- running title: Association of HS1,2 alleles with autoimmune diseases.

- Title : Association of autoimmune diseases to allele 2 of the 3' immuno-globulines HS1.2 enhancer bearing an NF-κB binding site.

Frezza Domenico 1, Barbara Tolusso 2, Vincenzo Giambra 3, Claudia Mattioli 1, Elisa Gremese 2, Silvia Bosello 2, Maria DeSantis 2, Barbara K. Birshtein 4, Gianfranco Ferraccioli 2.

1 Department of Biology "Enrico Calef", University of Roma Tor Vergata, Italy.

2 Division of Rheumatology, Catholic University of the Sacred Hearth, Rome, Italy

3 actual address Terry Fox Laboratory, British Columbia Cancer Research Centre, Vancouver, BC, Canada

4 Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY, USA.

ABSTRACT

We identified allele *2 in the HS1.2 enhancer of the Igh 3' regulatory region (3'RR-1) as a risk factor for systemic lupus erythematosus (SLE), rheumatoid arthritis (RA) and systemic sclerosis (SSc) (scleroderma). An NF-κB binding site specific to allele *2 may contribute to the association of this allele with chronic inflammatory autoimmune diseases.

Brief report

Autoimmune diseases are associated with environmental and genetic risk factors that affect severity, age of first symptoms and disease course. Genome-wide association (GWA) and linkage studies were developed to identify the genetic basis of these multiple factors. Complementary to this approach are studies on candidate genes involved in immune response or predicted by GWA or linkage disequilibrium analyses (1). Here we have analyzed a cis regulatory region of Igh locus for its association with possible functional allelic differences and autoimmune diseases using a case-controlled approach.

The regulation of the humoral immune response depends upon cis- and trans-acting elements, among which is a 3' regulatory region of the immunoglobulin (Ig) heavy chain locus (2, 3). Human Ig constant region genes are duplicated, with each segment followed by a 3' regulatory region (3'RR), (4) (supplementary Figure 1A). Each 3'RR contains enhancer elements HS3, HS1.2 and HS4, which, as shown in mice, are involved in regulating B cell receptor expression, Ig germline transcription and class switching (3). While HS3 and HS4 enhancers in 3' RR-1 are identical to those in 3'RR-2, the HS1.2 enhancer is polymorphic, with variable numbers of tandem repeats and spacers, accompanied by predicted differences in transcription factor binding sites (5, 6). In this study, we report the frequencies of the four alleles of HS1.2 of 3'RR-1 in three cohorts of patients [158 with SLE, 233 with RA (7), and 135 with SSC (8)] compared to those obtained in 573 control subjects of the general population of the same area (Table 1). HS1.2 allele frequency in each of the three cohorts (total=526) is very similar (<3% variation, Sup Table1, Sup Figure2). 3'RR-1 HS1.2 alleles *1 and *2 are the most prevalent, while HS1.2 alleles *3 and *4 are poorly represented in patients and general populations with no significant changes. Notably, in each patient cohort, allele *2 homozygousity is almost double that of the controls. These data suggest that HS1.2 allele *2 is associated with these chronic inflammatory autoimmune diseases (Table 1, suppl. Fig2). Compared to 3' RR-1, only modest changes in 3' RR-2 allele frequency were detected, therefore neglected in this study.

Sequence differences in 3' RR-1 HS1.2 alleles*1 and *2 are associated with predicted differences in transcription factor binding sites (5). Electrophoretic mobility shift assays (EMSA) have been used to locate binding sites for transcription factors SP1 and NF-κB in HS1,2 alleles (9). Additional EMSA with allele *1 or *2 as probes was carried out with nuclear protein extracts derived from the human pro-B cell line FLEB14 (Figure 1). Only one band, a, was detected with allele *1 while three bands, a, b, and c were detected with allele *2. Using segments of allele *2 as potential competitors, or specific antibodies against

transcription factors for supershift analysis, we found that the single band of allele *1 (a) is competed by SP1 (Panel A, lane 8) and not by the NF-kB consensus (lane 7). Sp1 binding also accounts for band a (and c (4)) of allele *2 (Panel B, lane 3). Sp1 sites are present on fragments 1 and 3 (Panel B and C), and localized in Fragment 1 to an 18 bp element (Panel D) (see supplementary Fig. 1). Complete (Panel C) or partial (Panel D) deletion of the 18 bp segment from fragment 1 eliminated the competition of this fragment for Sp1 binding to band a of allele *2. Panel B shows that NF-κB uniquely binds to allele *2, band b (lane 7) and to fragment 2 of this allele. Band c is present only in allele *2 complexes and is competed by fragments 1 and 2, and the 18 bp element. Previous studies showed that anti-Sp1 supershifted this band as well as band a (8). These results are consistent with competition of EMSA by fragments of allele *1 and *2, as shown in lanes 4, 5 and 6 of panels A and B.

The 3'RR-2 region also contains multiple alleles, two of which-- alleles *1 and *2-- are very poorly represented in the population (Frezza, unpublished observation): allele *3 is modestly present and allele *4 is predominantly present (5). These latter alleles bear at high (> 95%) frequency, a 29 bp element that replaces the SP1-binding 18 bp element at the 5' end of the polymorphic region. This 29 bp element is predicted to contain a USF site that could endow the 3'RR-2 with an insulator role (5). The binding of HS1.2 and HS4 enhancers to nuclear protein complexes, including NF-κB, was previously demonstrated (10, 11). We, therefore, anticipate that specific haplotypes linked to the HS1.2 enhancer could drive the risk for the diseases. In fact, in the 4 kb segment where HS1.2 is embedded, fewer haplotypes are linked to allele *2 than to allele *1 (results not shown). This suggests an instability of allele *2 in the presence of changes in the surrounding region, probably due to conformations, mobility after epigenetic changes when bound to protein factors.

This study shows a significant increase in the frequency of allele *2 compared to allele *1 of the HS1.2 enhancer downstream of Cα1 of the Ig heavy chain locus in patients with three

systemic autoimmune diseases. A similar association of HS1.2 allele *2 with celiac disease (12) and psoriasis (13) was also observed, thus reinforcing its relationship with autoimmunity. Recent observations revealed HS1.2 polymorphism related to IgA deficiency (14) and to relative levels of Igh expression (9). EMSA indicates that NFκB binds uniquely to a central region of allele *2, which is absent from allele *1. NF-kB is a pivotal factor in B cell maturation, involved either in innate as well as in adaptive immunity(15). On these basis we predict that HS1.2 allele *2, can bias toward a pathological immune response in conjunction with other factors that interact with HS1.2 enhancer binding complexes. It will be important to determine the influence, at the epigenetic level, of haplotypes of the Igh 3' RRs in the immune response.

Acknowledgments

This research was partially supported by D. Frezza grant MIUR PRIN N. 20073RH73W003, University of Tor Vergata annual fund, by ASRALES foundation to G.F.Ferraccioli, and by NIH R01AI13509 to B.K.Birshtein.

Legend to Figure 1

Localization of transcription factors binding sites of NFκB and Sp1 in HS1.2 alleles *1 and *2.

EMSA with nuclear extracts of FLEB human pro-B cell line incubated with probes of allele *1 (Panel A) and *2 (Panels B-D). Competitions were performed with fragments of the two alleles (schematized below), consensus sequences for NF-κB and Oct-1, specific antibody against Sp1 as described in supplementary material.

Panel A) A single band is shown for allele *1, competed by allele *2 (lane 3), fragments 1 (lane 4) and 3 (lane 6) of allele *2, and anti-Sp1 antibodies (lane 8), but not by NF-κB (lane 7) and Oct1 (lane 9) consensus sequences.

B) Three bands are detected with allele *2. Band a is competed by allele *1 (lane 3), fragments 1 (lane 4) and 3 (lane 6) of allele *2. Fragment 1 also competes for band c (lane 4), central fragment 2 competes for band b and c (lane 5). Band b is competed by an NF-κB consensus sequence (lane 7), but not by Oct1 consensus (lane 8).

C) An 18 bp element of Fragment 1 of allele *2 contains an Sp1 binding site in band a. The deletion of 18 bp element (lane 4), eliminates competition of Fragment 1 for bands a and c of allele *2 (lane 3). Deletion of 40 mer of fragment 1 has no effect on Fragment 1's binding (lane 5).

D) 18 bp element of Fragment 1 competes for bands a and c. Partial deletion of 18 bp sequence eliminates the binding of intact 18 bp element to a and c (lane 4).

Table I. HS1.2 allelic frequencies of SLE, RA and SSC patients and control population.

Genotypes	**Controls (TOT 573)**			**SLE (TOT 158)**			**RA (TOT 233)**			**SSC (TOT 135)**		
	Obs*	**%**	**Exp**	**Obs***	**%**	**Exp**	**Obs***	**%**	**Exp**	**Obs***	**%**	**Exp**
1/1	115	20	103	23	14.6	19.7	23	9.8	29.2	20	14.9	17.4
2/2	90	15.7	95	51	32.3	48.3	70	30	77.6	45	33.3	44.5
3/3	3	0.5	1.3	-	-	-	-	-	-	-	-	-
4/4	8	1.4	1.3	4	2.5	0.8	5	2.2	0.6	1	0.7	0.2
1/2	189	33	197.8	59	37.3	61.8	113	48.5	95.2	53	39.2	55.6
1/3	21	3.7	23.3	2	1.3	1.7	3	1.3	2.8	2	1.5	2.4
1/4	47	8.2	57.3	5	3.2	8.3	3	1.3	8.4	2	1.5	3.9
2/3	27	4.7	22.4	3	1.9	2.6	5	2.2	4.6	5	3.7	3.9
2/4	71	12.4	55	11	6.9	13.1	11	4.7	13.7	7	5.2	6.2
3/4	2	0.4	6.5	-	-	0.4	-	-	0.4	-	-	0.3

Alleles	Frequencies	Obs**	Frequencies	Obs**	Frequencies	Obs**	Frequencies	Obs**
***1**	0.424 ± 0.014	487	0.354 ± 0.026	112	0.354 ± 0.022	165	0.359 ± 0.029	97
***2**	0.407 ± 0.014	467	0.553 ± 0.028	175	0.577 ± 0.022	269	0.574 ± 0.030	155
***3**	0.048 ± 0.006	56	0.015 ± 0.006	5	0.017 ± 0.005	8	0.025 ± 0.009	7
***4**	0.118 ± 0.009	136	0.075 ± 0.015	24	0.051 ± 0.010	24	0.040 ± 0.011	11

Materials and experimental procedures

All RA patients met the American College of Rheumatology (formerly, the American Rheumatism Association) 1987 revised criteria for the classification of RA (Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988; **31**: 315-24).

All patients met the American College of Rheumatology (ACR) revised criteria for the classification of SLE (Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.).

SSc patients were required to fulfill the American College of Rheumatology criteria for SSc (Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum 1980;23:581–90.) and were classified as having limited cutaneous SSc (lcSSc) or diffuse cutaneous SSc (dcSSc) according to the criteria described by LeRoy et al (LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T,Medsger TA Jr, et al. Scleroderma (systemic sclerosis): classification,subsets and pathogenesis. J Rheumatol 1988;15:202–5).

DNA extraction was performed with standard methods for blood samples with FlexiGene DNA kit (Qiagen, Valencia, CA, USA), and for swab buccal samples, as

described in. Giambra V. et al. Gene 2005. Selective PCR of polymorphic HS1.2 3'RR1 region was performed as described in Giambra V. et al. Gene 2005 (6).

EMSA experiments

Electrophoretic mobility shift assay (EMSA)

Nuclear extract preparation and EMSA were performed as described **(**9**);**

The sequences for the probes were : allele *1 (227 bp) accession number AJ544218.2; GI:41223289 and allele *2 (279 bp) accession number AJ544219.2; GI:41223290. Both were deleted for the first 60 bp bearing the Oct1 consensus. The following correspond respectively to the 5' region (Fragment 1) 50 bp long 5'AGCTTGCACGATTCTCCCGCCCCCTCCCCCAGCGTGGCCAGGCTGGCTCA -3', the center (Fragment 2) was 60 bp long 5'GGCTGGCTCAGGCCTCCAGATTCGGGGACACCCCCCCACCACAGCGTGGCCAG GCTGGCT-3' , and 3' border of the *2A allele. (Fragment 3) 60 bp long 5'CAGGCTGGCTCAGGCCTCCAGATTCGGGGACACCCGCCCCCTCCCCCAGGACA GGCCCAC-3'. The competitor Fragment 1-Δ18 was 50 bp long 5'AGGCTGTTTCCAGAAATAGCTTGCACGATTAGCGTGGCCAGGCTGGCTCA3' and competitor Fragment 1-Δ40 was 50 bp long 5'TTCCAGAAATAGCTTGCACGATTCTCCCGCCCCCTCCCCCGGCTGGCTCA3', the competitor Fragment 18 was 25 bp long 5' GATTCTCCCGCCCCCTCCCCCAGCG 3', Fragment 18 mutated was 25 bp long 5' ATAGCTTGCACGATTCCCCCCAGCG 3'. The competitor sequence for octamer binding proteins was 5'- CCA TTT GCA TAT TTG CAT ATT TGC ATC C-3' (9). The competitor for NF-κB was: 5'-GAG AGG GGA TTC CCC GAT TAG CTT TCG GGG AAT CCC CTC T-3' (9). The antibody used for the SP1 subtractive assay was obtained from Santa Cruz Biotechnologies (Santa Cruz, California, USA) (PEP2) (sc-59) at a concentration of 200 µg/ml.

REFERENCES

1. Manolio T.A. et al. Nature. 461, 747-753 (2009) review.

2. Laurencikiene J., Tamosiunas V. and Severinson E. Blood 109, 159-167 (2007).

3. Cogne M. and Birshtein B.K. , Chapter 19 in Molecular Biolgoy of B cells . pp. 289-305. Elsevier, New York (2004).

4. Mills C.F. et al. J. Exptl. Med. 186, 845-858 (1997).

5. Chen C, Birshtein BK. J.Immunol. 159, 1310-1318 (1997).

6. Giambra V. et al. Gene. 346, 105-114 (2005).

7. Tolusso B. et al. Ann Rheum Dis. 68, 416-419, (2009).

8. Frezza D. et al. Ann Rheum Dis. 66, 1210-1215, (2007).

9. Frezza D. et al. Int. J. Immunopathol. Pharmacol. 22, 115-23, (2009).

10. Michaelson JS, et al. J.Immunol. 156, 2828-2839, (1996).

11. Sepulveda M.A. et al. J.Immunol. 172, 1054-1064, (2004).

12. Frezza D. et al. Scand. J. Gastroenterol. 39, 1083-1087, (2004).

13. Cianci R. et al. J. Invest. Dermatol. 128, 1920-1924, (2008).

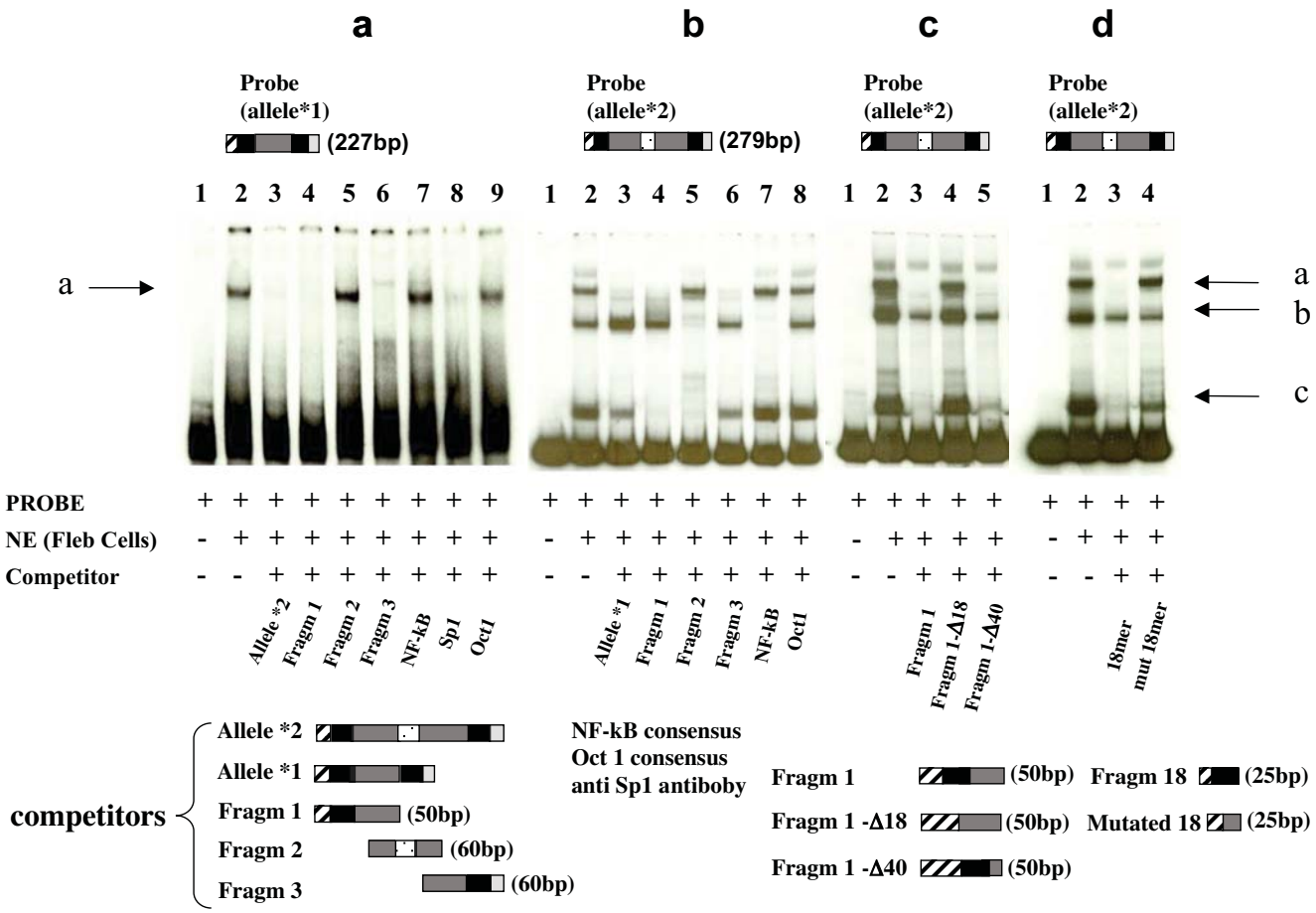
14. Giambra V. et al. J.Immunol. 183, 8280-8285, (2009).

15. Caamano J, Hunter C.A. Clin. Microbiol. Rev. 15, 414–429, (2002).

Table I. HS1.2 allelic frequencies of SLE, RA and SSC patients and control population.

Genotypes	Controls (TOT 573)			SLE (TOT 158)			RA (TOT 233)			SSC (TOT 135)		
	Obs*	%	Exp	Obs*	%	Exp	Obs*	%	Exp	Obs*	%	Exp
1/1	115	20	103	23	14.6	19.7	23	9.8	29.2	20	14.9	17.4
2/2	90	15.7	95	51	32.3	48.3	70	30	77.6	45	33.3	44.5
3/3	3	0.5	1.3	-	-	-	-	-	-	-	-	-
4/4	8	1.4	1.3	4	2.5	0.8	5	2.2	0.6	1	0.7	0.2
1/2	189	33	197.8	59	37.3	61.8	113	48.5	95.2	53	39.2	55.6
1/3	21	3.7	23.3	2	1.3	1.7	3	1.3	2.8	2	1.5	2.4
1/4	47	8.2	57.3	5	3.2	8.3	3	1.3	8.4	2	1.5	3.9
2/3	27	4.7	22.4	3	1.9	2.6	5	2.2	4.6	5	3.7	3.9
2/4	71	12.4	55	11	6.9	13.1	11	4.7	13.7	7	5.2	6.2
3/4	2	0.4	6.5	-	-	0.4	-	-	0.4	-	-	0.3

Alleles	Frequencies	Obs**	Frequencies	Obs**	Frequencies	Obs**	Frequencies	Obs**
***1**	0.424 ± 0.014	487	0.354 ± 0.026	112	0.354 ± 0.022	165	0.359 ± 0.029	97
***2**	0.407 ± 0.014	467	0.553 ± 0.028	175	0.577 ± 0.022	269	0.574 ± 0.030	155
***3**	0.048 ± 0.006	56	0.015 ± 0.006	5	0.017 ± 0.005	8	0.025 ± 0.009	7
***4**	0.118 ± 0.009	136	0.075 ± 0.015	24	0.051 ± 0.010	24	0.040 ± 0.011	11


a
b
c
d
Probe (allele*1)
(227bp)
Probe (allele*2)
(279bp)
Probe (allele*2)
Probe (allele*2)
1 2 3 4 5 6 7 8 9
1 2 3 4 5 6 7 8
1 2 3 4 5
1 2 3 4
a
a
b
c
PROBE
NE (Fleb Cells)
Competitor
Allele *2
Fragm 1
Fragm 2
Fragm 3
NF-kB
Sp1
Oct1
Allele *1
Fragm 1
Fragm 2
Fragm 3
NF-kB
Oct1
Fragm 1
Fragm 1-Δ18
Fragm 1-Δ40
18mer
mut 18mer
competitors
Allele *2
Allele *1
Fragm 1 (50bp)
Fragm 2 (60bp)
Fragm 3 (60bp)
NF-kB consensus
Oct 1 consensus
anti Sp1 antiboby
Fragm 1 (50bp)
Fragm 18 (25bp)
Fragm 1 -Δ18 (50bp)
Mutated 18 (25bp)
Fragm 1 -Δ40 (50bp)