Identification of two autoantigens recognised by circulating autoantibodies as potential biomarkers for diagnosing giant cell arteritis

E. Pesce[1,2], M. Bombaci[2], S. Croci[3], M. Bonacini[3], C. Marvisi[4], C. Ricordi[5], S. Monti[6], F. Muratore[4,5], S. Abrignani[1,2], R. Caporali[1,7], M.O. Borghi[1,6], C. Salvarani[4,5], P.M. Villiger[8], R. Grifantini[2], P.L. Meroni[1]

[1]*Dipartimento di Scienze Cliniche e di Comunità, Dipartimento di Eccellenza 2023-2027, University of Milan;*
[2]*Istituto Nazionale Genetica Molecolare, Milan;*
[3]*Unit of Clinical Immunology, Allergy and Advanced Biotechnologies, Azienda Unità Sanitaria Locale-IRCCS di Reggio Emilia, Reggio Emilia;*
[4]*Unit of Rheumatology, Azienda Unità Sanitaria Locale-IRCCS di Reggio Emilia, Reggio Emilia;*
[5]*University of Modena and Reggio Emilia, Modena;*
[6]*IRCCS Istituto Auxologico Italiano, Laboratory of Immunorheumatologic Researches, Milan;*
[7]*Department of Rheumatology, ASST PINI-CTO, Milan, Italy;*
[8]*Medical Center Monbijou and University of Bern, Switzerland.*

Supplementary material and methods

Protein array platform: staining and image acquisition

In the discovery phase, two different types of protein array were used to detect and identify autoantibodies in human sera. The first was the ProtoArray v5.0 Human Protein Microarray (PAH05251020, Thermo Fisher Scientific), containing over 9,000 unique antigens. All proteins were expressed as glutathione s-transferase (GST) fusion proteins in insect cells and purified under native conditions. The second was a proprietary human protein array comprising about 1,650 recombinant human proteins predicted to be secreted or membrane-associated, most of which are still poorly characterised. These proteins were expressed as His-tag proteins in *E. coli* (1, 2).

Microarrays were initially blocked using Blocking Buffer (PA017, Thermo Fisher Scientific) for 1 hour and then incubated with sera diluted to 1:500 in washing buffer TPBS (PBS buffer with 0.05% Tween 20) for 1 hour. After washing, the arrays were probed with α-human IgG conjugated to AlexaFluor-647 (A-21445, Thermo Fisher Scientific) at a dilution of 1:1000 for 1 hour at room temperature. Finally, the microarrays were washed, dried, and scanned using the ScanArray Gx PLUS (PerkinElmer). 16-bit images were generated and analyzed using Protoarray Prospector software v5.2.3 (ThermoFisher Scientific) for the ProtoArray v5.0 and ImaGene 8.0 software (Biodiscovery Inc) for the in house developed array.

Protein array data analysis

For each sample, the Mean Fluorescence Intensity (MFI) value was determined by subtracting the background given by the controls used and then averaging the replicates. Subsequently, the value was normalised based on the human IgG curve to allow comparison of data from different experiments. MFI values for each protein were normalised accordingly (3). A normalised MFI value of 5.000, corresponding to the normalised MFI value of negative controls, was chosen as the lowest signal threshold for scoring a protein as positively recognised by human sera. Recognition frequency was defined as the percentage of sera reacting with a particular antigen in the protein array with an MFI ≥12.500, and it was calculated for each group of sera compared to HC sera. The localisation prediction of the selected proteins was conducted using the Deeploc 2.0 predictive algorithm (4).

Expression and purification of candidate autoantigens

Candidate autoantigens were expressed as His-tagged recombinant forms in the *E. coli* BL21 (DE3) system. The recombinant proteins were affinity-purified from the bacterial inclusion bodies by means of Immobilized Metal Affinity Chromatography (IMAC) (GE). Purification of the recombinant protein under denaturing condition was performed using Ni Sepharose 6 Fast Flow (Cytiva) according to the manufacturer's recommendations. Purity and identity of the recombinant protein was confirmed by 4-12% Bolt Bis-Tris (Thermo Fisher Scientific) sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE.

ELISA

96-well microplate (Nunc Maxisorp) were coated with 2.5 μg/ml of each purified recombinant proteins in PBS (pH 7.4) overnight at 4°C. After washing four times with PBS containing 0.05% Tween 20 (PBST), plates were blocked with 10% BSA in PBST for 1h at 37°C, then incubated with 1:100 diluted sera in 1% BSA in PBST for 1h at 37°C. Plates were washed five times with PBST solution and probed with 1:2000 HRP-conjugated α-human IgG secondary antibody (Cytiva) for 45 minutes at RT. After washing, plates were developed by addition 100 μL of 3, 3´,5,5´-tetramethylbenzidine chromogenic (TMB) substrate (Thermo Fisher Scientific) for 10 minutes. The reaction was stopped by adding 100μl 2 M H_2SO_4 after 5-10 min, and the absorbance was measured at λ=450 nm by InfiniteF200 PRO plate reader (Tecan). The cutoff was calculated by adding two standard deviations to the mean optical density (OD) value of ELISA runs on the samples of HC.

Supplementary Table S1. List of 59 autoantigens, selected by protein array analysis, that react specifically with GCA sera. The table shows Gene annotation, MFI values, and recognition frequency for GCA and HC classes. Reactivity was represented with the flag.

GCA pool			Prot Id	Description	Gene Symbol	MFI		Frequency		Ratio
01	**02**	**03**				**HC**	**GCA**	**HC**	**GCA**	
	√		Q8ND87	*nucleosome assembly protein 1-like 5 (NAP1L5)*	NAP1L5	986	19158	0	33	*19.4*
	√		P19957	*peptidase inhibitor 3, skin-derived (SKALP) (PI3)*	PI3	1946	19158	0	33	*9.8*
	√		Q9H3Q1	*CDC42 effector protein (Rho GTPase binding) 4 (CDC42EP4)*	CDC42EP4	3085	19006	0	33	*6.8*
	√		Q9ULJ7	*ankyrin repeat domain 50 (ANKRD50)*	ANKRD50	4927	20927	0	33	*3.2*
√	√		Q8N1L9	*basic leucine zipper transcription factor, ATF-like 2 (BATF2)*	BATF2	2778	15822	0	67	*4.6*
√			Q13432	*unc-119 homolog (C. elegans) (UNC119), transcript variant 1*	UNC119	1783	12875	0	33	*7.6*
	√		Q9BRL4	*PCTAIRE protein kinase 1 (PCTK1), transcript variant 3*	PCTK1	3464	13606	0	33	*5.3*
	√		P17677	*growth associated protein 43 (GAP43)*	GAP43	1236	18314	0	33	*12.9*
	√		Q86XR8	*centrosomal protein 57kDa (CEP57)*	CEP57	2945	15887	0	33	*5.6*
	√		Q9BTY8	*EF-hand domain family, member D2 (EFHD2)*	EFHD2	2365	16579	0	33	*6.1*
	√		Q14498	*RNA binding motif protein 39 (RBM39), transcript variant 3*	RBM39	467	14372	0	33	*41.0*
√	√		Q04206	*v-rel reticuloendotheliosis viral oncogene homolog A, nuclear factor of kappa light polypeptide gene enhancer in B-cells 3, p65 (avian) (RELA)*	RELA	1710	19147	0	67	*7.7*
	√	√	Q96AA8	*janus kinase and microtubule interacting protein 2 (JAKMIP2)*	JAKMIP2	1178	13185	0	67	*17.1*
	√	√	NA	*cDNA clone IMAGE:4829245*	LOC100505540	1678	20193	0	67	*7.9*
	√		P55347	*PBX/knotted 1 homeobox 1 (PKNOX1)*	PKNOX1	946	13256	0	33	*20.7*
√	√		P49137	*MAP kinase-activated protein kinase 2*	MAPKAPK2	1494	19541	0	67	*9.8*
√			Q13432	*unc-119 homolog (C. elegans) (UNC119), transcript variant 2*	UNC119	191	14648	0	33	*67.2*
√			Q96BX8	*MOB1, Mps One Binder kinase activator-like 2A (yeast) (MOBKL2A)*	MOBKL2A	1309	12846	0	33	*11.5*
	√		Q9H4G0	*erythrocyte membrane protein band 4.1-like 1 (EPB41L1), transcript variant 2*	EPB41L1	2607	15019	0	33	*7.2*
	√		Q9H3Q1	*CDC42 effector protein (Rho GTPase binding) 4 (CDC42EP4)*	CDC42EP4	1960	18841	0	33	*8.2*
√			O14645	*dynein, axonemal, light intermediate chain 1 (DNALI1)*	DNALI1	3453	16156	0	33	*5.3*
	√		Q9BZE9	*alveolar soft part sarcoma chromosome region, candidate 1 (ASPSCR1)*	ASPSCR1	1977	18273	0	33	*6.3*
	√		Q9NP73	*asparagine-linked glycosylation 13 homolog (S. cerevisiae) (ALG13)*	ALG13	200	12510	0	33	*71.6*
	√	√	A8MT08	*Leukocyte specific transcript 1 (LST1), transcript variant 2, mRNA*	LST1	3532	14298	0	67	*5.0*
		√	P61604	*heat shock 10kDa protein 1 (chaperonin 10) (HSPE1)*	HSPE1	780	17713	0	33	*20.4*
	√		Q9H668	*oligonucleotide/oligosaccharide-binding fold containing 1 (OBFC1)*	OBFC1	2327	15865	0	33	*5.8*
	√	√	P11766	*alcohol dehydrogenase 5 (class III), chi polypeptide (ADH5)*	ADH5	710	13606	0	67	*21.3*
√	√		Q66K14	*TBC1 domain family, member 9B (with GRAM domain) (TBC1D9B)*	TBC1D9B	119	15122	0	67	*158.9*
	√	√	Q9NZ09	*ubiquitin associated protein 1 (UBAP1)*	UBAP1	2642	18979	0	67	*5.7*
	√	√	O43679	*LIM domain binding 2 (LDB2)*	LDB2	1225	15130	0	67	*15.4*
	√		Q8IW43	*hypothetical protein MGC23270 (MGC23270)*	MGC23270	1344	18873	0	33	*10.0*
	√		O00506	*serine/threonine kinase 25 (STE20 homolog, yeast) (STK25)*	STK25	2521	13382	0	33	*8.1*
	√		P09917	*arachidonate 5-lipoxygenase (ALOX5)*	ALOX5	1106	20452	0	33	*17.0*
√	√		Q9UNP9	*peptidylprolyl isomerase E (cyclophilin E) (PPIE), transcript variant 3*	PPIE	862	18787	0	67	*23.0*
	√		Q8TB22	*spermatogenesis associated 20 (SPATA20)*	SPATA20	1304	19793	0	67	*10.9*
√	√		O75940	*survival motor neuron domain containing 1 (SMNDC1)*	SMNDC1	1750	14174	0	67	*8.1*
	√		Q13094	*lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte protein of 76kDa) (LCP2)*	LCP2	2901	14198	0	33	*5.4*
	√		Q14498	*RNA-binding protein 39*	RBM39	762	15806	0	33	*22.3*
	√		Q9Y3A3	*Mps one binder kinase activator-like 3*	MOBKL3	233	16979	0	33	*84.7*
	√		Q9BY44	*eukaryotic translation initiation factor 2A, 65kDa (EIF2A)*	EIF2A	2300	19773	0	33	*6.3*
	√		O00506	*serine/threonine kinase 25 (STE20 homolog, yeast) (STK25)*	STK25	1483	14423	0	33	*12.0*
	√		P22459	*Potassium voltage-gated channel, shaker-related subfamily, member 4 (KCNA4), mRNA*	KCNA4	609	17823	0	33	*22.3*
	√		Q9BSF8	*BTB (POZ) domain containing 10 (BTBD10)*	BTBD10	610	13576	0	33	*30.0*
	√	√	Q9H7Z7	*prostaglandin E synthase 2 (PTGES2), transcript variant 2*	PTGES2	323	18265	0	67	*72.2*
√		√	Q8N8Z6	*Discoidin, CUB and LCCL domain-containing protein 1 Precursor*	DCBLD1	4705	23290	0	67	*2.8*
√	√	√	Q02985	*Complement factor H-related protein 3 Precursor*	CFHR3	2010	12977	0	100	*8.6*
√		√	Q8N3G9	*Transmembrane protein 130*	TMEM130	2960	17340	0	67	*5.6*
√	√	√	O75951	*Lysozyme-like protein 6 Precursor*	LYZL6	3925	16500	0	100	*4.1*
√		√	Q8N755	*PQ-loop repeat-containing protein 3 Precursor*	PQLC3	3103	16135	0	67	*4.1*
√		√	Q9H6B9	*Abhydrolase domain-containing protein 9 Precursor*	EPHX3	4779	12788	0	67	*2.9*
√	√	√	Q86YB7	*Enoyl-CoA hydratase domain-containing protein 2, mitochondrial Precursor*	ECHDC2	4359	13629	0	100	*3.8*
√		√	A2A2V5	*Serine-rich and transmembrane domain-containing protein 1*	SERTM1	3431	16604	0	67	*3.8*
√		√	Q6ZPD9	*Protein dpy-19 homolog 3*	DPY19L3	4686	12990	0	67	*2.7*
√		√	O95377	*Gap junction beta-5 protein*	GJB5	4716	12688	0	67	*2.7*
		√	Q8N2F6	*Armadillo repeat-containing protein 10*	ARMC10	3930	12593	0	33	*3.3*
√			Q12893	*Transmembrane protein 115*	TMEM115	1991	12852	0	33	*7.0*
√	√	√	Q86VR7	*V-set and immunoglobulin domain containing 10 like*	VSIG10L	4345	13840	0	100	*3.8*
√		√	O15335	*Chondroadherin Precursor*	CHAD	4600	16678	0	67	*3.0*
√		√	Q8TEZ7	*Membrane progestin receptor beta*	PAQR8	2883	13765	0	67	*4.9*

References

1. GRIFANTINI R, PAGANI M, PIERLEONI A *et al*.: A novel polyclonal antibody library for expression profiling of poorly characterized, membrane and secreted human proteins. *J Proteomics* 2011; 75(2):532-47. https://doi.org/ 10.1016/j.jprot.2011.08.018
2. ZINGARETTI C, ARIGÒ M, CARDACI A *et al*.: Identification of new autoantigens by protein array indicates a role for IL4 neutralization in autoimmune hepatitis. *Mol Cell Proteomics* 2012; 11(12):1885-97. https://doi.org/10.1074/mcp.M112.018713
3. BOMBACI M, GRIFANTINI R, MORA M, REGUZZI V, PETRACCA R, MEONI E *et al*.: Protein array profiling of tic patient sera reveals a broad range and enhanced immune response against Group A Streptococcus antigens. *PLoS One* 2009; 4(7): e6332. https://doi.org/10.1371/journal.pone.0006332
4. THUMULURI V, ALMAGRO ARMENTEROS JJ, JOHANSEN AR, NIELSEN H, WINTHER O: DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Res* 2022; 50(W1):W228-W234. https://doi.org/10.1093/nar/gkac278