



Short sequence-paper

Z39Ig is co-expressed with activated macrophage genes

Michael G. Walker

Incyte Genomics, 3160 Porter Drive, Palo Alto, CA 94304, USA

Received 21 August 2001; accepted 15 November 2001

Abstract

Z39Ig is a recently-identified gene with immunoglobulin-like domains whose function is unknown. We examined expression of Z39Ig in 1432 human cDNA libraries, and found it primarily in synovium of patients with rheumatoid arthritis, in placenta, and in lung. We analyzed its co-expression pattern using the Guilt-by-Association (GBA) algorithm, and found that it is most similar in expression to early genes in the classical complement system (C1qA, C1qB, C1qC, C1r, and C1 inhibitor), MHC class II genes (HLA-DR alpha, HLA-DR beta 1, and HLA-DP alpha 1), Fc receptors (Fc gamma RIIa and Fc epsilon R1), lysosomal protein (LAPTM5), tissue transglutaminase, and macrophage receptors (MARCO and CD163/M130). The sequence and expression data suggest that Z39Ig is a cell surface receptor, expressed in activated macrophages, and linked with the classical complement system, most likely in phagocytosis preceding antigen presentation. Knowledge of this gene may contribute to better understanding of the role of complement and activated macrophages in rheumatoid arthritis and systemic lupus. © 2001 Published by Elsevier Science B.V.

Keywords: Z39Ig; Gene expression analysis; Activated macrophage

Z39Ig is a recently-identified gene of unknown function. In a search for genes possibly involved in mental retardation, Langnaese and colleagues studied Z39Ig, which maps to a region of the human X chromosome associated with retardation [1]. They examined expression in 16 pooled cDNA libraries: eight fetal (brain, heart, kidney, liver, lung, skeletal muscle, spleen, and thymus), and eight adult (brain, heart, kidney, liver, lung, skeletal muscle, placenta, and pancreas), and found it expressed widely in fetal tissue, but primarily in lung and placenta in adults. (Their libraries did not include synovium.) Sequence analysis identified a signal peptide region, a transmembrane region, and two immunoglobulin-like domains. The expression results indicated that Z39Ig was not a good candidate for a role in mental retardation, but the gene and its potential role in disease is otherwise uncharacterized. To further characterize Z39Ig, we examined its expression pattern in a set of human cDNA libraries from diverse tissues, and compared its expression pattern to the expression patterns of genes with known functions.

We examined the expression of genes in 1432 human cDNA libraries. These libraries were from diverse anatomic and pathologic states, mainly from surgery, biopsy, or post-mortem samples, or were prepared from cell lines,

and include all the libraries that were in the Incyte LifeSeq database at the time of the analysis. Surgical samples often contained mixed tissue; no samples were microdissected. Some libraries were subtracted or normalized to enrich rare mRNA. Approximately 5000 cDNAs from each library were sequenced by gel electrophoresis, assembled, and aligned against known genes. All genes that were detected in at least five of the 1432 libraries were included in the analysis described here, which yielded 37 071 genes, gene fragments, or splice variants. Similar gene expression databases are available at the US National Cancer Institute (www.ncbi.nlm.nih.gov/UniGene/ddd.cgi), or will be shortly through the European Bioinformatics Institute (www.ebi.ac.uk) or the Gene Expression Omnibus at the US National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/geo/).

To identify genes with a similar expression pattern to Z39Ig, we performed co-expression analysis using the Guilt-by-Association (GBA) algorithm. We previously used GBA to identify genes that are co-expressed with prostate cancer genes [2], cell cycle genes [3], cardiovascular genes [4], and neurotransmitter genes [5]. Briefly, in a GBA analysis, we consider a gene to be present (expressed) in a library if cDNA corresponding to that gene is detected in the sample taken from that library. We consider a gene to be absent (not expressed) in a library when no cDNA for that gene is detected in the library. For a

E-mail address: mwalker@stanfordalumni.org (M.G. Walker).

given pair of genes, the co-expression data can be summarized in a two-by-two contingency table, where the contingency table entries indicate one of four categories: both genes detected, neither detected, and one or the other gene detected. From the contingency table, we determine the probability that the co-expression occurs by chance using a chi-square test or a Fisher Exact Test [6].

Z39Ig mRNA is detected in 148 of the 1432 cDNA libraries. It is most abundant in synovium (eight of 10 samples, including six of seven from rheumatoid arthritis patients). It is next most abundant in placenta (detected in seven of 16 libraries, all post week 12), followed by lung (in 26 of 87 libraries). It is detected at lower levels in endocrine and exocrine tissues, detected occasionally in cardiovascular, digestive, and nervous tissues, and rare in other tissues.

Table 1 shows genes with expression patterns similar to that of Z39Ig by GBA analysis. For all the co-expressed genes in Table 1, the probability that the co-expression with Z39Ig is due to chance is less than $1.0 \times e^{-15}$ by the Fisher Exact Test. What do the genes with which Z39Ig is co-expressed indicate about the cell types in which it is likely expressed?

The scavenger receptors MARCO [7–9] and CD163/M130 are expressed specifically in monocytes and macrophages, are upregulated in activated macrophages, and participate in phagocytosis of pathogens and waste products [10–14]. The MHC class II genes (HLA-DR alpha, HLA-DR beta 1, and HLA-DP alpha 1) are expressed specifically by antigen presenting cells, including macrophages, dendrites, and B cells; they are expressed at low levels in resting macrophages, and upregulated by activation. Susceptibility to develop rheumatoid arthritis is associated with specific HLA-DR beta 1 alleles [15–22]. Like the MHC class II genes, the early genes of the classical complement system (C1qA, C1qB, C1qC, C1r, and C1 inhibitor) are expressed at low levels in resting macro-

phages and upregulated by activation [23–26]. Complement protein synthesis by synovial macrophages and complement activation are characteristic of rheumatoid arthritis [27–31]. Other cell types also express early genes of the classical system. The Fc receptor subunits Fc gamma RIIa and Fc epsilon R1 are expressed in macrophages and other leukocytes including antigen presenting cells; Fc gamma receptors phagocytose antigen-antibody complexes, inducing antigen presentation by MHC II proteins, and also participate in signal transduction [32–38]. Altered Fc epsilon R1 gamma expression has been linked to SLE [39]. Lysosomal protein (LAPTM5) is a multitransmembrane protein that is expressed preferentially in hematopoietic cell lines, is localized to lysosomes, and binds to ubiquitin [40]. Adra and colleagues found high expression of LAPTM5 in peripheral blood leukocytes, thymus, spleen, and liver in adult human tissues, while in fetal tissue it was broadly expressed. Tissue transglutaminase (TGM2) is an enzyme that catalyzes protein cross-linking and is expressed in endothelial cells and macrophages [41–43]. Among its many functions, it participates in antigen processing and may specifically participate in receptor-dependent endocytosis and subsequent antigen presentation [44,45]. Activity is elevated in aging degenerative cartilage [46], and it has been linked to autoimmune disorders, particularly celiac disease [47–49]. HAIRB, recently isolated from dendritic cells, is similar in sequence to Fc epsilon RI beta-related proteins of the MS4A family [50]. Human epididymous secreted protein (HE2) appears to be a descendant of defensin; it is similar in sequence, located in the same chromosomal region, and splice variants act as defensins, which are antimicrobial and chemokines for adaptive immune cells [51,52]. Co-expression of Z39Ig with these genes indicates that it is expressed in activated macrophages and is likely involved in phagocytosis preceding antigen presentation.

Is expression of Z39Ig in activated macrophages consist-

Table 1
Genes co-expressed with Z39Ig (*P*-values from Fisher Exact Test)

<i>P</i> -value	GenBank	Gene short name	Gene description
$1.84 \times e^{-23}$	g1255239	LAPTM5	lysosomal-associated membrane protein
$1.34 \times e^{-22}$	g50229	C1q C	C1q C complement chain
$1.70 \times e^{-22}$	g29537	C1q B	C1q B complement chain
$1.74 \times e^{-22}$	g12005800	HAIRB	HAIRB
$2.85 \times e^{-21}$	g4894853	C1q A	C1q A complement chain
$3.24 \times e^{-20}$	g188231	HLA-DR alpha	HLA-DR alpha
$7.75 \times e^{-19}$	g31328	Fc gamma RII	Fc receptor gamma II
$4.01 \times e^{-18}$	g36405	HLA-DP alpha 1	HLA-DP alpha 1
$8.81 \times e^{-18}$	g339520	TGase	Tissue transglutaminase
$1.20 \times e^{-17}$	g188240	HLA-DR beta 1	HLA-DR beta 1
$1.43 \times e^{-17}$	g179620	C1 inhibitor	C1 complement inhibitor
$2.33 \times e^{-17}$	g182487	Fc epsilon R1 gamma	Fc epsilon receptor gamma 1
$2.56 \times e^{-17}$	g29538	C1r	C1r complement
$3.06 \times e^{-17}$	g312143	CD163	CD163/M130 macrophage receptor
$5.23 \times e^{-17}$	g818880	HE1	epididymal secretory protein
$2.95 \times e^{-16}$	g3002790	MARCO	macrophage receptor MARCO

tent with the types of tissues in which it is most commonly detected? The tissues in which Z39Ig is most abundant – synovium from rheumatoid arthritis patients, placenta, and lung – are tissues in which activated macrophages are particularly abundant, and in which the complement system genes C1q A, C1q B, C1q C, C1r, and the C1 inhibitor are highly expressed. The chronic joint inflammation associated with RA is mediated primarily by macrophages responding to reactive T cells. The occurrence of Z39Ig in placental and fetal libraries is consistent with the known expression of the complement genes early in gestation.

Expression analysis using a database of cDNA libraries provides hypotheses about the likely tissue specificity and function of Z39Ig, but these hypotheses need confirmation in direct experiments with activated macrophages. The primary utility of an expression database analysis is to suggest experiments that are most likely to be fruitful, thereby saving research time and expense.

The GBA analysis makes several assumptions that are violated to greater or lesser degrees by aspects of the library selection and preparation. For example, libraries are not completely independent, because more than one library may be obtained from a single patient (for example, multiple organs, or matched tumor and non-tumor tissue, or normalized and non-normalized). Normalizing or subtracting makes the detection of associations between genes expressed at different levels more difficult. Genes expressed at low levels may not be detected because of random sampling. The cDNA libraries used in this analysis were prepared at different times and with different methods, and may not be consistent. The effect of different cDNA library samples, different normalization, different preparation methods, or preparation at different times is most likely to be to obscure true relationships. Such differences will make the calculated probability of association less accurate. However, it is unlikely that a pattern that is consistent across 1432 libraries, has good *P*-values, and is consistent with known biological relationships, would be introduced by the cumulative random effects of such differences. Thus, the GBA analysis will yield false negatives, but it is unlikely to yield false positive results with a database of this size.

The sequence and expression pattern suggest that Z39Ig is a cell surface receptor, expressed in activated macrophages, linked with the classical complement system, and likely involved in phagocytosis leading to antigen presentation. Better understanding of this gene may help us mediate the adverse effects of complement and activated macrophages in rheumatoid arthritis and systemic lupus.

Colleagues at Incyte Genomics provided the LifeSeq database and contributed to the development of the GBA analysis method. LifeSeq is a trademark of Incyte Genomics, Inc.

References

- [1] K. Langnaese, L. Colleaux, D.U. Kloos, M. Fontes, P. Wieacker, *Biochim. Biophys. Acta* 1492 (2000) 522–525.
- [2] M.G. Walker, W. Volkmuth, E. Sprinzak, D. Hodgson, T. Klingler, *Genome Res.* 9 (1999) 1198–1203.
- [3] M.G. Walker, *Curr. Cancer Drug Targets* 1 (2001) 73–83.
- [4] M.G. Walker, *Mini Rev. Med. Chem.* 1 (2001) 197–205.
- [5] M.G. Walker, W. Volkmuth, T.M. Klingler (1999) in: *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, pp. 282–286.
- [6] A. Agresti, *Categorical Data Analysis*, Wiley, New York, 1990.
- [7] G. Kraal, L.J. van der Laan, O. Elomaa, K. Tryggvason, *Microbes Infect.* 2 (2000) 313–316.
- [8] L.J. van der Laan, E.A. Dopp, R. Haworth, T. Pikkarainen, M. Kangas, O. Elomaa, C.D. Dijkstra, S. Gordon, K. Tryggvason, G. Kraal, *J. Immunol.* 162 (1999) 939–947.
- [9] T. Pikkarainen, A. Brannstrom, K. Tryggvason, *J. Biol. Chem.* 274 (1999) 10975–10982.
- [10] S.K. Law, K.J. Micklem, J.M. Shaw, X.P. Zhang, Y. Dong, A.C. Willis, D.Y. Mason, *Eur. J. Immunol.* 23 (1993) 2320–2325.
- [11] D.J. Schaer, F.S. Boretti, A. Hongegger, D. Poehler, P. Linnscheid, H. Staeger, C. Muller, G. Schoedon, A. Schaffner, *Immunogenetics* 53 (2001) 170–177.
- [12] M. Kristiansen, J.H. Graversen, C. Jacobsen, O. Sonne, H.J. Hoffmann, S.K. Law, S.K. Moestrup, *Nature* 409 (2001) 198–201.
- [13] M. Ritter, C. Buechler, M. Kapinsky, G. Schmitz, *Eur. J. Immunol.* 31 (2001) 999–1009.
- [14] N.R. Ratcliffe, S.M. Kennedy, M. Peter, *Immunol. Lett.* 77 (2001) 169–174.
- [15] E. Toussiro, D. Wendling, *Presse Med.* 24 (1995) 855–858.
- [16] I. Auger, J. Roudier, *Autoimmunity* 26 (1997) 123–128.
- [17] C.H. van Jaarsveld, H.G. Otten, J.W. Jacobs, A.A. Kruijze, H.L. Brus, J.W. Bijlsma, *Clin. Exp. Rheumatol.* 16 (1998) 483–488.
- [18] C.H. Van Jaarsveld, H.G. Otten, J.W. Jacobs, A.A. Kruijze, H.L. Brus, J.W. Bijlsma, *Br. J. Rheumatol.* 37 (1998) 411–416.
- [19] C. Wilson, H. Tiwana, A. Ebringer, *Microbes Infect.* 2 (2000) 1489–1496.
- [20] S. Kerlan-Candon, P. Louis-Pence, A. Wiedemann, B. Combe, J. Clot, J.F. Eliaou, V. Pinet, *Arthritis Rheum.* 44 (2001) 1281–1292.
- [21] S. Kerlan-Candon, B. Combe, R. Vincent, J. Clot, V. Pinet, J.F. Eliaou, *Clin. Exp. Immunol.* 124 (2001) 142–149.
- [22] J.P. Hellier, J.F. Eliaou, J.P. Daures, J. Sany, B. Combe, *Ann. Rheum. Dis.* 60 (2001) 531–533.
- [23] J.C. Bensa, A. Reboul, M.G. Colomb, *Biochem. J.* 216 (1983) 385–392.
- [24] M.T. Rimoldi, A.J. Tenner, D.A. Bobak, K.A. Joiner, *J. Clin. Invest.* 84 (1989) 1982–1989.
- [25] A.R. McPhaden, K. Whaley, *Immunol. Res.* 12 (1993) 213–232.
- [26] P. Gulati, C. Lemercier, D. Guc, D. Lappin, K. Whaley, *Behring Inst. Mitt.*, 1993, pp. 196–203.
- [27] K. Whaley, D. Guc, P. Gulati, D. Lappin, *Immunopharmacology* 24 (1992) 83–89.
- [28] D. Guc, P. Gulati, C. Lemercier, D. Lappin, G.D. Birnie, K. Whaley, *Rheumatol. Int.* 13 (1993) 139–146.
- [29] P. Gulati, D. Guc, C. Lemercier, D. Lappin, K. Whaley, *Rheumatol. Int.* 14 (1994) 13–19.
- [30] S. Breitner, S. Storkel, W. Reichel, M. Loos, *Arthritis Rheum.* 38 (1995) 492–498.
- [31] E.T. Molenaar, A.E. Voskuyl, A. Familian, G.J. van Mierlo, B.A. Dijkmans, C.E. Hack, *Arthritis Rheum.* 44 (2001) 997–1002.
- [32] H. Heiken, R.J. Schulz, J.V. Ravetch, E.L. Reinherz, S. Koyasu, *Eur. J. Immunol.* 26 (1996) 1935–1943.
- [33] S. Amigorena, C. Bonnerot, *Immunol. Rev.* 172 (1999) 279–284.
- [34] S. Amigorena, C. Bonnerot, *Semin. Immunol.* 11 (1999) 385–390.
- [35] V.C. Huber, J.M. Lynch, D.J. Bucher, J. Le, D.W. Metzger, *J. Immunol.* 166 (2001) 7381–7388.

- [36] S. Kraft, T. Bieber, *Int. Arch. Allergy Immunol.* 125 (2001) 9–15.
- [37] S. Kraft, N. Katoh, N. Novak, S. Koch, T. Bieber, *Int. Arch. Allergy Immunol.* 124 (2001) 35–37.
- [38] S.A. Minskoff, K. Matter, I. Mellman, *J. Immunol.* 161 (1998) 2079–2083.
- [39] E.J. Enyedy, M.P. Nambiar, S.N. Liossis, G. Dennis, G.M. Kammer, G.C. Tsokos, *Arthritis Rheum.* 44 (2001) 1114–1121.
- [40] C.N. Adra, S. Zhu, J.L. Ko, J.C. Guillemot, A.M. Cuervo, H. Kobayashi, T. Horiuchi, J.M. Lelias, J.D. Rowley, B. Lim, *Genomics* 35 (1996) 328–337.
- [41] V. Gentile, M. Saydak, E.A. Chiocca, O. Akande, P.J. Birckbichler, K.N. Lee, J.P. Stein, P.J. Davies, *J. Biol. Chem.* 266 (1991) 478–483.
- [42] L. Lavie, O. Weinreb, *Mech. Ageing Dev.* 90 (1996) 129–143.
- [43] Z.A. Haroon, T.S. Lai, J.M. Hettasch, R.A. Lindberg, M.W. Dewhirst, C.S. Greenberg, *Lab. Invest.* 79 (1999) 1679–1686.
- [44] K. Teshigawara, R. Kannagi, N. Noro, T. Masuda, *Microbiol. Immunol.* 29 (1985) 737–750.
- [45] J.S. Jennings, P.D. Howdle, *Curr. Opin. Gastroenterol.* 17 (2001) 118–126.
- [46] K. Johnson, S. Hashimoto, M. Lotz, K. Pritzker, R. Terkeltaub, *Am. J. Pathol.* 159 (2001) 149–163.
- [47] K. Sjoberg, R. Wassmuth, S. Reichstetter, K.F. Eriksson, U.B. Eriksson, S. Eriksson, *Autoimmunity* 32 (2000) 217–228.
- [48] A.V. Dahele, M.C. Aldhous, K. Humphreys, S. Ghosh, *Qjm* 94 (2001) 195–205.
- [49] V. Kumar, M. Rajadhyaksha, J. Wortsman, *Clin. Diagn. Lab. Immunol.* 8 (2001) 678–685.
- [50] Y. Liang, T.F. Tedder, *Genomics* 72 (2001) 119–127.
- [51] O. Frohlich, C. Po, L.G. Young, *Biol. Reprod.* 64 (2001) 1072–1079.
- [52] H.P. Jia, B.C. Schutte, A. Schudy, R. Linzmeier, J.M. Guthmiller, G.K. Johnson, B.F. Tack, J.P. Mitros, A. Rosenthal, T. Ganz, P.B. McCray Jr., *Gene* 263 (2001) 211–218.