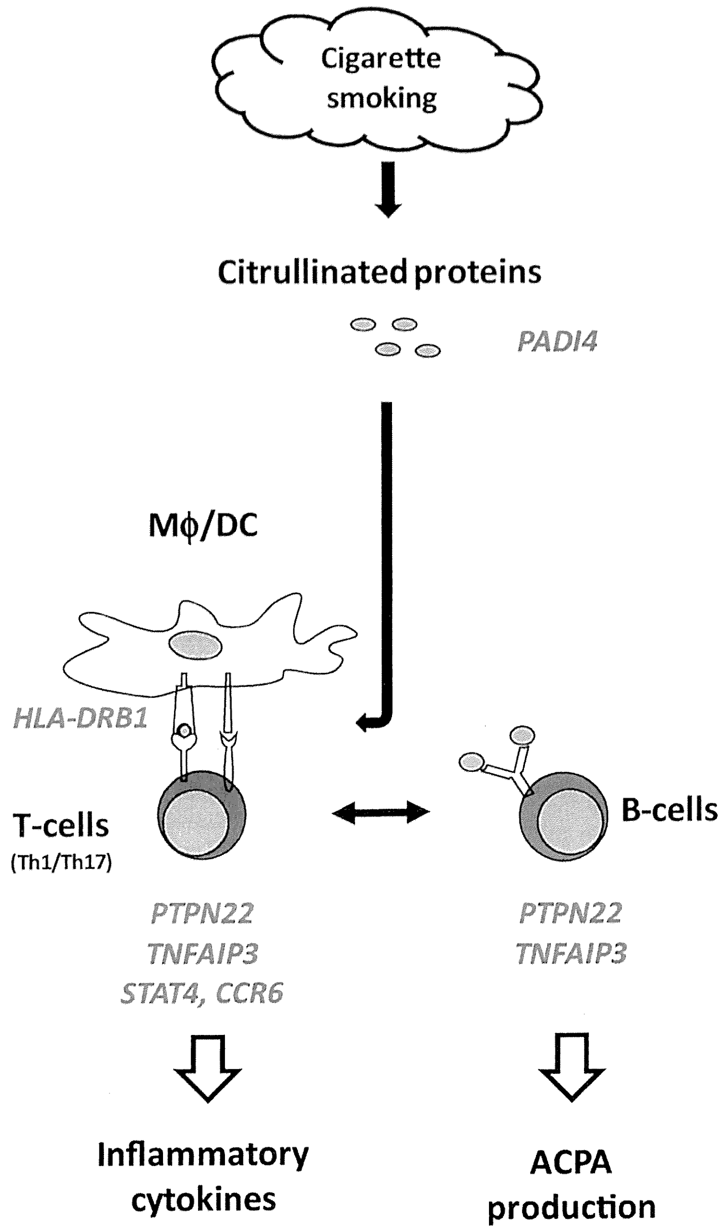

Fig. 2. Genetic factors involving autoimmunity to citrullinated proteins.

3. Insights from GWASs

In a GWAS, ~1 million single-nucleotide polymorphisms (SNPs) are simultaneously genotyped for affected patients (cases) and non-affected individuals (controls). The null hypothesis of a GWAS is that there is no association between a given SNP and disease susceptibility and is tested by comparing the allele frequency or genotype frequency between cases and controls. If the null hypothesis is rejected with a genome-wide significance level, which is usually set at $\alpha = 5 \times 10^{-8}$, the genetic marker indicates the presence of a causal variant(s) in the locus. Following the first GWAS, which was performed in a Japanese population and covered only the gene regions and not the intergenic regions [34], multiple GWASs have been performed in worldwide populations. These individual GWASs identified a number of RA-susceptibility loci, including *PADI4* [34], *PTPN22* [35], *TNFAIP3* [36], *TRAF1/C5* [37], *REL* [38], and *CCR6* [39]. However, each of these studies lacked sufficient statistical power to detect loci that have a moderate effect size. To overcome this limitation, meta-analyses of GWASs of both European and Asian populations were conducted, which increased the number of risk loci [40,41]. More recently, a multi-ethnic meta-analysis of GWASs was performed that involved collaboration between 25 study groups worldwide and a total of over 100,000 subjects (29,880 RA cases and 73,758 controls) of European or Asian ancestry [42]. This was the largest meta-analysis of autoimmune disease GWASs ever performed and identified 101 RA risk loci.

Although a GWAS can identify disease risk loci, it can directly identify neither the responsible genes nor disease-causing variants in the loci. The disease-causing variants can affect the function of the responsible genes by (1) introducing stop codons or frame-shift mutations, (2) changing the amino acid sequence, (3) affecting alternative splicing, or (4) regulating the level of transcript expression. Among the 100 risk-associated SNPs in non-HLA RA risk loci, only 16% are in linkage disequilibrium with missense SNPs, indicating that the majority of causal variants in the risk loci affect splicing or the level of gene expression. In fact, RA-risk SNPs were found in 44 *cis*-acting expression quantitative trait loci (*cis*-eQTL) identified in peripheral blood mononuclear cells [43], indicating that disease-causing variants in the risk loci affect the expression level of genes in *cis*. Similar observations have been reported for other autoimmune diseases, indicating that the accumulation of quantitative differences in risk genes leads to disease onset [44,45].

As the regulation of gene expression in cells of the immune system, including T cells, B cells, and macrophages, is quite sophisticated, the *cis*-eQTL effects may also be cell specific. Data from recent human genome studies, such as the Encyclopedia of DNA Elements (*ENCODE*) project, provide clues that may help elucidate the underlying mechanism of cell-specific eQTL effects for many loci identified in GWASs. Using omics data (e.g., genomic, epigenomic, transcriptomic) obtained via next-generation sequencing technologies, the ENCODE project developed a comprehensive "parts list" of functional elements in the human genome that included descriptions of regulatory elements that control cells and the circumstances under which a given gene is active [46]. Analyses of non-HLA RA risk loci for enrichment in epigenetic chromatin marks revealed significant trimethylation of histone H3 at lysine 4 (H3K4me3), which is a promoter- and enhancer-specific modification associated with active transcription [47]. Among 34 cell types investigated, H3K4me3 peaks were particularly enriched in primary $CD4^+$ regulatory T cells (Treg cells) [42]. This observation suggests that a substantial proportion of RA risk variants are involved in transcriptional regulation of genes in Treg cells and that modulating the activity of Treg cells by targeting these GWAS-identified genes could be used to treat RA.

4. RA risk genes and pathogenesis

As mentioned above, GWASs have identified more than 100 RA risk loci. Although the effect of each individual locus is moderate (e.g., the odds ratio for most individual alleles ranges between 1.1 and 1.3), detailed analyses of individual loci to identify disease-causing variants and to determine the effect of the identified variants on responsible genes (e.g., gain-of-function or loss-of-function) would enhance our understanding of the disease. Examples of RA risk genes and their role in the pathogenesis of RA are discussed below.

4.1. PADI4

PADI4 was the first RA susceptibility gene identified in a GWAS of an Asian population [34]. *PADI4* is a member of the peptidyl arginine deaminase gene family and encodes an enzyme that converts arginine into citrulline in a posttranslational modification (Fig. 1). Although the physiological role of citrullination of proteins is not well understood, the specific presence of autoantibodies to citrullinated proteins (i.e., ACPAs) in RA supports the hypothesis that citrullination of autoantigens leads to autoimmunity in RA. Through in vitro assays, we have shown that transcripts of the risk haplotype of *PADI4* are more stable than transcripts of the non-risk haplotype, suggesting that increased expression and function of PADI4 could increase the risk of developing RA. Interestingly, the effect size of *PADI4* variants on the risk of developing disease differs between European and Asian populations, with greater effects observed in Asian populations [48]. One possible explanation for this genetic heterogeneity is the impact of environmental factors. For example, we demonstrated that *PADI4* variants exert an epistatic effect in conjunction with cigarette smoking, especially in males [49]. Because 40–60% of East Asian males smoke, compared with 10–30% of European males, the higher effect size of *PADI4* in Asian populations may be partially explained by the difference in smoking rates.

4.2. PTPN22

Among the RA-associated common variants outside the *HLA* region that have been identified by GWASs, the missense variant of the protein tyrosine phosphatase nonreceptor 22 (*PTPN22*) gene has the strongest effect [50]. To date, this missense variant (*PTPN22* R620W) has been associated with over 20 different autoimmune diseases in European populations, including systemic lupus erythematosus (SLE), type 1 diabetes, and Graves disease and is considered a common autoimmune gene [51]. Interestingly, this variant is very rare or is not polymorphic in Asian and African populations [52,53] and provides another example of genetic heterogeneity among populations.

PTPN22 encodes lymphoid tyrosine phosphatase (LYP), which dephosphorylates the phosphotyrosine residues of target proteins in lymphocytes. The disease-associated variant exchanges the arginine residue at position 620 in the proline-rich 1 motif to tryptophan. In vitro assays demonstrated that expression of the R620W risk allele leads to interference with the physical association between LYP and c-Src kinase (CSK), resulting in increased LYP activity. Parallel to this observation, both T-cell receptor (TCR) and B-cell receptor signaling were found to be reduced in the lymphocytes of risk allele carriers [54,55]. These observations suggest that the R620W LYP variant is a gain-of-function mutant. However, when this mutation was introduced at residue 619 of the murine ortholog of human LYP, Pep (Pep_619W), the phenotype of the knock-in mice was similar to that of Pep-deficient mice, characterized by splenomegaly and spontaneous germinal-center reactions [56]. This observation suggests that in contrast to human LYP_620W, murine Pep_619W is a loss-of-function variant. This is also supported by evidence demonstrating that TCR signaling is enhanced in both Pep-deficient and Pep_619W knock-in mice. Interestingly, an autoimmune response was observed in B6×129 Pep_619W knock-in mice, suggesting that Pep_619W triggers autoimmunity in mice in combination with other genetic factors [57]. The conflicting observations in human and mouse studies demonstrate the difficulty of translating findings from mouse models to human diseases (for a full review, see [58]).

4.3. TNFAIP3

Several GWASs almost simultaneously reported an association between the tumor necrosis factor-α-induced protein 3 (*TNFAIP3*) locus and RA [36] and SLE [59,60]. *TNFAIP3* encodes the ubiquitin-editing enzyme A20, which is a key player in the negative feedback regulation of NF-κB signaling. Two functional variants that may cause the diseases have been identified to date, one of which is a missense variant involving substitution of phenylalanine with cysteine at amino acid position 127, which in vitro assays have shown results in impaired A20 function [60]. The other variant is a TT>A polymorphic dinucleotide located 42 kb downstream of the *TNFAIP3* promoter that reduces the avidity of NF-κB subunit binding, resulting in reduced *TNFAIP3* expression [61]. Both of these variants impair the function of A20 and consequently augment NF-κB signaling. As these two variants are in linkage disequilibrium, both may simultaneously contribute to disease development. In mice, specific ablation of *Tnfaip3* in myeloid cells results in spontaneous development of a severe polyarthritis resembling RA [62]. These mice have high serum levels of inflammatory cytokines, including TNF-α, which is consistent with sus-

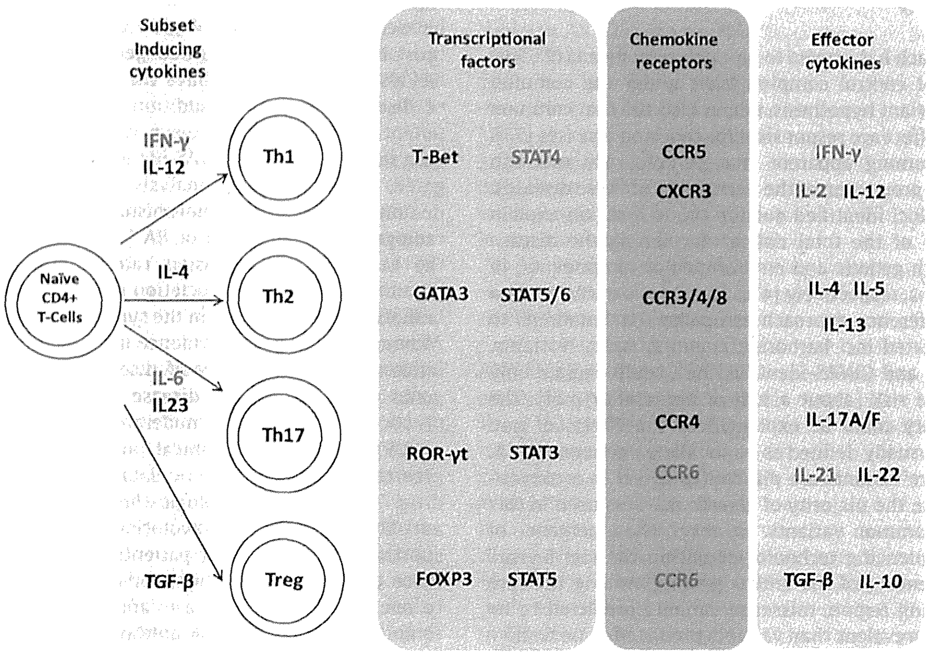

Fig. 3. $CD4^+$ helper T-cell subsets and their regulating factors.

tained NF-κB activation in macrophages. Interestingly, A20 is frequently inactivated by somatic mutations and/or deletions in B-lineage lymphomas [63]. Similar observations have been reported with somatic mutations in genes associated with lymphomas, such as *REL*, *FCRL3*, and *DDX6*, in which common variants cause RA [42], revealing contrasts in the etiologies of these diseases.

4.4. STAT4

An association between the signal transducer and activator of transcription 4 (*STAT4*) gene locus and the autoimmune diseases RA and SLE was first reported in a European population [64], and this association has been repeatedly confirmed in GWASs in multiple populations. An association between this locus and other autoimmune diseases, including systemic sclerosis and Sjögren syndrome, has also been reported [65]. Allelic expression analyses revealed that *STAT4* is overexpressed in carriers of the risk allele, suggesting that an increase in the function of the *STAT4* gene product results in these autoimmune diseases [66]. Members of the STAT family of proteins act as transcription factors; they are activated upstream by JAK proteins and mediate signaling associated with many inflammatory cytokines. STAT4 plays a critical role in type I interferon signaling. In fact, expression of the risk variant of *STAT4* is associated with greater IFN-α-induced gene expression in peripheral blood cells of SLE patients [67]. STAT family members also play essential and non-redundant roles in the differentiation of Th1, Th2, and Th17 helper proinflammatory T cells. Both Th1 and Th17 cells are thought to be involved in autoimmune disorders, as *STAT4* is involved in the differentiation of Th1 cells (Fig. 3). Interestingly, GWASs have revealed that *STAT3*, which plays a role in the differentiation of Th17 cells, is associated with the development of Crohn's disease [68], psoriasis [69], and multiple sclerosis [70]. Because Th17 cells are thought to play a significant role in the pathology of these autoimmune diseases, the *STAT3* variant may drive the activity of Th17 cells in these diseases. These observations suggest that different helper T-cell subsets may play the central role in different diseases in which the genetic factors that drive the activity of each subset also differ.

4.5. CCR6

In a GWAS in a Japanese population, we identified an association between the C—C chemokine receptor type 6 (*CCR6*) locus and RA [39]. We then examined the *CCR6* region for causal variants and identified a dinucleotide polymorphism (*CCR6DNP*) in the 5′-flanking region that influences the binding of nuclear proteins and enhances the transcriptional activity of *CCR6*. The risk allele exhibits greater enhancing activity and the level of *CCR6* transcription is higher in cells with the risk genotype. *CCR6DNP* is also associated with the positive status of IL-17A in the serum of RA patients, suggesting that *CCR6DNP* influences the activity of Th17 cells. In the SKG mouse model of arthritis, which involves a mutation in *Zap70*, Ccr6+Th17 cells are recruited into inflamed joints by the Ccr6 ligand Ccl20. Administration of anti-Ccr6 blocking antibodies substantially relieves the inflammation, suggesting that Ccr6+Th17 cells play a role in the pathogenesis of arthritis [71]. Although these data strongly support the hypothesis that Th17 cells are involved in RA, an association between the *CCR6* locus and disease development has only been confirmed for Crohn's and Basedow's diseases and not for other Th17-related diseases, such as psoriasis and multiple sclerosis, in which the *STAT3* variant increases the risk, as mentioned above.

Because the gene is also expressed in other T-cell subsets, *CCR6* can influence the activity of Treg and γδ T cells. Recently, another *CCR6*-expressing T-cell type, designated exFoxp3 Th17, was identified in a murine arthritis model [72]. In this model, CD25 lo Foxp3+ $CD4^+$ T cells cease to express Foxp3 and undergo transdifferentiation into Th17 cells. These exFoxp3 Th17 cells are more potent osteoclastogenic T cells than are naïve $CD4^+$ T cell-derived Th17 cells, although the roles of counterpart cells in humans is not clear. Taken together, these observations suggest that the differential effects of *CCR6* variants in different autoimmune diseases may be linked to the cells that drive the diseases.

5. Missing heritability

The GWAS approach has proven to be a powerful means of identifying risk loci that control complex traits under the common disease–common variant hypothesis, which assumes that common variants of modest effect are responsible for common diseases [73]. However, it is becoming apparent that common variants can explain only a small proportion of the heritability of these diseases. In RA, the 100 risk loci identified outside the *HLA* region explain only 5.5% and 4.7% of the total risk of developing the disease (which involves both genetic and environmental components) in Europeans and Asians, respectively [42]. An analysis of GWAS data using a Bayesian inference approach estimated that hundreds to thousands of associated loci harboring common causal variants, including *HLA-DRB1* and GWAS-identified loci, could explain only ~30% of the disease risk (about a half of heritability) [74]. The remaining heritability could be explained by the effects of rare variants, which are usually defined as minor allele frequencies <1%. The signals from rare variants are difficult to detect in a conventional GWAS because the majority of genetic markers used in this type of study are common variants. However, the emergence of next-generation sequencing technologies within the last 5 years now enable resequencing of the entire genome. Among the rare variants in the coding region, missense variants predicted to be damaging are more prevalent than variants predicted to be benign, whereas most common variants are predicted to be benign, consistent with studies demonstrating that rare variants in coding regions are under purifying selection [75]. This evidence suggests that the contribution of each individual variant to disease development should be higher for a rare missense variant than a common missense variant, warranting the sequencing of protein-coding regions based on priority. In a recent study attempting to determine the roles of rare variants in RA, deep exon sequencing of 25 biological candidate genes from GWAS-identified loci was performed and resulted in the identification of an accumulation of missense variants in the *IL2RA* and *IL2RB* genes [76]. A more comprehensive approach involves whole-exome sequencing, which decodes all protein-coding genes. However, to date no study has succeeded in identifying RA-associated rare variants using whole-exome sequencing, primarily due to insufficient statistical power. For example, a causal rare variant with a frequency of 0.2% and relative risk of 10 using a sample set of 200 cases and 200 controls has only 0.2% power to be detected at the conventional GWAS significance threshold ($\alpha = 5 \times 10^{-8}$) [75], indicating that lower-cost sequencing technologies that provide greater statistical power are needed to analyze rare variants.

6. Clinical use of genetic data in RA

In the final section of this review, we discuss the use of genetic data in clinical practice as it pertains to treating RA. The use of genetic data represents a challenge in the post-GWAS era because RA is a very heterogeneous disease with an outcome that is difficult to predict. The heterogeneity of RA can be partially explained by genetic factors; that is, the specific combination of genetic factors in an individual can determine the outcome of the disease. In this context, GWAS data can be used to predict an individual's disease phenotype. Phenotype prediction has been intensely investigated in RA with respect to two outcomes: disease severity and drug response.

The nature of progressive joint damage in RA varies considerably between individuals. Patients who would experience more rapid progression need more extensive therapy, such as the early use of biologics. Disease severity can be quantified by assessing and scoring the degree of joint damage using radiographic imaging. Regression analyses can then be performed to test associations between changes in radiologic scores and variant genotypes. The most extensively investigated gene to date is *HLA-DRB1*, which has also been shown to have the strongest effect on the severity of disease [77–79]. In addition, several studies have reported potential associations between various candidate genes and disease severity [80–84]. GWAS-identified loci have also been investigated, and a recent analysis involving a Japanese cohort demonstrated that polymorphisms in *PADI4* are associated with radiographic progression of RA [79]. More recently, a GWAS on the radiological progression rate in autoantibody-positive RA patients identified an association in an SNP at *SPAG16* gene, which is shown to be expressed in the synovial tissues of RA patients [85]. Although these lines of evidence indicate that genetic variants can influence on the severity of disease, individual alleles of these genes could not predict disease severity sufficiently for clinical practice use due to their moderate effect.

Another important clinical phenotype that ideally should be predictable based on genetic data is an individual's response to a drug. The advent of biologic therapies such as treatment with anti-TNF antibodies has revolutionized the treatment of RA, but a substantial proportion of patients (20–40%) will not respond to these therapies. In addition, some patients who do not respond to one biologic therapy (e.g., anti-TNF antibodies) may respond to another (e.g., anti-IL-6R antibodies). Therefore, if the response to a biologic agent can be predicted, unnecessary costs and potential side effects can be avoided. Several GWASs examining the response to anti-TNF antibody therapy provided evidence suggesting an association between drug response and genes involved in signaling, including the *CD84* locus [86–88]. However, as with the prediction of disease severity, individual loci cannot sufficiently predict an individual's drug response. The observations resulting from attempts to predict both disease severity and drug response clearly indicate that single genetic factors are insufficient for predicting clinical phenotype and that we need to establish a polygenic approach in combination with analyses of environmental factors using appropriate statistical models.

Acknowledgments

This work was supported by grants from the Ministry of Education, Culture, Sports, Sciences and Technology and the Ministry of Health, Labor and Welfare of the Japanese government.

References

[1] D. Aletaha, T. Neogi, A.J. Silman, J. Funovits, D.T. Felson, C.O. Bingham 3rd, N.S. Birnbaum, G.R. Burmester, V.P. Bykerk, M.D. Cohen, B. Combe, K.H. Costenbader, M. Dougados, P. Emery, G. Ferraccioli, J.M. Hazes, K. Hobbs, T.W. Huizinga, A. Kavanaugh, J. Kay, T.K. Kvien, T. Laing, P. Mease, H.A. Menard, L.W. Moreland, R.L. Naden, T. Pincus, J.S. Smolen, E. Stanislawska-Biernat, D. Symmons, P.P. Tak, K.S. Upchurch, J. Vencovsky, F. Wolfe, G. Hawker, Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative, Arthritis Rheum. 62 (2010) (2010) 2569–2581.

[2] G.A. Schellekens, B.A. de Jong, F.H. van den Hoogen, L.B. van de Putte, W.J. van Venrooij, Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies, J. Clin. Invest. 101 (1998) 273–281.

[3] T. Frisell, M. Holmqvist, H. Kallberg, L. Klareskog, L. Alfredsson, J. Askling, Familial risks and heritability of rheumatoid arthritis: role of rheumatoid factor/anti-citrullinated protein antibody status, number and type of affected relatives, sex, and age, Arthritis Rheum. 65 (2013) 2773–2782.

[4] A.J. Silman, A.J. MacGregor, W. Thomson, S. Holligan, D. Carthy, A. Farhan, W.E. Ollier, Twin concordance rates for rheumatoid arthritis: results from a nationwide study, Br. J. Rheumatol. 32 (1993) 903–907.

[5] A.J. MacGregor, H. Snieder, A.S. Rigby, M. Koskenvuo, J. Kaprio, K. Aho, A.J. Silman, Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins, Arthritis Rheum. 43 (2000) 30–37.

[6] D. van der Woude, J.J. Houwing-Duistermaat, R.E. Toes, T.W. Huizinga, W. Thomson, J. Worthington, A.H. van der Helm-van Mil, R.R. de Vries, Quantitative heritability of anti-citrullinated protein antibody-positive and

anti-citrullinated protein antibody-negative rheumatoid arthritis, Arthritis Rheum. 60 (2009) 916–923.
[7] International HapMap Consortium, The International HapMap Project, Nature 426 (2003) 789–796.
[8] K.P. Liao, L. Alfredsson, E.W. Karlson, Environmental influences on risk for rheumatoid arthritis, Curr. Opin. Rheumatol. 21 (2009) 279–283.
[9] R.A. Hoovestol, T.R. Mikuls, Environmental exposures and rheumatoid arthritis risk, Curr. Rheumatol. Rep. 13 (2011) 431–439.
[10] G.P. Astorga, R.C. Williams Jr., Altered reactivity in mixed lymphocyte culture of lymphocytes from patients with rheumatoid arthritis, Arthritis Rheum. 12 (1969) 547–554.
[11] A.J. McMichael, T. Sasazuki, H.O. McDevitt, R.O. Payne, Increased frequency of HLA-Cw3 and HLA-Dw4 in rheumatoid arthritis, Arthritis Rheum. 20 (1977) 1037–1042.
[12] P. Stastny, Association of the B-cell alloantigen DRw4 with rheumatoid arthritis, N. Engl. J. Med. 298 (1978) 869–871.
[13] L. Legrand, G.M. Lathrop, A. Marcelli-Barge, A. Dryll, T. Bardin, N. Debeyre, J.C. Poirier, M. Schmid, A. Ryckewaert, J. Dausset, HLA-DR genotype risks in seropositive rheumatoid arthritis, Am. J. Hum. Genet. 36 (1984) 690–699.
[14] P.K. Gregersen, J. Silver, R.J. Winchester, The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis, Arthritis Rheum. 30 (1987) 1205–1213.
[15] B.M. Freed, R.P. Schuyler, M.T. Aubrey, Association of the HLA-DRB1 epitope LA(67, 74) with rheumatoid arthritis and citrullinated vimentin binding, Arthritis Rheum. 63 (2011) 3733–3739.
[16] S. Raychaudhuri, C. Sandor, E.A. Stahl, J. Freudenberg, H.S. Lee, X. Jia, L. Alfredsson, L. Padyukov, L. Klareskog, J. Worthington, K.A. Siminovitch, S.C. Bae, R.M. Plenge, P.K. Gregersen, P.I. de Bakker, Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis, Nat. Genet. 44 (2012) 291–296.
[17] I. Auger, M. Sebbag, C. Vincent, N. Balandraud, S. Guis, L. Nogueira, B. Svensson, A. Cantagrel, G. Serre, J. Roudier, Influence of HLA-DR genes on the production of rheumatoid arthritis-specific autoantibodies to citrullinated fibrinogen, Arthritis Rheum. 52 (2005) 3424–3432.
[18] F.A. van Gaalen, J. van Aken, T.W. Huizinga, G.M. Schreuder, F.C. Breedveld, E. Zanelli, W.J. van Venrooij, C.L. Verweij, R.E. Toes, R.R. de Vries, Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis, Arthritis Rheum. 50 (2004) 2113–2121.
[19] S. Bas, T.V. Perneger, E. Mikhnevitch, M. Seitz, J.M. Tiercy, P. Roux-Lombard, P.A. Guerne, Association of rheumatoid factors and anti-filaggrin antibodies with severity of erosions in rheumatoid arthritis, Rheumatology (Oxford) 39 (2000) 1082–1088.
[20] K. Shimane, Y. Kochi, A. Suzuki, Y. Okada, T. Ishii, T. Horita, K. Saito, A. Okamoto, N. Nishimoto, K. Myouzen, M. Kubo, M. Hirakata, T. Sumida, Y. Takasaki, R. Yamada, Y. Nakamura, N. Kamatani, K. Yamamoto, An association analysis of HLA-DRB1 with systemic lupus erythematosus and rheumatoid arthritis in a Japanese population: effects of *09:01 allele on disease phenotypes, Rheumatology (Oxford) 52 (2013) 1172–1182.
[21] J.A. Hill, S. Southwood, A. Sette, A.M. Jevnikar, D.A. Bell, E. Cairns, Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule, J. Immunol. 171 (2003) 538–541.
[22] S.W. Scally, J. Petersen, S.C. Law, N.L. Dudek, H.J. Nel, K.L. Loh, L.C. Wijeyewickrema, S.B. Eckle, J. van Heemst, R.N. Pike, J. McCluskey, R.E. Toes, N.L. La Gruta, A.W. Purcell, H.H. Reid, R. Thomas, J. Rossjohn, A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis, J. Exp. Med. 210 (2013) 2569–2582.
[23] L. Klareskog, J. Ronnelid, K. Lundberg, L. Padyukov, L. Alfredsson, Immunity to citrullinated proteins in rheumatoid arthritis, Annu. Rev. Immunol. 26 (2008) 651–675.
[24] D. van der Woude, S. Rantapaa-Dahlqvist, A. Ioan-Facsinay, C. Onnekink, C.M. Schwarte, K.N. Verpoort, J.W. Drijfhout, T.W. Huizinga, R.E. Toes, G.J. Pruijn, Epitope spreading of the anti-citrullinated protein antibody response occurs before disease onset and is associated with the disease course of early arthritis, Ann. Rheum. Dis. 69 (2010) 1554–1561.
[25] L. Klareskog, P. Stolt, K. Lundberg, H. Kallberg, C. Bengtsson, J. Grunewald, J. Ronnelid, H.E. Harris, A.K. Ulfgren, S. Rantapaa-Dahlqvist, A. Eklund, L. Padyukov, L. Alfredsson, A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination, Arthritis Rheum. 54 (2006) 38–46.
[26] S.P. Linn-Rasker, A.H. van der Helm-van Mil, F.A. van Gaalen, M. Kloppenburg, R.R. de Vries, S. le Cessie, F.C. Breedveld, R.E. Toes, T.W. Huizinga, Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles, Ann. Rheum. Dis. 65 (2006) 366–371.
[27] H. Kallberg, L. Padyukov, R.M. Plenge, J. Ronnelid, P.K. Gregersen, A.H. van der Helm-van Mil, R.E. Toes, T.W. Huizinga, L. Klareskog, L. Alfredsson, Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis, Am. J. Hum. Genet. 80 (2007) 867–875.
[28] K. Lundberg, C. Bengtsson, N. Kharlamova, E. Reed, X. Jiang, H. Kallberg, I. Pollak-Dorocic, L. Israelsson, C. Kessel, L. Padyukov, R. Holmdahl, L. Alfredsson, L. Klareskog, Genetic and environmental determinants for disease risk in subsets of rheumatoid arthritis defined by the anticitrullinated protein/peptide antibody fine specificity profile, Ann. Rheum. Dis. 72 (2013) 652–658.
[29] H. Mahdi, B.A. Fisher, H. Kallberg, D. Plant, V. Malmstrom, J. Ronnelid, P. Charles, B. Ding, L. Alfredsson, L. Padyukov, D.P. Symmons, P.J. Venables, L. Klareskog, K. Lundberg, Specific interaction between genotype, smoking and autoimmunity to citrullinated alpha-enolase in the etiology of rheumatoid arthritis, Nat. Genet. 41 (2009) 1319–1324.
[30] K.N. Verpoort, F.A. van Gaalen, A.H. van der Helm-van Mil, G.M. Schreuder, F.C. Breedveld, T.W. Huizinga, R.R. de Vries, R.E. Toes, Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis, Arthritis Rheum. 52 (2005) 3058–3062.
[31] P. Irigoyen, A.T. Lee, M.H. Wener, W. Li, M. Kern, F. Batliwalla, R.F. Lum, E. Massarotti, M. Weisman, C. Bombardier, E.F. Remmers, D.L. Kastner, M.F. Seldin, L.A. Criswell, P.K. Gregersen, Regulation of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: contrasting effects of HLA-DR3 and the shared epitope alleles, Arthritis Rheum. 52 (2005) 3813–3818.
[32] C. Terao, K. Ohmura, K. Ikari, Y. Kochi, E. Maruya, M. Katayama, K. Yurugi, K. Shimada, A. Murasawa, S. Honjo, K. Takasugi, K. Matsuo, K. Tajima, A. Suzuki, K. Yamamoto, S. Momohara, H. Yamanaka, R. Yamada, H. Saji, F. Matsuda, T. Mimori, ACPA-negative RA consists of two genetically distinct subsets based on RF positivity in Japanese, PLoS One 7 (2012) e40067.
[33] B. Han, D. Diogo, S. Eyre, H. Kallberg, A. Zhernakova, J. Bowes, L. Padyukov, Y. Okada, M.A. Gonzalez-Gay, S. Rantapaa-Dahlqvist, J. Martin, T.W. Huizinga, R.M. Plenge, J. Worthington, P.K. Gregersen, L. Klareskog, P.I. de Bakker, S. Raychaudhuri, Fine mapping seronegative and seropositive rheumatoid arthritis to shared and distinct HLA alleles by adjusting for the effects of heterogeneity, Am. J. Hum. Genet. 94 (2014) 522–532.
[34] A. Suzuki, R. Yamada, X. Chang, S. Tokuhiro, T. Sawada, M. Suzuki, M. Nagasaki, M. Nakayama-Hamada, R. Kawaida, M. Ono, M. Ohtsuki, H. Furukawa, S. Yoshino, M. Yukioka, S. Tohma, T. Matsubara, S. Wakitani, R. Teshima, Y. Nishioka, A. Sekine, A. Iida, A. Takahashi, T. Tsunoda, Y. Nakamura, K. Yamamoto, Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis, Nat. Genet. 34 (2003) 395–402.
[35] Wellcome Trust Case Control Consortium, Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls, Nature 447 (2007) 661–678.
[36] W. Thomson, A. Barton, X. Ke, S. Eyre, A. Hinks, J. Bowes, R. Donn, D. Symmons, S. Hider, I.N. Bruce, A.G. Wilson, I. Marinou, A. Morgan, P. Emery, A. Carter, S. Steer, L. Hocking, D.M. Reid, P. Wordsworth, P. Harrison, D. Strachan, J. Worthington, Rheumatoid arthritis association at 6q23, Nat. Genet. 39 (2007) 1431–1433.
[37] R.M. Plenge, M. Seielstad, L. Padyukov, A.T. Lee, E.F. Remmers, B. Ding, A. Liew, H. Khalili, A. Chandrasekaran, L.R. Davies, W. Li, A.K. Tan, C. Bonnard, R.T. Ong, A. Thalamuthu, S. Pettersson, C. Liu, C. Tian, W.V. Chen, J.P. Carulli, E.M. Beckman, D. Altshuler, L. Alfredsson, L.A. Criswell, C.I. Amos, M.F. Seldin, D.L. Kastner, L. Klareskog, P.K. Gregersen, TRAF1-C5 as a risk locus for rheumatoid arthritis – a genomewide study, N. Engl. J. Med. 357 (2007) 1199–1209.
[38] P.K. Gregersen, C.I. Amos, A.T. Lee, Y. Lu, E.F. Remmers, D.L. Kastner, M.F. Seldin, L.A. Criswell, R.M. Plenge, V.M. Holers, T.R. Mikuls, T. Sokka, L.W. Moreland, S.L. Bridges Jr., G. Xie, A.B. Begovich, K.A. Siminovitch, REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis, Nat. Genet. 41 (2009) 820–823.
[39] Y. Kochi, Y. Okada, A. Suzuki, K. Ikari, C. Terao, A. Takahashi, K. Yamazaki, N. Hosono, K. Myouzen, T. Tsunoda, N. Kamatani, T. Furuichi, S. Ikegawa, K. Ohmura, T. Mimori, F. Matsuda, T. Iwamoto, S. Momohara, H. Yamanaka, R. Yamada, M. Kubo, Y. Nakamura, K. Yamamoto, A regulatory variant in CCR6 is associated with rheumatoid arthritis susceptibility, Nat. Genet. 42 (2010) 515–519.
[40] E.A. Stahl, S. Raychaudhuri, E.F. Remmers, G. Xie, S. Eyre, B.P. Thomson, Y. Li, F.A. Kurreeman, A. Zhernakova, A. Hinks, C. Guiducci, R. Chen, L. Alfredsson, C.I. Amos, K.G. Ardlie, A. Barton, J. Bowes, E. Brouwer, N.P. Burtt, J.J. Catanese, J. Coblyn, M.J. Coenen, K.H. Costenbader, L.A. Criswell, J.B. Crusius, J. Cui, P.I. de Bakker, P.L. De Jager, B. Ding, P. Emery, E. Flynn, P. Harrison, L.J. Hocking, T.W. Huizinga, D.L. Kastner, X. Ke, A.T. Lee, X. Liu, P. Martin, A.W. Morgan, L. Padyukov, M.D. Posthumus, T.R. Radstake, D.M. Reid, M. Seielstad, M.F. Seldin, N.A. Shadick, S. Steer, P.P. Tak, W. Thomson, A.H. van der Helm-van Mil, I.E. van der Horst-Bruinsma, C.E. van der Schoot, P.L. van Riel, M.E. Weinblatt, A.G. Wilson, G.J. Wolbink, B.P. Wordsworth, C. Wijmenga, E.W. Karlson, R.E. Toes, N. de Vries, A.B. Begovich, J. Worthington, K.A. Siminovitch, P.K. Gregersen, L. Klareskog, R.M. Plenge, Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci, Nat. Genet. 42 (2010) 508–514.
[41] Y. Okada, C. Terao, K. Ikari, Y. Kochi, K. Ohmura, A. Suzuki, T. Kawaguchi, E.A. Stahl, F.A. Kurreeman, N. Nishida, H. Ohmiya, K. Myouzen, M. Takahashi, T. Sawada, Y. Nishioka, M. Yukioka, T. Matsubara, S. Wakitani, R. Teshima, S. Tohma, K. Takasugi, K. Shimada, A. Murasawa, S. Honjo, K. Matsuo, H. Tanaka, K. Tajima, T. Suzuki, T. Iwamoto, Y. Kawamura, H. Tanii, Y. Okazaki, T. Sasaki, P.K. Gregersen, L. Padyukov, J. Worthington, K.A. Siminovitch, M. Lathrop, A. Taniguchi, A. Takahashi, K. Tokunaga, M. Kubo, Y. Nakamura, N. Kamatani, T. Mimori, R.M. Plenge, H. Yamanaka, S. Momohara, R. Yamada, F. Matsuda, K. Yamamoto, Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population, Nat. Genet. 44 (2012) 511–516.
[42] Y. Okada, D. Wu, G. Trynka, T. Raj, C. Terao, K. Ikari, Y. Kochi, K. Ohmura, A. Suzuki, S. Yoshida, R.R. Graham, A. Manoharan, W. Ortmann, T. Bhangale, J.C. Denny, R.J. Carroll, A.E. Eyler, J.D. Greenberg, J.M. Kremer, D.A. Pappas, L. Jiang, J. Yin, L. Ye, D.F. Su, J. Yang, G. Xie, E. Keystone, H.J. Westra, T. Esko, A. Metspalu, X. Zhou, N. Gupta, D. Mirel, E.A. Stahl, D. Diogo, J. Cui, K. Liao, M.H.

Guo, K. Myouzen, T. Kawaguchi, M.J. Coenen, P.L. van Riel, M.A. van de Laar, H.J. Guchelaar, T.W. Huizinga, P. Dieude, X. Mariette, S.L. Bridges Jr., A. Zhernakova, R.E. Toes, P.P. Tak, C. Miceli-Richard, S.Y. Bang, H.S. Lee, J. Martin, M.A. Gonzalez-Gay, L. Rodriguez-Rodriguez, S. Rantapaa-Dahlqvist, L. Arlestig, H.K. Choi, Y. Kamatani, P. Galan, M. Lathrop, R. consortium, G. consortium, S. Eyre, J. Bowes, N. de Vries, L.W. Moreland, L.A. Criswell, E.W. Karlson, A. Taniguchi, R. Yamada, M. Kubo, J.S. Liu, S.C. Bae, J. Worthington, L. Padyukov, L. Klareskog, P.K. Gregersen, S. Raychaudhuri, B.E. Stranger, P.L. De Jager, L. Franke, P.M. Visscher, M.A. Brown, H. Yamanaka, T. Mimori, A. Takahashi, H. Xu, T.W. Behrens, K.A. Siminovitch, S. Momohara, F. Matsuda, K. Yamamoto, R.M. Plenge, Genetics of rheumatoid arthritis contributes to biology and drug discovery, Nature 506 (2014) 376–381.

[43] H.J. Westra, M.J. Peters, T. Esko, H. Yaghootkar, C. Schurmann, J. Kettunen, M.W. Christiansen, B.P. Fairfax, K. Schramm, J.E. Powell, A. Zhernakova, D.V. Zhernakova, J.H. Veldink, L.H. Van den Berg, J. Karjalainen, S. Withoff, A.G. Uitterlinden, A. Hofman, F. Rivadeneira, P.A. t Hoen, E. Reinmaa, K. Fischer, M. Nelis, L. Milani, D. Melzer, L. Ferrucci, A.B. Singleton, D.G. Hernandez, M.A. Nalls, G. Homuth, M. Nauck, D. Radke, U. Volker, M. Perola, V. Salomaa, J. Brody, A. Suchy-Dicey, S.A. Gharib, D.A. Enquobahrie, T. Lumley, G.W. Montgomery, S. Makino, H. Prokisch, C. Herder, M. Roden, H. Grallert, T. Meitinger, K. Strauch, Y. Li, R.C. Jansen, P.M. Visscher, J.C. Knight, B.M. Psaty, S. Ripatti, A. Teumer, T.M. Frayling, A. Metspalu, J.B. van Meurs, L. Franke, Systematic identification of trans eQTLs as putative drivers of known disease associations, Nat. Genet. 45 (2013) 1238–1243.

[44] P.C. Dubois, G. Trynka, L. Franke, K.A. Hunt, J. Romanos, A. Curtotti, A. Zhernakova, G.A. Heap, R. Adany, A. Aromaa, M.T. Bardella, L.H. van den Berg, N.A. Bockett, E.G. de la Concha, B. Dema, R.S. Fehrmann, M. Fernandez-Arquero, S. Fiatal, E. Grandone, P.M. Green, H.J. Groen, R. Gwilliam, R.H. Houwen, S.E. Hunt, K. Kaukinen, D. Kelleher, I. Korponay-Szabo, K. Kurppa, P. MacMathuna, M. Maki, M.C. Mazzilli, O.T. McCann, M.L. Mearin, C.A. Mein, M.M. Mirza, V. Mistry, B. Mora, K.I. Morley, C.J. Mulder, J.A. Murray, C. Nunez, E. Oosterom, R.A. Ophoff, I. Polanco, L. Peltonen, M. Platteel, A. Rybak, V. Salomaa, J.J. Schweizer, M.P. Sperandeo, G.J. Tack, G. Turner, J.H. Veldink, W.H. Verbeek, R.K. Weersma, V.M. Wolters, E. Urcelay, B. Cukrowska, L. Greco, S.L. Neuhausen, R. McManus, D. Barisani, P. Deloukas, J.C. Barrett, P. Saavalainen, C. Wijmenga, D.A. van Heel, Multiple common variants for celiac disease influencing immune gene expression, Nat. Genet. 42 (2010) 295–302.

[45] Y. Okada, K. Shimane, Y. Kochi, T. Tahira, A. Suzuki, K. Higasa, A. Takahashi, T. Horita, T. Atsumi, T. Ishii, A. Okamoto, K. Fujio, M. Hirakata, H. Amano, Y. Kondo, S. Ito, K. Takada, A. Mimori, K. Saito, M. Kamachi, Y. Kawaguchi, K. Ikari, O.W. Mohammed, K. Matsuda, C. Terao, K. Ohmura, K. Myouzen, N. Hosono, T. Tsunoda, N. Nishimoto, T. Mimori, F. Matsuda, Y. Tanaka, T. Sumida, H. Yamanaka, Y. Takasaki, T. Koike, T. Horiuchi, K. Hayashi, M. Kubo, N. Kamatani, R. Yamada, Y. Nakamura, K. Yamamoto, A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus erythematosus in Japanese, PLoS Genet. 8 (2012) e1002455.

[46] E.P. Consortium, A. Kundaje, S.F. Aldred, P.J. Collins, C.A. Davis, F. Doyle, C.B. Epstein, S. Frietze, J. Harrow, R. Kaul, J. Khatun, B.R. Lajoie, S.G. Landt, B.K. Lee, F. Pauli, K.R. Rosenbloom, P. Sabo, A. Safi, A. Sanyal, N. Shoresh, J.M. Simon, L. Song, N.D. Trinklein, R.C. Altshuler, E. Birney, J.B. Brown, C. Cheng, S. Djebali, X. Dong, J. Ernst, T.S. Furey, M. Gerstein, B. Giardine, M. Greven, R.C. Hardison, R.S. Harris, J. Herrero, M.M. Hoffman, S. Iyer, M. Kelllis, P. Kheradpour, T. Lassmann, Q. Li, X. Lin, G.K. Marinov, A. Merkel, A. Mortazavi, S.C. Parker, T.E. Reddy, J. Rozowsky, F. Schlesinger, R.E. Thurman, J. Wang, L.D. Ward, T.W. Whitfield, S.P. Wilder, W. Wu, H.S. Xi, K.Y. Yip, J. Zhuang, B.E. Bernstein, E.D. Green, C. Gunter, M. Snyder, M.J. Pazin, R.F. Lowdon, L.A. Dillon, L.B. Adams, C.J. Kelly, J. Zhang, J.R. Wexler, P.J. Good, E.A. Feingold, G.E. Crawford, J. Dekker, L. Elinitski, P.J. Farnham, M.C. Giddings, T.R. Gingeras, R. Guigo, T.J. Hubbard, M. Kellis, W.J. Kent, J.D. Lieb, E.H. Margulies, R.M. Myers, J.A. Starnatoyannopoulos, S.A. Tennebaum, Z. Weng, K.P. White, B. Wold, Y. Yu, J. Wrobel, B.A. Risk, H.P. Gunawardena, H.C. Kuiper, C.W. Maier, L. Xie, X. Chen, et al., An integrated encyclopedia of DNA elements in the human genome, Nature 489 (2012) 57–74.

[47] G. Trynka, C. Sandor, B. Han, H. Xu, B.E. Stranger, X.S. Liu, S. Raychaudhuri, Chromatin marks identify critical cell types for fine mapping complex trait variants, Nat. Genet. 45 (2013) 124–130.

[48] Y. Kochi, A. Suzuki, R. Yamada, K. Yamamoto, Ethnogenetic heterogeneity of rheumatoid arthritis-implications for pathogenesis, Nat. Rev. Rheumatol. 6 (2010) 290–295.

[49] Y. Kochi, M.M. Thabet, A. Suzuki, Y. Okada, N.A. Daha, R.E. Toes, T.W. Huizinga, K. Myouzen, M. Kubo, R. Yamada, Y. Nakamura, K. Yamamoto, PADI4 polymorphism predisposes male smokers to rheumatoid arthritis, Ann. Rheum. Dis. 70 (2011) 512–515.

[50] A.B. Begovich, V.E. Carlton, L.A. Honigberg, S.J. Schrodi, A.P. Chokkalingam, H.C. Alexander, K.G. Ardlie, Q. Huang, A.M. Smith, J.M. Spoerke, M.T. Conn, M. Chang, S.Y. Chang, R.K. Saiki, J.J. Catanese, D.U. Leong, V.E. Garcia, L.B. McAllister, D.A. Jeffery, A.T. Lee, F. Batliwalla, E. Remmers, L.A. Criswell, M.F. Seldin, D.L. Kastner, C.I. Amos, J.J. Sninsky, P.K. Gregersen, A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis, Am. J. Hum. Genet. 75 (2004) 330–337.

[51] J. Zheng, S. Ibrahim, F. Petersen, X. Yu, Meta-analysis reveals an association of PTPN22 C1858T with autoimmune diseases, which depends on the localization of the affected tissue, Genes Immun. 13 (2012) 641–652.

[52] M. Mori, R. Yamada, K. Kobayashi, R. Kawaida, K. Yamamoto, Ethnic differences in allele frequency of autoimmune-disease-associated SNPs, J. Hum. Genet. 50 (2005) 264–266.

[53] T.C. Lins, R.G. Vieira, D. Grattapaglia, R.W. Pereira, Allele and haplotype frequency distribution in PTPN22 gene across variable ethnic groups: Implications for genetic association studies for autoimmune diseases, Autoimmunity 43 (2010) 308–316.

[54] A.F. Arechiga, T. Habib, Y. He, X. Zhang, Z.Y. Zhang, A. Funk, J.H. Buckner, Cutting edge: the PTPN22 allelic variant associated with autoimmunity impairs B cell signaling, J. Immunol. 182 (2009) 3343–3347.

[55] T. Vang, M. Congia, M.D. Macis, L. Musumeci, V. Orru, P. Zavattari, K. Nika, L. Tautz, K. Tasken, F. Cucca, T. Mustelin, N. Bottini, Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant, Nat. Genet. 37 (2005) 1317–1319.

[56] J. Zhang, N. Zahir, Q. Jiang, H. Miliotis, S. Heyraud, X. Meng, B. Dong, G. Xie, F. Qiu, Z. Hao, C.A. McCulloch, E.C. Keystone, A.C. Peterson, K.A. Siminovitch, The autoimmune disease-associated PTPN22 variant promotes calpain-mediated Lyp/Pep degradation associated with lymphocyte and dendritic cell hyperresponsiveness, Nat. Genet. 43 (2011) 902–907.

[57] X. Dai, R.G. James, T. Habib, S. Singh, S. Jackson, S. Khim, R.T. Moon, D. Liggitt, A. Wolf-Yadlin, J.H. Buckner, D.J. Rawlings, A disease-associated PTPN22 variant promotes systemic autoimmunity in murine models, J. Clin. Invest. 123 (2013) 2024–2036.

[58] J. Zheng, F. Petersen, X. Yu, The role of PTPN22 in autoimmunity: learning from mice, Autoimmun. Rev. 13 (2014) 266–271.

[59] R.R. Graham, C. Cotsapas, L. Davies, R. Hackett, C.J. Lessard, J.M. Leon, N.P. Burtt, C. Guiducci, M. Parkin, C. Gates, R.M. Plenge, T.W. Behrens, J.E. Wither, J.D. Rioux, P.R. Fortin, D.C. Graham, A.K. Wong, T.J. Vyse, M.J. Daly, D. Altshuler, K.L. Moser, P.M. Gaffney, Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus, Nat. Genet. 40 (2008) 1059–1061.

[60] S.L. Musone, K.E. Taylor, T.T. Lu, J. Nititham, R.C. Ferreira, W. Ortmann, N. Shifrin, M.A. Petri, M. Ilyas Kamboh, S. Manzi, M.F. Seldin, P.K. Gregersen, T.W. Behrens, A. Ma, P.Y. Kwok, L.A. Criswell, Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus, Nat. Genet. 40 (2008) 1062–1064.

[61] I. Adrianto, F. Wen, A. Templeton, G. Wiley, J.B. King, C.J. Lessard, J.S. Bates, Y. Hu, J.A. Kelly, K.M. Kaufman, J.M. Guthridge, M.E. Alarcon-Riquelme, Biolupus, G. Networks, J.M. Anaya, S.C. Bae, S.Y. Bang, S.A. Boackle, E.E. Brown, M.A. Petri, C. Gallant, R. Ramsey-Goldman, J.D. Reveille, L.M. Vila, L.A. Criswell, J.C. Edberg, B.I. Freedman, P.K. Gregersen, G.S. Gilkeson, C.O. Jacob, J.A. James, D.L. Kamen, R.P. Kimberly, J. Martin, J.T. Merrill, T.B. Niewold, S.Y. Park, B.A. Pons-Estel, R.H. Scofield, A.M. Stevens, B.P. Tsao, T.J. Vyse, C.D. Langefeld, J.B. Harley, K.L. Moser, C.F. Webb, M.B. Humphrey, C.G. Montgomery, P.M. Gaffney, Association of a functional variant downstream of TNFAIP3 with systemic lupus erythematosus, Nat. Genet. 43 (2011) 253–258.

[62] M. Matmati, P. Jacques, J. Maelfait, E. Verheugen, M. Kool, M. Sze, L. Geboes, E. Louagie, C. Mc Guire, L. Vereecke, Y. Chu, L. Boon, S. Staelens, P. Matthys, B.N. Lambrecht, M. Schmidt-Supprian, M. Pasparakis, D. Elewaut, R. Beyaert, G. van Loo, A20 (TNFAIP3) deficiency in myeloid cells triggers erosive polyarthritis resembling rheumatoid arthritis, Nat. Genet. 43 (2011) 908–912.

[63] M. Kato, M. Sanada, I. Kato, Y. Sato, J. Takita, K. Takeuchi, A. Niwa, Y. Chen, K. Nakazaki, J. Nomoto, Y. Asakura, S. Muto, A. Tamura, M. Iio, Y. Akatsuka, Y. Hayashi, H. Mori, T. Igarashi, M. Kurokawa, S. Chiba, S. Mori, Y. Ishikawa, K. Okamoto, K. Tobinai, H. Nakagama, T. Nakahata, T. Yoshino, Y. Kobayashi, S. Ogawa, Frequent inactivation of A20 in B-cell lymphomas, Nature 459 (2009) 712–716.

[64] E.F. Remmers, R.M. Plenge, A.T. Lee, R.R. Graham, G. Hom, T.W. Behrens, P.I. de Bakker, J.M. Le, H.S. Lee, F. Batliwalla, W. Li, S.L. Masters, M.G. Booty, J.P. Carulli, L. Padyukov, L. Alfredsson, L. Klareskog, W.V. Chen, C.I. Amos, L.A. Criswell, M.F. Seldin, D.L. Kastner, P.K. Gregersen, STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus, N. Engl. J. Med. 357 (2007) 977–986.

[65] J. Zheng, J. Yin, R. Huang, F. Petersen, X. Yu, Meta-analysis reveals an association of STAT4 polymorphisms with systemic autoimmune disorders and anti-dsDNA antibody, Hum. Immunol. 74 (2013) 986–992.

[66] S. Sigurdsson, G. Nordmark, S. Garnier, E. Grundberg, T. Kwan, O. Nilsson, M.L. Eloranta, I. Gunnarsson, E. Svenungsson, G. Sturfelt, A.A. Bengtsson, A. Jonsen, L. Truedsson, S. Rantapaa-Dahlqvist, C. Eriksson, G. Alm, H.H. Goring, T. Pastinen, A.C. Syvanen, L. Ronnblom, A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with anti-dsDNA and shows additive effects with two risk alleles of IRF5, Hum. Mol. Genet. 17 (2008) 2868–2876.

[67] S.N. Kariuki, K.A. Kirou, E.J. MacDermott, L. Barillas-Arias, M.K. Crow, T.B. Niewold, Cutting edge: autoimmune disease risk variant of STAT4 confers increased sensitivity to IFN-alpha in lupus patients in vivo, J. Immunol. 182 (2009) 34–38.

[68] J.C. Barrett, S. Hansoul, D.L. Nicolae, J.H. Cho, R.H. Duerr, J.D. Rioux, S.R. Brant, M.S. Silverberg, K.D. Taylor, M.M. Barmada, A. Bitton, T. Dassopoulos, L.W. Datta, T. Green, A.M. Griffiths, E.O. Kistner, M.T. Murtha, M.D. Regueiro, J.I. Rotter, L.P. Schumm, A.H. Steinhart, S.R. Targan, R.J. Xavier, C. Libioulle, C. Sandor, M. Lathrop, J. Belaiche, O. Dewit, I. Gut, S. Heath, D. Laukens, M. Mni, P. Rutgeerts, A. Van Gossum, D. Zelenika, D. Franchimont, J.P. Hugot, M. de Vos, S. Vermeire, E. Louis, L.R. Cardon, C.A. Anderson, H. Drummond, E. Nimmo, T. Ahmad, N.J. Prescott, C.M. Onnie, S.A. Fisher, J. Marchini, J. Ghori, S. Bumpstead,

R. Gwilliam, M. Tremelling, P. Deloukas, J. Mansfield, D. Jewell, J. Satsangi, C.G. Mathew, M. Parkes, M. Georges, M.J. Daly, Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease, Nat. Genet. 40 (2008) 955–962.

[69] L.C. Tsoi, S.L. Spain, J. Knight, E. Ellinghaus, P.E. Stuart, F. Capon, J. Ding, Y. Li, T. Tejasvi, J.E. Gudjonsson, H.M. Kang, M.H. Allen, R. McManus, G. Novelli, L. Samuelsson, J. Schalkwijk, M. Stahle, A.D. Burden, C.H. Smith, M.J. Cork, X. Estivill, A.M. Bowcock, G.G. Krueger, W. Weger, J. Worthington, R. Tazi-Ahnini, F.O. Nestle, A. Hayday, P. Hoffmann, J. Winkelmann, C. Wijmenga, C. Langford, S. Edkins, R. Andrews, H. Blackburn, A. Strange, G. Band, R.D. Pearson, D. Vukcevic, C.C. Spencer, P. Deloukas, U. Mrowietz, S. Schreiber, S. Weidinger, S. Koks, K. Kingo, T. Esko, A. Metspalu, H.W. Lim, J.J. Voorhees, M. Weichenthal, H.E. Wichmann, V. Chandran, C.F. Rosen, P. Rahman, D.D. Gladman, C.E. Griffiths, A. Reis, J. Kere, Collaborative Association Study of Psoriasis, Genetic Analysis of Psoriasis Consortium, Psoriasis Association Genetics Extension, Wellcome Trust Case Control Consortium, R.P. Nair, A. Franke, J.N. Barker, G.R. Abecasis, J.T. Elder, R.C. Trembath, Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity, Nat. Genet. 44 (2012) 1341–1348.

[70] E. Jakkula, V. Leppa, A.M. Sulonen, T. Varilo, S. Kallio, A. Kemppinen, S. Purcell, K. Koivisto, P. Tienari, M.L. Sumelahti, I. Elovaara, T. Pirttila, M. Reunanen, A. Aromaa, A.B. Oturai, H.B. Sondergaard, H.F. Harbo, I.L. Mero, S.B. Gabriel, D.B. Mirel, S.L. Hauser, L. Kappos, C. Polman, P.L. De Jager, D.A. Hafler, M.J. Daly, A. Palotie, J. Saarela, L. Peltonen, Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene, Am. J. Hum. Genet. 86 (2010) 285–291.

[71] K. Hirota, H. Yoshitomi, M. Hashimoto, S. Maeda, S. Teradaira, N. Sugimoto, T. Yamaguchi, T. Nomura, H. Ito, T. Nakamura, N. Sakaguchi, S. Sakaguchi, Preferential recruitment of CCR6-expressing Th17 cells to inflamed joints via CCL20 in rheumatoid arthritis and its animal model, J. Exp. Med. 204 (2007) 2803–2812.

[72] N. Komatsu, K. Okamoto, S. Sawa, T. Nakashima, M. Oh-Hora, T. Kodama, S. Tanaka, J.A. Bluestone, H. Takayanagi, Pathogenic conversion of Foxp3(+) T cells into TH17 cells in autoimmune arthritis, Nat. Med. 20 (2014) 62–68.

[73] W.Y. Wang, B.J. Barratt, D.G. Clayton, J.A. Todd, Genome-wide association studies: theoretical and practical concerns, Nat. Rev. Genet. 6 (2005) 109–118.

[74] E.A. Stahl, D. Wegmann, G. Trynka, J. Gutierrez-Achury, R. Do, B.F. Voight, P. Kraft, R. Chen, H.J. Kallberg, F.A. Kurreeman, Diabetes Genetics Replication and Meta-analysis Consortium, Myocardial Infarction Genetics Consortium, S. Kathiresan, C. Wijmenga, P.K. Gregersen, L. Alfredsson, K.A. Siminovitch, J. Worthington, P.I. de Bakker, S. Raychaudhuri, R.M. Plenge, Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis, Nat. Genet. 44 (2012) 483–489.

[75] A. Kiezun, K. Garimella, R. Do, N.O. Stitziel, B.M. Neale, P.J. McLaren, N. Gupta, P. Sklar, P.F. Sullivan, J.L. Moran, C.M. Hultman, P. Lichtenstein, P. Magnusson, T. Lehner, Y.Y. Shugart, A.L. Price, P.I. de Bakker, S.M. Purcell, S.R. Sunyaev, Exome sequencing and the genetic basis of complex traits, Nat. Genet. 44 (2012) 623–630.

[76] D. Diogo, F. Kurreeman, E.A. Stahl, K.P. Liao, N. Gupta, J.D. Greenberg, M.A. Rivas, B. Hickey, J. Flannick, B. Thomson, C. Guiducci, S. Ripke, I. Adzhubey, A. Barton, J.M. Kremer, L. Alfredsson, Consortium of Rheumatology Researchers of North America, Rheumatoid Arthritis Consortium International, S. Sunyaev, J. Martin, A. Zhernakova, J. Bowes, S. Eyre, K.A. Siminovitch, P.K. Gregersen, J. Worthington, L. Klareskog, L. Padyukov, S. Raychaudhuri, R.M. Plenge, Rare, low-frequency, and common variants in the protein-coding sequence of biological candidate genes from GWASs contribute to risk of rheumatoid arthritis, Am. J. Hum. Genet. 92 (2013) 15–27.

[77] A. MacGregor, W. Ollier, W. Thomson, D. Jawaheer, A. Silman, HLA-DRB1*0401/0404 genotype and rheumatoid arthritis: increased association in men, young age at onset, and disease severity, J. Rheumatol. 22 (1995) 1032–1036.

[78] C.M. Weyand, K.C. Hicok, D.L. Conn, J.J. Goronzy, The influence of HLA-DRB1 genes on disease severity in rheumatoid arthritis, Ann. Intern. Med. 117 (1992) 801–806.

[79] T. Suzuki, K. Ikari, K. Yano, E. Inoue, Y. Toyama, A. Taniguchi, H. Yamanaka, S. Momohara, PADI4 and HLA-DRB1 are genetic risks for radiographic progression in RA patients, independent of ACPA status: results from the IORRA cohort study, PLoS One 8 (2013) e61045.

[80] A. Barton, J. Bowes, S. Eyre, D. Symmons, J. Worthington, A. Silman, Investigation of polymorphisms in the PADI4 gene in determining severity of inflammatory polyarthritis, Ann. Rheum. Dis. 64 (2005) 1311–1315.

[81] B. Joven, N. Gonzalez, F. Aguilar, B. Santiago, M. Galindo, J. Alcami, J.L. Pablos, Association between stromal cell-derived factor 1 chemokine gene variant and radiographic progression of rheumatoid arthritis, Arthritis Rheum. 52 (2005) 354–356.

[82] R. Knevel, A. Krabben, A.G. Wilson, E. Brouwer, M.K. Leijsma, E. Lindqvist, D.P. de Rooy, N.A. Daha, M.P. van der Linden, S. Tsonaka, A. Zhernakova, H.J. Westra, L. Franke, J.J. Houwing-Duistermaat, R.E. Toes, T.W. Huizinga, T. Saxne, A.H. van der Helm-van Mil, A genetic variant in granzyme B is associated with progression of joint destruction in rheumatoid arthritis, Arthritis Rheum. 65 (2013) 582–589.

[83] R.J. Mathews, J.I. Robinson, M. Battellino, C. Wong, J.C. Taylor, G. Biologics in Rheumatoid Arthritis, S. Genomics Study, S. Eyre, S.M. Churchman, A.G. Wilson, J.D. Isaacs, K. Hyrich, A. Barton, D. Plant, S. Savic, G.P. Cook, P. Sarzi-Puttini, P. Emery, J.H. Barrett, A.W. Morgan, M.F. McDermott, Evidence of NLRP3-inflammasome activation in rheumatoid arthritis (RA); genetic variants within the NLRP3-inflammasome complex in relation to susceptibility to RA and response to anti-TNF treatment, Ann. Rheum. Dis. 73 (2014) 1202–1210.

[84] D.P. de Rooy, A. Zhernakova, R. Tsonaka, A. Willemze, B.A. Kurreeman, G. Trynka, L. van Toorn, R.E. Toes, T.W. Huizinga, J.J. Houwing-Duistermaat, P.K. Gregersen, A.H. van der Helm-van Mil, A genetic variant in the region of MMP-9 is associated with serum levels and progression of joint damage in rheumatoid arthritis, Ann. Rheum. Dis. 73 (2014) 1163–1169.

[85] R. Knevel, K. Klein, K. Somers, C. Ospelt, J.J. Houwing-Duistermaat, J.A. van Nies, D.P. de Rooy, L. de Bock, F.A. Kurreeman, J. Schonkeren, G. Stoeken-Rijsbergen, Q. Helmer, M.P. van der Linden, M. Kern, N. Manjarrez-Orduno, L. Rodriguez-Rodriquez, P. Stinissen, T.W. Huizinga, R.E. Toes, S. Gay, P.K. Gregersen, V. Somers, A.H. van der Helm-van Mil, Identification of a genetic variant for joint damage progression in autoantibody-positive rheumatoid arthritis, Ann. Rheum. Dis. (2013), Available online.

[86] C. Liu, F. Batliwalla, W. Li, A. Lee, R. Roubenoff, E. Beckman, H. Khalili, A. Damle, M. Kern, R. Furie, J. Dupuis, R.M. Plenge, M.J. Coenen, T.W. Behrens, J.P. Carulli, P.K. Gregersen, Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in rheumatoid arthritis, Mol. Med. 14 (2008) 575–581.

[87] D. Plant, J. Bowes, C. Potter, K.L. Hyrich, A.W. Morgan, A.G. Wilson, J.D. Isaacs, Wellcome Trust Case Control Consortium, British Society for Rheumatology Biologics Register, A. Barton, Genome-wide association study of genetic predictors of anti-tumor necrosis factor treatment efficacy in rheumatoid arthritis identifies associations with polymorphisms at seven loci, Arthritis Rheum. 63 (2011) 645–653.

[88] J. Cui, E.A. Stahl, S. Saevarsdottir, C. Miceli, D. Diogo, G. Trynka, T. Raj, M.U. Mirkov, H. Canhao, K. Ikari, C. Terao, Y. Okada, S. Wedren, J. Askling, H. Yamanaka, S. Momohara, A. Taniguchi, K. Ohmura, F. Matsuda, T. Mimori, N. Gupta, M. Kuchroo, A.W. Morgan, J.D. Isaacs, A.G. Wilson, K.L. Hyrich, M. Herenius, M.E. Doorenspleet, P.P. Tak, J.B. Crusius, I.E. van der Horst-Bruinsma, G.J. Wolbink, P.L. van Riel, M. van de Laar, H.J. Guchelaar, N.A. Shadick, C.F. Allaart, T.W. Huizinga, R.E. Toes, R.P. Kimberly, S.L. Bridges Jr., L.A. Criswell, L.W. Moreland, J.E. Fonseca, N. de Vries, B.E. Stranger, P.L. De Jager, S. Raychaudhuri, M.E. Weinblatt, P.K. Gregersen, X. Mariette, A. Barton, L. Padyukov, M.J. Coenen, E.W. Karlson, R.M. Plenge, Genome-wide association study and gene expression analysis identifies CD84 as a predictor of response to etanercept therapy in rheumatoid arthritis, PLoS Genet. 9 (2013) e1003394.

APLAR — OPINION

Genetics of rheumatoid arthritis in Asia—present and future

Kazuhiko Yamamoto, Yukinori Okada, Akari Suzuki and Yuta Kochi

Abstract | Genome-wide association studies (GWAS) have uncovered numerous susceptibility genes for rheumatoid arthritis (RA) in patients of European, Asian and other ethnic ancestries. Although previous transethnic GWAS meta-analyses enabled the identification of several novel loci, the genetic heterogeneity observed in the *PADI4* and *PTPN22* genes suggests that ethnic variation should be considered. In addition, the effects of genetic polymorphisms on gene expression profiles are important when assessing the association of genetic information with disease pathogenesis and will influence the development of personalized medicine. Gene expression is controlled by epigenetic modifications, which in turn can be affected by environmental stimuli. Altogether, genetic and epigenetic information of Asian populations will contribute considerably to future rheumatology research.

Yamamoto *et al. Nat. Rev. Rheumatol.* advance online publication XX Month 2015; doi:10.1038/nrrheum.2015.7

Introduction

Clinical practice of rheumatology in Asia has been tremendously improved for more than a decade, mainly owing to the implementation of effective DMARDs and several biologic agents. As a result, the induction and maintenance of remission in patients with rheumatoid arthritis (RA) has risen considerably. In Japan, seven biologic agents are approved for the treatment of patients with RA. The concept of 'treat-to-target' (T2T)[1] has also been accepted by the majority of rheumatologists and clinical professionals. However, the costs of biologic agents are high, and not all patients can afford to be treated. Furthermore, patients with RA are not equally responsive to individual biologic drugs, perhaps owing to the heterogeneity of the disease.

RA is a highly heterogeneous disease with outcomes that are difficult to predict. The extent of joint damage in patients with RA varies considerably among individuals. Patients with more rapid progression seem to need more extensive therapy, such as early treatment with biologic agents, than patients with slower progressing disease. The heterogeneity of RA can be explained, at least in part, by genetic factors; however, several observations indicate that single genetic factors are insufficient to predict clinical phenotype. Thus, the specific combination of genetic factors in an individual might determine the outcome of the disease.

Another important clinical phenotype that could be predicted based on genetic information is an individual's response to therapy. The development of biologic agents such as TNF inhibitors has revolutionized the treatment of RA, but a substantial proportion of patients will not respond to these drugs. In addition, adverse effects to aggressive treatment in patients from Asian ancestry seem to differ to those reported in patients of white European ancestry. For example, the incident ratio of interstitial pneumonia and pneumocystis pneumonia are especially high in Japanese patients with RA.[2] Thus, several unmet needs in clinical practice of rheumatology in Asia remain, with multiple factors potentially affecting this ethnic heterogeneity. To elucidate these differences, genetic and environmental factors should be examined in detail. Furthermore, better knowledge of the interface between genetic and environmental factors could lead to improvements in the understanding of the pathogenesis of RA and other autoimmune diseases, and push the field of rheumatology forwards. In this Perspectives article, we discuss recent advances from genome-wide association studies (GWAS) in patients with RA in Asia, and explore open issues and future research avenues that will help our understanding of the disease.

Competing interests
The authors declare no competing interests.

Genetic studies in Asia

Large-scale GWAS have led to a considerable increase in the availability of genetic information from patients with RA, and have linked common single-nucleotide polymorphisms (SNPs, minor allele frequency >5%) with the risk of developing RA. Several genetic risk factors for this disease have been identified across multiple ethnic groups, including in Asian populations. In the National Human Genome Research Institute catalogue of published GWAS,[3] six out of the 25 GWAS pertaining to RA were of East Asian ethnicities, mainly of Japanese and Korean populations. Results from several GWAS of Asian populations are summarized in Supplementary Table 1 online. Although most associations identified in these studies were later replicated in independent sample sets from the same or distinct populations, some were not. This failure to replicate some of the findings can probably be explained by the tendency of statistically significant findings, especially in GWAS, to overestimate the magnitude of the effects (or the relative risk of disease) and by the requirement for greater statistical power. In some cases, the association of genetic variants with risk of developing RA was specific to the original population but not transposable to other populations, a clear sign of the genetic heterogeneity of the disease. A comprehensive survey of GWAS performed across 28 diseases showed that the replicability rate in East Asian populations was 45.8% (113/225 replication attempts were successful [$P<0.05$]), which was considerably lower than that of European populations (85.6%).[4] Studies of East Asian populations currently tend to have smaller sample sizes and, thus, lower power than those of European populations;[4] the heterogeneity of East Asian populations can also be an important factor contributing to this difference.

Europe–Asia divide

Strong examples of genetic heterogeneity between Asian and European patients

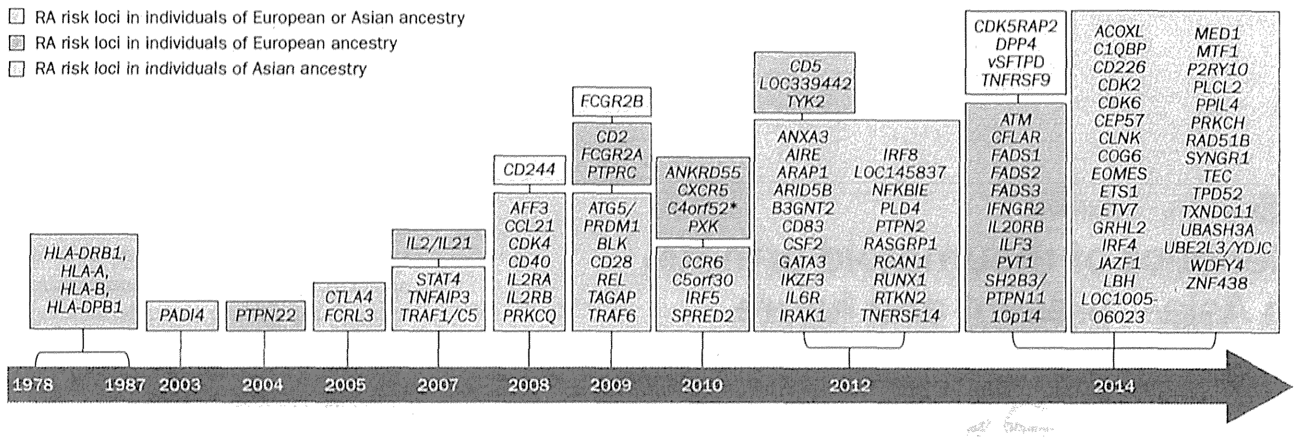

Figure 1 | Historical overview of disease susceptibility polymorphisms identified in RA. Risk loci selected with a genome-wide significance threshold of $P < 5.0 \times 10^{-8}$ in the previous studies are included. *Also known as *SMIM20*. Abbreviations: RA, rheumatoid arthritis.

with RA include the *PADI4* and *PTPN22* genes. In 2003, we reported the first RA-associated SNPs outside the *HLA* gene locus, in *PADI4* (by high-throughput SNP genotyping), which contributed to the risk of RA in Japanese populations.[5] Even though the same *PADI4* polymorphisms were also identified in subsequent studies of Asian populations, these results differ from those observed in European populations.[6] Nevertheless, meta-analyses have confirmed the association of *PADI4* variants in European ancestries, although the effect size of the variants identified was smaller than that in Asian populations.[7] Owing to the epistatic interaction between *PADI4* variants and cigarette smoking, particularly in male patients, the difference in the effect size of *PADI4* variants on the risk of developing RA might be partially attributed to different smoking rates among populations.[8]

Thus, *PADI4* variants are a good example of genetic heterogeneity in which population-specific environmental factors, such as tobacco smoking, can influence the effect size of particular gene variants. Conversely, the Arg620Trp SNP (rs2476601) of *PTPN22*, one of the most studied RA-associated polymorphisms to date, has a relatively large effect size (OR = 1.81) but the SNP is not polymorphic in Asian populations.[9] Given that the allele frequency of this polymorphism varies even among European populations (being higher in northern countries), Arg620Trp might be under strong natural selective pressure from environmental factors. A study of SLE genetics has revealed signs of positive selection in several SLE risk loci—including the *PTPN22* locus—as a possible mechanism underlying the genetic heterogeneity of RA.[10]

Multiethnic risk variants

Although *PADI4* and *PTPN22* variants indicate the presence of genetic heterogeneity in RA, many other common risk variants (including those identified in *CD40*, *TNFAIP3* and *CCR6*) are shared between multiethnic populations.[11] Simulations and empirical analysis of published GWAS data revealed that the unexplained genetic heritability of common diseases could be clarified partially by the accumulation of common alleles with relatively small effect sizes,[12] and by how these polygenic effects are shared among different populations.[11,13] A strong correlation was also observed between the odds ratios for SNPs identified in European GWAS and those in the largest East Asian study performed to date.[11] Overall, the identification of a large number of susceptibility genes and variants is important for understanding the pathogenesis of RA.

Transethnic GWAS

Motivated by the findings described previously, we conducted a transethnic GWAS meta-analysis for RA through international collaboration partnerships involving >20 study cohorts in both European and Asian populations.[14] More than 100,000 case–control participants were incorporated, and the associations of ~10 million autosomal or X-chromosomal SNPs were assessed. Our study identified 42 novel RA risk loci (defined by a genome-wide significance threshold of $P < 5.0 \times 10^{-8}$), increasing the total number of RA risk loci to 101 (Figure 1). We found that >80% of the heritability attributed to these 101 loci was shared among ethnicities, supporting the hypothesis that the genetic risk of RA is, in general, common among distinct populations.[6] This study was one of the first to involve multiple populations in the discovery stage of the GWAS meta-analysis, and also demonstrated the value of transethnic studies for the identification of novel disease risk loci. Furthermore, the detailed transethnic comparison of risk signals in each loci revealed that ethnically diverse linkage disequilibrium structures can help fine-tune the mapping of causal alleles of risk loci.

Given the successes of GWAS meta-analyses in identifying new loci, we hypothesized that a systematic approach to integrate GWAS results using multiple biological and clinical data sources will allow novel insights into disease biology and clues to new therapeutic developments, including drug discovery.[6] We systematically evaluated the overlap between RA risk genes (as well as genes indicated by protein–protein interactions) and genes targeted by approved RA treatments; we observed a considerable number of network connections (Figure 2), which empirically suggests disease-associated genes to be promising resources for drug development.[6] These results highlight the potential of human genetics research to not only identify novel disease susceptibility loci, but also to contribute to the discovery of novel drugs.[6,14]

RA genetics—perspectives

HLA loci

One important assignment for RA genetic studies in Asian populations is the functional dissection of the role of the *HLA-DRB1*0901* allele. The association of *HLA-DRB1*0901* with the risk of developing RA is moderate, but has been reported repeatedly in Asian populations.[15–18] Interestingly, this allele does not contain the classical 'shared epitope' observed in other risk alleles such as *HLA-DRB1*0401* and *HLA-DRB1*0405*. Raychaudhuri and colleagues have established a sophisticated model that could explain the association of *HLA-DRB1* with RA by using amino-acid sequences at locations 11 and 13, in addition to the classical shared epitope at locations 71 and 74.[19] This model could also explain the association of *HLA-DRB1*0901* with the risk of developing RA in Asian populations by adding the amino-acid sequence at location 57 to the model.[20]

Interestingly, although our previous study showed that the *HLA DRB1*0901* allele, similarly to classical shared epitope alleles, was associated with the presence of anti-citrullinated protein antibodies (ACPA), no obvious evidence of an epistatic interaction between *HLA-DRB1*0901* and cigarette smoking was detected, contrary to data regarding shared epitope alleles.[21] Moreover, *HLA-DRB1*0901* has been negatively associated with ACPA titres.[22] *HLA-DRB1*0901* shares the glutamine residue at position 74 with non-shared epitope alleles such as *HLA-DRB1*0403* and *HLA-DRB1*0406*, whereas all shared epitope alleles have an alanine at this amino acid position. This residue, in combination with the residues at position 13 and 71, is important for the formation of the P4 pocket and the binding of citrullinated peptide on the MHC receptor, as shown by crystal structure analysis.[23] Changes in residue 74 might affect the avidity of *HLA-DRB1*0901* to citrullinated peptides—indeed, a peptide-binding study has demonstrated low binding avidity of *HLA-DRB1*0901* to a citrullinated vimentin peptide.[24] These lines of evidence suggest that *HLA-DRB1*0901* (or other genetic factors on the *HLA-DRB1*0901* haplotype) is involved in autoimmunity to citrullinated proteins with a mechanism differing from that of classical shared epitope alleles. Therefore, further functional analyses at the molecular level are needed to dissect the role of the *HLA-DRB1*0901* allele in RA.

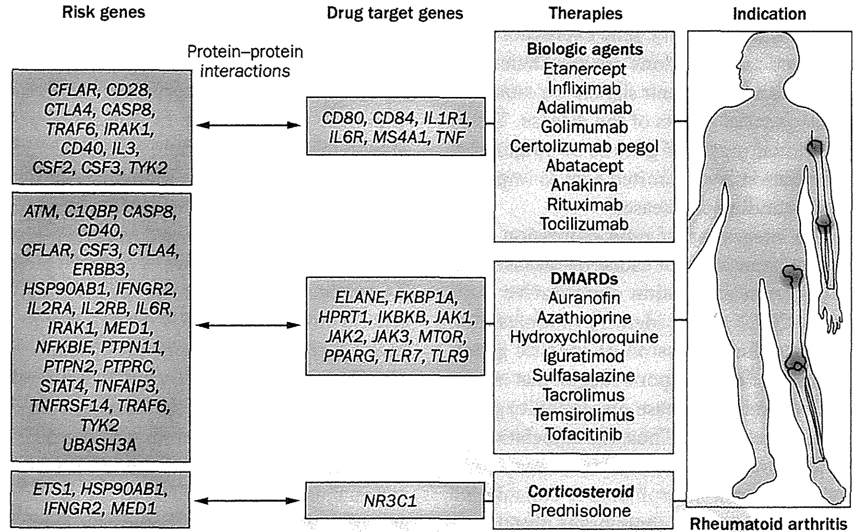

Figure 2 | Association between risk genes and approved RA drugs. Risk genes connect target genes of drugs approved for treatment of RA via protein–protein interactions. Abbreviation: RA, rheumatoid arthritis.

Non-*HLA* loci

Regarding non-*HLA* loci, although the multiethnic GWAS meta-analysis mentioned previously identified an enormous number of disease risk loci with sufficient statistical power, the results of a recent GWAS performed in Han Chinese patients with RA[25] suggest that additional GWAS for specific populations could still be fruitful. This study identified two genetic loci associated with risk of developing RA—*DPP4* at 2q24.3 and *CDK5RAP2* at 9q33.2—neither of which is among the 101 loci identified by the transethnic GWAS meta-analysis. Although the reason for the specificity of the two loci to this population is not clear and needs further investigation, several potential explanations exist. The risk variants of complex traits, including autoimmune diseases, are often under positive natural selection or balancing selection, suggesting they are beneficial for the survival of an individual in their specific environment.[26] According to this view, the risk variants in *DPP4* and *CDKRAP2* could have been positively selected in the Han Chinese population. Similar associations with the risk of RA were reported[11,27] for the 10q21 locus in a Japanese population: three independent associations were detected within 10q21 in the *ARID5B*, *RTKN2* and *EGR2* genes, but only the signal at the *ARID5B* variant was also detected in European populations.[11] In addition, we have shown evidence that *RTKN2* variants have been positively selected in the Japanese population, which might explain the specificity of the disease association in this population. Another explanation could be the interaction of variants with environmental factors, as mentioned for *PADI4*. As *DPP4* plays an important part in the regulation of type 17 T helper (T_H17) cells,[28] population-specific environmental factors, such as microbial activation, might trigger dysregulation of these cells through effects of *DPP4* variants.

Rare variants

Compared with single-ethnicity studies, genetic analysis of multiethnic populations could improve the discovery of rare variants. Given that common gene variants can explain only approximately half of the heritability of RA (as estimated by GWAS data[29]), the remaining heritability might be attributed to rare variants with signals beyond the sensitivity of conventional GWAS. The revolutionary advance in next-generation sequencing technologies has enabled the resequencing of individual genomes, as well as the analysis of rare variants of diseases. However, the analysis of rare variants is sensitive to population stratification, which can substantially bias the results of case–control association tests.[30] Therefore, rare variant analysis should be performed carefully, ideally in genetically homogenous populations with lower admixture, for example in island countries such as Japan. Moreover, given that disease-related rare variants might be unique to a specific population, studies in multiple populations are needed to fully clarify the impact of rare variants in RA.

Beyond genetics

One of the ultimate goals for genetic studies of human diseases is the understanding of pathogeneses. Importantly, genotype is established before disease onset, which supports the concept that genotype leads to phenotype (but not *vice versa*). Thus, compiled information on RA genetics could support approaches to personalized medicine in the future. Nevertheless, few examples of disease susceptibility polymorphisms have been identified which are correlated with clinical outcomes for autoimmune diseases. Additionally, gene expression and epigenetic studies alone are insufficient to conclude whether their signals are causal or consequential events of the disease. Thus, individual pieces of genetic or epigenetic information are of limited help to improve understanding of diseases.

The integration of gene expression data and epigenetics with cause-of-disease can profit from expression quantitative trait loci (eQTL), which identify modules that regulate the expression of specific genes (Figure 3). Several reports suggest that nearly half of common disease susceptibility genes are within eQTL.[31] Thus, the association of gene expression with disease susceptibility genetic polymorphisms is considered a causal event that contributes directly to pathogenesis, providing important information for new therapeutic targets and personalized medicines. For example, Lee *et al.*[32] reported that haplotypes containing the minor allele of a SNP in *FOXO3*, encoding a transcription factor, are associated with elevated transcription of *FOXO3* in monocytes after lipopolysaccharide stimulation. Also in monocytes, this allele was associated with lower production of TNF, a proinflammatory cytokine, and up-regulation of IL-10, an anti-inflammatory cytokine. In keeping with the data on cytokine expression, this minor allele was also associated with a milder course of RA and decreased joint damage, being also correlated with increased susceptibility to severe malaria infection. This is a clear example of one (out of many) polymorphism in a disease susceptibility locus which is truly correlated with disease outcome.

Epigenetic modifications are often important factors when considering associations of gene expression with disease susceptibility genetic polymorphisms involved in disease pathogenesis. In this case, epigenetic modifications could be considered as the cause—not the consequence—of the disease, a process that can reveal potential drug targets and personalized treatments. In fact, SNPs associated with risk of developing RA overlap considerably with peaks of epigenetic modification (H3K4me3) in regulatory T cells.[6,33] Epigenetic modifications are strongly influenced by environmental stimuli. As discussed previously, transethnic polymorphisms associated with diseases, as well as ethnic-specific SNPs affected by regional environments, add to the usefulness of epigenetic research. Studying genetic contributions from multiple Asian geographical regions is important in clinical as well as basic rheumatology.

Conclusions

Genetic studies of patients with RA are necessary for our understanding of the disease, and findings in Asian populations have made major contributions to the field. Furthermore, transethnic GWAS are powerful tools that could become even more useful with the inclusion of additional populations, such as patients from Africa and other Asian countries. The combination of genetic findings with gene expression, epigenetics, protein expression and other intermediate biological processes will facilitate our overall understanding of pathogenesis, helping the process of drug discovery and the development of personalized medicine.

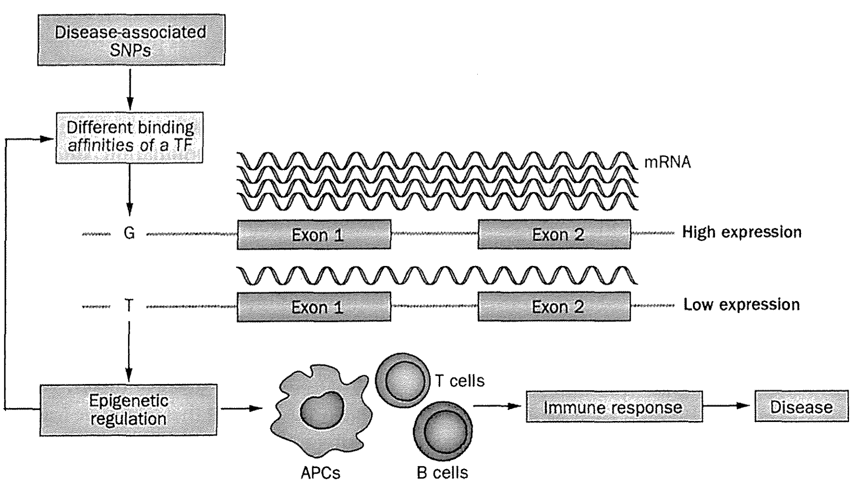

Figure 3 | Disease susceptibility polymorphisms can determine disease pathogenesis. Example of how susceptibility polymorphisms can determine gene expression, epigenetics, immune responses and pathogenesis in rheumatoid arthritis. Abbreviations: APC, antigen-presenting cell; SNP, single nucleotide polymorphism; TF, transcription factor.

*Department of Allergy and Rheumatology, Graduate School of Medicine, the University of Tokyo 113-0033, Japan (**K.Y.**); Laboratory for Autoimmune Diseases, Center for Integrated Medical Sciences, RIKEN, Yokohama 230-0045, Japan (**K.Y., A.S., Y.K.**); Department of Human Genetics and Disease Diversity, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo 113-8510, Japan (**Y.O.**).*
Correspondence to: K.Y.
yamamoto-tky@umin.ac.jp

1. Smolen, J. S. *et al.* Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann. Rheum. Dis.* **69**, 631–637 (2010).
2. Cavagna, L. *et al.* The multifaceted aspects of interstitial lung disease in rheumatoid arthritis. *Biomed. Res. Int.* **2013**, 759760 (2013).
3. National Institutes of Health. *A catalog of published genome-wide association studies* [online], http://www.genome.gov/gwastudies (2014).
4. Marigorta, U. M. & Navarro, A. High trans-ethnic replicability of GWAS results implies common causal variants. *PLoS Genet.* **9**, e1003566 (2013).
5. Suzuki, A. *et al.* Functional haplotypes of *PADI4*, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat. Genet.* **34**, 365–402 (2003).
6. Okada, Y. *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**, 376–381 (2014).
7. Stahl, E. A. *et al.* Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat. Genet.* **42**, 508–514 (2010).
8. Kochi, Y. *et al. PADI4* polymorphism predisposes male smokers to rheumatoid arthritis. *Ann. Rheum. Dis.* **70**, 512–515 (2011).
9. Mori, M., Yamada, R., Kobayashi, K., Kawaida, R. & Yamamoto, K. Ethnic differences in allele frequency of autoimmune-disease-associated SNPs. *J. Hum. Genet.* **50**, 264–266 (2005).
10. Ramos, P. S., Shaftman, S. R., Ward, R. C. & Langefeld, C. D. Genes associated with SLE are targets of recent positive selection. *Autoimmune Dis.* **2014**, 203435 (2014).
11. Okada, Y. *et al.* Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. *Nat. Genet.* **44**, 511–516 (2012).
12. Stahl, E. A. *et al.* Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat. Genet.* **44**, 483–489 (2012).
13. Kurreeman, F. A. *et al.* Use of a multiethnic approach to identify rheumatoid- arthritis-susceptibility loci, 1p36 and 17q12. *Am. J. Hum. Genet.* **90**, 524–532 (2012).
14. Plenge, R. M., Scolnick, E. M. & Altshuler, D. Validating therapeutic targets through human genetics. *Nat. Rev. Drug Discov.* **12**, 581–594 (2013).
15. Wakitani, S. *et al.* The homozygote of HLA-DRB1*0901, not its heterozygote, is associated with rheumatoid arthritis in Japanese. *Scand. J. Rheumatol.* **27**, 381–382 (1998).
16. Lee, H. S. *et al.* Increased susceptibility to rheumatoid arthritis in Koreans heterozygous for HLA-DRB1*0405 and *0901. *Arthritis Rheum.* **50**, 3468–3475 (2004).
17. Kong, K. F., Yeap, S. S., Chow, S. K. &

Phipps, M. E. HLA-DRB1 genes and susceptibility to rheumatoid arthritis in three ethnic groups from Malaysia. *Autoimmunity* **35**, 235–239 (2002).
18. Kochi, Y. *et al.* Analysis of single-nucleotide polymorphisms in Japanese rheumatoid arthritis patients shows additional susceptibility markers besides the classic shared epitope susceptibility sequences. *Arthritis Rheum.* **50**, 63–71 (2004).
19. Raychaudhuri, S. *et al.* Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat. Genet.* **44**, 291–296 (2012).
20. Okada, Y. *et al.* Risk for ACPA-positive rheumatoid arthritis is driven by shared HLA amino acid polymorphisms in Asian and European populations. *Hum. Mol. Genet.* **23**, 6916–6926 (2014).
21. Shimane, K. *et al.* An association analysis of *HLA-DRB1* with systemic lupus erythematosus and rheumatoid arthritis in a Japanese population: effects of *09:01 allele on disease phenotypes. *Rheumatology (Oxford)* **52**, 1172–1182 (2013).
22. Okada, Y. *et al.* HLA-DRB1*0901 lowers anti-cyclic citrullinated peptide antibody levels in Japanese patients with rheumatoid arthritis. *Ann. Rheum. Dis.* **69**, 1569–1570 (2010).
23. Scally, S. W. *et al.* A molecular basis for the association of the *HLA-DRB1* locus, citrullination, and rheumatoid arthritis. *J. Exp. Med.* **210**, 2569–2582 (2013).
24. Freed, B. M., Schuyler, R. P. & Aubrey, M. T. Association of the HLA-DRB1 epitope LA(67, 74) with rheumatoid arthritis and citrullinated vimentin binding. *Arthritis Rheum.* **63**, 3733–3739 (2011).
25. Jiang, L. *et al.* Novel risk loci for rheumatoid arthritis in Han Chinese and congruence with risk variants in Europeans. *Arthritis Rheumatol.* **66**, 1121–1132 (2014).
26. Barreiro, L. B. & Quintana-Murci, L. From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nat. Rev. Genet.* **11**, 17–30 (2010).
27. Myouzen, K. *et al.* Functional variants in *NFKBIE* and *RTKN2* involved in activation of the NF-κB pathway are associated with rheumatoid arthritis in Japanese. *PLoS Genet.* **8**, e1002949 (2012).
28. Bengsch, B. *et al.* Human T_H17 cells express high levels of enzymatically active dipeptidylpeptidase IV (CD26). *J. Immunol.* **188**, 5438–5447 (2012).
29. Stahl, E. A. *et al.* Bayesian inference analyses of the polygenicarchitecture of rheumatoid arthritis. *Nat. Genet.* **44**, 483–489 (2012).
30. Jiang, Y., Epstein, M. P. & Conneely, K. N. Assessing the impact of population stratification on association studies of rare variation. *Hum. Hered.* **76**, 28–35 (2013).
31. Dubois, P. C. et *al.* Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet.* **42**, 295–302 (2010).
32. Lee, J. C. *et al.* Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway. *Cell* **155**, 57–69 (2013).
33. Trynka, G. *et al.* Chromatin marks identify critical cell-types for fine-mapping complex trait variants. *Nat. Genet.* **45**, 124–130 (2013).

Author contributions
Y.O. contributed substantially to the discussion of content. K.Y., A.S. and Y.K. reviewed and edited the manuscript before submission. All authors wrote the manuscript.

Supplementary information is linked to the online version of the paper at www.nature.com/nrrheum.

ARTHRITIS & RHEUMATOLOGY
Vol. 67, No. 1, January 2015, pp 63–73
DOI 10.1002/art.38895

Activation of Syk in Peripheral Blood B Cells in Patients With Rheumatoid Arthritis

A Potential Target for Abatacept Therapy

Shigeru Iwata, Shingo Nakayamada, Shunsuke Fukuyo, Satoshi Kubo, Naoki Yunoue, Sheau-Pey Wang, Maiko Yoshikawa, Kazuyoshi Saito, and Yoshiya Tanaka

Objective. **B cells play a pivotal role in the pathogenesis of autoimmune diseases. Although Syk functions as a key molecule in B cell receptor signaling, the pathologic role of Syk in B cells in rheumatoid arthritis (RA) remains unclear. The purpose of this study was to assess the relevance of activation of Syk in B cells to the pathologic development of RA and to the responsiveness of RA patients to treatment with biologics.**

Methods. **Healthy subjects (n = 36) and patients with moderate or severe RA disease activity (n = 70) were studied. The phosphorylation of Syk (pSyk) in peripheral blood B cells was measured by flow cytometry, and its correlation with clinical characteristics and changes after administration of biologic agents was evaluated.**

Results. **Levels of pSyk in peripheral blood B cells were preferentially higher in patients with RA compared to healthy subjects. Patients with significantly higher pSyk levels were strongly positive for anti–citrullinated protein antibodies (ACPAs). High pSyk levels were not correlated with the severity of disease activity. Treatment with abatacept, but not tumor necrosis factor inhibitors, significantly reduced the levels of pSyk in RA peripheral blood B cells. Abatacept also significantly reduced the proportion of follicular helper T (Tfh) cells.**

Conclusion. **Levels of pSyk in peripheral blood B cells were significantly elevated in patients with RA, and these patients also exhibited strong positivity for ACPAs. These data suggest that abatacept seems to inhibit the phosphorylation of Syk in B cells, as well as the development of Tfh cells, thus highlighting the relevance of B cell–T cell interactions as a potential target of abatacept therapy in RA.**

Activated autoreactive B cells produce autoantibodies and inflammatory cytokines such as interleukin-6 (IL-6) and tumor necrosis factor α (TNFα). The expression of costimulatory molecules, such as CD40 and CD80, is enhanced on B cells and is involved in the interactive activation with surrounding immunocompetent cells, including T cells. B cells have an antigen-presenting activity, particularly in autoimmune diseases, and are associated with the activation of autoreactive T cells. Therefore, B cells play an important role in the pathogenetic processes of rheumatoid arthritis (RA).

Rituximab, a chimeric anti-CD20 antibody, eliminates B cells through antibody- and complement-dependent cytotoxic activities. The efficacy of rituximab

Supported in part by a Research Grant on Rare and Intractable diseases and a Research Grant-In-Aid for Scientific Research from the Ministry of Health, Labor, and Welfare of Japan, the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and the University of Occupational and Environmental Health Japan and by a Grant for Advanced Research from the University of Occupational and Environmental Health Japan.

Shigeru Iwata, MD, PhD, Shingo Nakayamada, MD, PhD, Shunsuke Fukuyo, MD, PhD, Satoshi Kubo, MD, PhD, Naoki Yunoue, MD, Sheau-Pey Wang, MS, Maiko Yoshikawa, MD, Kazuyoshi Saito, MD, PhD, Yoshiya Tanaka, MD, PhD: University of Occupational and Environmental Health Japan, Kitakyushu, Japan.

Dr. Tanaka has received consulting fees, speaking fees, and/or honoraria from Astellas, Takeda, Santen, Mitsubishi-Tanabe, Pfizer, Janssen, Eisai, Daiichi-Sankyo, UCB, GlaxoSmithKline, Bristol-Myers Squibb, Abbott Japan, Astra-Zeneca, Eli Lilly Japan, Quintiles, MSD, and Asahi-Kasei (less than $10,000 each) and from AbbVie and Chugai (more than $10,000 each) and has received research grants from Bristol-Myers Squibb, Mitsubishi-Tanabe, AbbVie, MSD, Chugai, Astellas, and Daiichi-Sankyo.

Address correspondence to Yoshiya Tanaka, MD, PhD, First Department of Internal Medicine, School of Medicine, University of Occupational and Environmental Health Japan, 1-1 Iseigaoka, Yahata-nishi, Kitakyushu 807-8555, Japan. E-mail: tanaka@med.uoeh-u.ac.jp.

Submitted for publication November 28, 2013; accepted in revised form September 23, 2014.

has been demonstrated in RA patients with high disease activity (in the Dose-Ranging Assessment: International Clinical Evaluation of Rituximab in Rheumatoid Arthritis [DANCER] trial [1]) and in RA patients resistant to TNF inhibitor therapy (in the Randomized Evaluation of Long-term Efficacy of Rituximab in Rheumatoid Arthritis [REFLEX] trial [2]). Rituximab was approved for the treatment of RA in the US in 2006 and is currently considered the second-line biologic agent, subsequent to TNF inhibitor therapy. In addition to these studies, some clinical studies have demonstrated the efficacy of a humanized anti-CD20 antibody, ocrelizumab, and a fully human anti-CD20 antibody, ofatumumab, in patients with RA resistant to TNF inhibitor therapy, indicating that B cells are an evident therapeutic target for RA.

Syk is a 72-kd nonreceptor tyrosine kinase discovered by Taniguchi et al (3) in 1991. Syk is involved in the signaling pathway through Fc receptors, which are broadly expressed on immunocompetent cells, such as B cells, dendritic cells, mast cells, macrophages, and neutrophils, and on molecules associated with cell adhesion, such as integrin (4,5).

Recently, the importance of Syk in the pathologic processes of RA has been reported. The results of a phase II clinical study of R406, a Syk inhibitor, in patients resistant to treatment with methotrexate (MTX) indicated that phosphorylation of Syk (measured as levels of pSyk) was increased in the synovial tissue of RA patients compared to healthy subjects and patients with osteoarthritis (6–8). Another experimental study using the synovial cells from these patients demonstrated that R406 inhibits TNFα-induced activation of mitogen-activated protein kinases and the expression of the matrix metalloproteinase 3 (MMP-3) gene, thus highlighting the significant role of Syk in synovial fibroblasts of RA patients (9).

In addition, previous studies elucidated the role of Syk in B cells. Syk has important roles in B cell maturation and survival (10,11). The Toll-like receptor 9 (TLR-9) signaling pathway is involved in the activation of B cells and autoantibody production by B cells (12,13). In this regard, we have recently demonstrated that signaling through Syk results in effective signal transduction of TLR-9 by inducing optimal expression of TNF receptor–associated factor 6 (TRAF6), and that this signaling is important for antibody production by B cells (14). Based on these results, we hypothesized that Syk phosphorylation in B cells is involved in the pathologic processes of RA through the production of autoantibodies, such as rheumatoid factor (RF) and anti–citrullinated protein antibodies (ACPAs).

T cells (especially Th1 and Th17 cells) also play a pivotal role in the pathogenesis of RA (15,16). Recently, follicular helper T (Tfh) cells, whose primary task is to drive the formation of B cell responses, have been recognized as a critical regulator of autoimmunity (17,18). We and other investigators have elucidated the mechanism of Tfh cell differentiation (19,20); however, the exact role of this T helper cell subset in RA remains elusive.

Abatacept, a fusion protein containing CTLA-4 and Ig, which is referred to as a T cell–selective costimulatory regulator, inhibits the activation of T cells. However, little is known about the T cell populations targeted by abatacept. The effect of abatacept on antigen-presenting cells has also been reported (21–23). The inhibitory effect of abatacept on T cell–dependent antibody production has been reported in mice and cynomolgus monkeys (24,25). Evidence suggests that abatacept also has an inhibitory effect on bone destruction, by suppressing the production of RF and ACPAs (26). However, the effect of abatacept on human B cells is unknown. Based on these observations, abatacept is predicted to regulate the activation of not only T cells but also B cells, directly and/or indirectly.

In this study, we observed significantly elevated Syk phosphorylation in the peripheral blood B cells of patients with RA compared to healthy subjects, and we demonstrated that the levels of pSyk were significantly high in patients who were strongly positive for ACPAs. Moreover, treatment with abatacept, but not with TNF inhibitors, significantly inhibited Syk phosphorylation in B cells. Interestingly, treatment with abatacept significantly reduced the proportion of Tfh cells, which could be a possible mechanism for the reduction in Syk phosphorylation in B cells. The results suggest that Syk plays an important role in ACPA production by B cells in patients with RA, and that abatacept inhibits both Syk phosphorylation in B cells and the development of Tfh cells.

PATIENTS AND METHODS

Patients. Table 1 summarizes the baseline characteristics of the 70 patients with RA. The healthy control subjects (n = 36) were either staff members of our hospital or healthy subjects who visited our hospital for medical examinations. Patients with RA who were resistant to treatment comprised those whose score of RA disease activity was >3.1 on the Disease Activity Score in 28 joints using erythrocyte sedimentation rate (DAS28-ESR) (27), despite having received treat-

Table 1. Characteristics of the study patients with rheumatoid arthritis (n = 70)*

Age, mean ± SD years	61.4 ± 15.1
Sex, no. female/no. male	60/10
Disease duration, mean ± SD months	91.5 ± 114.4
Prednisolone (or equivalent)	
No. not receiving treatment/total no.	11/70
Dosage, mean ± SD mg/day	3.4 ± 1.9
Methotrexate	
No. not receiving treatment/total no.	53/70
Dosage, mean ± SD mg/week	13.0 ± 3.6
Tender joint count, mean ± SD	8.5 ± 7.3
Swollen joint count, mean ± SD	7.3 ± 6.3
CRP, mean ± SD mg/dl	2.0 ± 3.0
ESR, mean ± SD mm/hour	53.2 ± 33.3
IgG, mean ± SD mg/dl	1,512.5 ± 452.5
RF	
Mean ± SD IU/ml	149.7 ± 407.7
No. negative/no. positive	21/49
ACPA status, no.	
Negative	22
Positive	6
Strongly positive	42
MMP-3, mean ± SD ng/ml	194.8 ± 246.7
DAS28-CRP, mean ± SD	4.7 ± 1.4
DAS28-ESR, mean ± SD	5.5 ± 1.4
CDAI, mean ± SD	26.3 ± 15.0
SDAI, mean ± SD	28.3 ± 16.8
HAQ score, mean ± SD	1.3 ± 0.9
No. not treated with biologics/total no.	57/70

* CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; RF = rheumatoid factor; ACPA = anti-citrullinated protein antibody; MMP-3 = matrix metalloproteinase 3; DAS28-CRP = Disease Activity Score in 28 joints using CRP level; CDAI = Clinical Disease Activity Index; SDAI = Simplified Disease Activity Index; HAQ = Health Assessment Questionnaire.

ment with adequate doses of antirheumatic drugs, mainly MTX, for a minimum of 3 months, and who showed no response or only a moderate response to treatment according to the European League Against Rheumatism (EULAR) improvement criteria (28). The Human Ethics Review Committee of the university reviewed and approved our study, including the collection of peripheral blood samples from healthy adults and patients with RA. Each subject provided a signed participation consent form.

Measurements. The background factors investigated were sex, age, duration of RA, and doses of corticosteroids and MTX. We also evaluated the severity of morning stiffness, number of swollen joints, number of tender joints, and patient's evaluations of pain and overall health by visual analog scales, in addition to global evaluations of health by the attending physician. The laboratory tests included measurements of the C-reactive protein (CRP) level, ESR, IgG, RF, ACPAs, and MMP-3. We consulted the American College of Rheumatology (ACR)/EULAR 2010 classification criteria for RA (29) to select the cutoff values for stratification of ACPA positivity. Low-positive ACPA refers to IU values that are higher than the upper limit of normal (ULN) but ≤3 times the ULN for the laboratory and assay, whereas high-positive ACPA refers to IU values that are >3 times the ULN for the laboratory and assay. The variables investigated included the DAS28 using CRP level (DAS28-CRP), DAS28-ESR, the Clinical Disease Activity Index (CDAI) (30), the Simplified Disease Activity Index (SDAI) (31), the Health Assessment Questionnaire (HAQ) (32), and history of biologics use.

Flow cytometry analysis. Peripheral blood mononuclear cells (PBMCs) from 36 normal healthy volunteers and from 70 patients with RA whose diagnosis met the ACR 1987 revised classification criteria for RA (33) were isolated from the peripheral blood using lymphocyte separation medium (ICN/Cappel Pharmaceuticals). For surface and intracellular staining, 2×10^5 PBMCs, which were acquired after strict deletion of dust by threshold adjustment, were subjected to fluorescence-activated cell sorting analysis. PBMCs were fixed with phosphate buffered saline (PBS) containing 1% formaldehyde and then permeabilized with PBS containing 0.1% saponin. After washing, the PBMCs were resuspended in saponin–PBS and stained with mouse anti-human Syk monoclonal antibodies (mAb) (Abcam) and mouse anti-human pSyk (pY348) mAb (BD PharMingen), followed by washing with saponin–PBS. Phycoerythrin-labeled goat anti-mouse IgG polyclonal antibody (BD PharMingen) was used as a secondary antibody. After washing with saponin–PBS, the PBMCs were stained with fluorescein isothiocyanate–labeled mouse anti-human CD19 antibodies (BD PharMingen).

The rate of pSyk expression in B cells was calculated as the percentage of pSyk-positive CD19+ B cells relative to total CD19+ B cells. We defined pSyk-positive CD19+ B cells as cells in which the intensity of staining was higher than the background staining with IgG control antibody. The proportion of CD19+ B cells (relative to total cells) in healthy donors and RA patients was a mean ± SD 15,199 ± 7,482 cells (7.6 ± 3.7%) and 12,844 ± 7,120 cells (6.6 ± 3.6%), respectively.

Tfh cells were stained with anti-CD4, anti-CXCR5, and anti–programmed death 1 (anti–PD-1) antibodies (BD PharMingen). The proportion of CD4+ cells (relative to total cells) was 20,364 ± 17,727 cells (8.2 ± 7.0%), while that of CD4+CXCR5+PD-1+ cells (relative to total cells) was 1,841 ± 3,940 cells (0.7 ± 1.5%). Stained cells were analyzed on a flow cytometer (FACSCalibur; BD PharMingen). The cells were collected and analyzed with FlowJo software (Tree Star).

In vitro B cell activation analysis. CD19+ B cells were purified from the peripheral blood of the healthy control subjects and RA patients. The cells were cultured in stimulation-free medium for 3 days to assess the production of IL-6 or for 5 days to assess the production of IgG. IL-6 production was determined using a BD Cytometric Bead Array human Flex set (BD PharMingen). Flow cytometry was carried out using a FACSCalibur and CellQuest software (Becton Dickinson). IgG levels in the culture medium were determined using a human IgG enzyme-linked immunosorbent assay quantitation kit (Bethyl Laboratories).

Statistical analysis. Data are expressed as the mean ± SD. Differences between groups for variables with normal distribution and homoscedasticity were compared using Student's *t*-test. Differences between groups for variables with skewed distribution were compared using Wilcoxon's rank sum test. Analysis of variance followed by the Bonferroni/Dunn post hoc test was used to compare data from 3 groups with normal distribution. The Kruskal-Wallis test followed by the Bonferroni/Dunn post hoc test was used to compare data from

>3 groups with skewed distribution. Correlation analysis was performed using Spearman's correlation coefficients. Baseline and posttreatment values within each sample were compared using Wilcoxon's matched-pairs signed-rank test. P values less than 0.05 were considered significant. All analyses were conducted using PASW Statistics software version 18.0 (IBM).

RESULTS

Patient background. This study was conducted in 70 patients with RA who were receiving treatment in our hospital in Japan. The clinical features of the RA patients are described in Table 1. The washout period in patients who had previously received biologics (etanercept, golimumab, adalimumab, tocilizumab, abatacept) was more than 1 month. Infliximab required a 60-day washout.

High Syk phosphorylation in B cells of ACPA-positive RA patients. PBMCs were isolated from 36 healthy donors (as controls) and 70 patients with RA

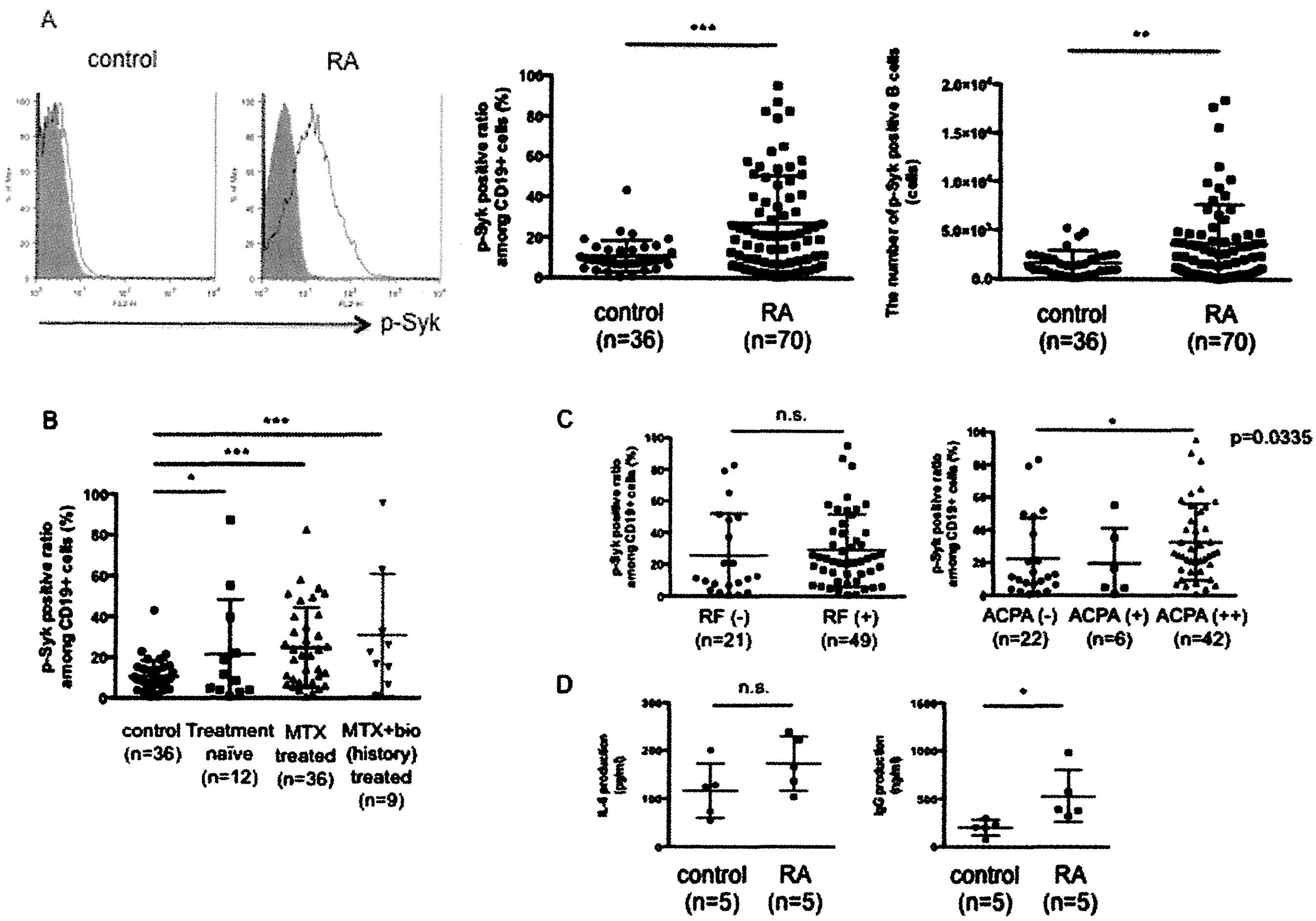

Figure 1. Phosphorylation of Syk in CD19+ B cells of healthy donors (controls) and patients with rheumatoid arthritis (RA). **A,** Representative histograms showing Syk phosphorylation in peripheral blood B cells from 70 RA patients and 36 healthy control subjects (left), and the ratio of pSyk-positive cells among CD19+ B cells (middle) and absolute number of pSyk-positive CD19+ B cells (right) in RA patients compared to healthy controls. **B,** Ratio of pSyk-positive cells among CD19+ B cells in 3 groups of RA patients: treatment-naive (n = 12), methotrexate (MTX)–treated (n = 36), and MTX + biologics (bio) (history)–treated (n = 9). RA patients treated with other disease-modifying antirheumatic drugs and/or corticosteroids were excluded. **C,** Ratio of pSyk-positive cells among CD19+ B cells in RA patients negative for rheumatoid factor (RF) (defined as <15 IU/ml, based on the normal limit at our hospital) or positive for RF (defined as ≥15 IU/ml), and RA patients negative (−), positive (+), or strongly positive (++) for anti–citrullinated protein antibodies (ACPAs) (defined as <4.5 units/ml, 4.5–13.5 units ml, and >13.5 units/ml, respectively, based on the normal limit at our hospital). **D,** Production of interleukin-6 (IL-6) (left) and IgG (right) by CD19+ B cells purified from the peripheral blood of healthy controls and RA patients. B cells were cultured in stimulus-free RPMI medium for 3 days (for IL-6) or 5 days (for IgG). Production of IL-6 in the supernatants was assayed by cytometric bead array, while IgG in the supernatants was quantified by enzyme-linked immunosorbent assay. Symbols represent individual subjects; bars show the mean ± SD. * = $P < 0.05$, ** = $P < 0.01$; *** = $P < 0.001$. NS = not significant.