

Novel coding genetic variants of the *GBP1* gene in wild and domestic pigs (*Sus scrofa*)

Shanyuan Chen^a, Rui Gomes^a, Vânia Costa^a, Isabel Rocha^a, Attila Zsolnai^{b,c}, István Anton^b, Rui Charneca^d, Pedro Santos^d, José Luis Nunes^d, József Buzgó^e, Gyula Varga^e, Ya-ping Zhang^f, Albano Beja-Pereira^{a,*}

^a Centro de Investigação em Biodiversidade e Recursos Genéticos da Universidade do Porto (CIBIO-UP), Rua Padre Armando Quintas 7, 4485-661 Vairão, Portugal

^b Research Institute for Animal Breeding and Nutrition, H-2053 Herceghalom, Hungary

^c University of Kaposvár, H-7400, Kaposvár, Hungary

^d ICAAM - Instituto de Ciências Agrárias e Ambientais Mediterrânicas, Universidade de Évora, Herdade da Mitra, 7002-554 Évora, Portugal

^e SEFAG Forest Management and Wood Industry Share Company, Kaposvár, Hungary

^f State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China

ARTICLE INFO

Article history:

Received 13 July 2011

Received in revised form 30 January 2012

Accepted 7 February 2012

Keywords:

Domestic pig

Wild boar

GBP1

Genetic variability

SNP

Rare alleles

ABSTRACT

The interferon-induced guanylate-binding protein 1 gene (*GBP1*) plays an important role in host defense against viral, bacterial and protozoan infections. To explore novel genetic variants in this gene, we re-sequenced a 587-bp fragment spanning the exon 2 of the *GBP1* gene in a sample panel consisting of 34 wild boars and 59 local domestic pigs from three geographic regions (China, Iberian Peninsula, and Central Europe) and 12 individuals of three commercial breeds (Pietrain, Landrace, and Large White). In a final 543-bp sequence fragment, there were 14 single nucleotide polymorphisms (SNPs), of which five were coding (three novel mutations). A total of 19 haplotypes were reconstructed and most haplotypes were shared by two or more sample groups. Those shared haplotypes revealed a clear signature of genetic introgression from Chinese domestic pigs into European domestic pigs. In addition, there were six haplotypes with frequencies below 1%, but none of them were present in the three commercial breeds (Pietrain, Landrace, and Large White). Although a limited number of individuals and breeds were analyzed, the absence of rare alleles (or haplotypes) in the commercial breeds is an indication that a significant proportion of genetic diversity in domestic species is not present in commercial breeds. This study demonstrated the potential to find sufficient genetic variation for population genetic analyses of demography versus selection, in functional candidate genes of domestic pigs and wild boars worldwide.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Interferon (IFN) proteins are essential for host defense against intracellular pathogens and tumor cells, by triggering transcriptional up-regulation of thousands of genes (Borden et al., 2007). Many of the most abundant proteins induced by IFNs are guanosine triphosphatases (GTPases), which constitute four major families – the p47 immunity-related

GTPases (IRGs), the p65 guanylate-binding proteins (GBPs), Mx proteins, and very large inducible GTPases (VLIG) (MacMicking, 2004; Martens and Howard, 2006).

In recent years, the GBP gene family has received much attention, due to their multiple functions (reviewed in Vestal and Jeyaratnam, 2010). Originally, two members of GBPs (GBP1 and 2) were identified as proteins in human fibroblasts treated with IFNs, particularly IFN- γ (Cheng et al., 1983). An early study showed that human GBP1 mediated an antiviral effect against vesicular stomatitis virus and encephalomyocarditis virus (Anderson et al., 1999). Recently,

* Corresponding author. Tel.: +351 252 660411; fax: +351 252 661780.
E-mail address: albanobp@fc.up.pt (A. Beja-Pereira).

extensive transcriptome and genome analyses revealed additional members of GBPs in human (hGBP1 to 7) and mice (mGBP1 to 11) (Kresse et al., 2008; Olszewski et al., 2006). Moreover, recent functional studies also demonstrated that some members of the GBP family (e.g., GBP1, GBP6, GBP7 and GBP10) played important roles in host defense against bacterial and protozoan infections (Degrandi et al., 2007; Kim et al., 2011). Besides those antiviral and antimicrobial activities, members of the GBP family also have other functional effects, such as inhibition of proliferation of endothelial cells, inhibition of matrix metalloproteinase expression, and intestinal epithelial development (Schnoor et al., 2009; Vestal and Jayaratnam, 2010).

GBP1 is the best studied member of the GBP family and its structure and activity has been well characterized. The structure of human GBP1 contains an N-terminal large G domain, a C-terminal α -helical elongated domain, and a short intermediate region connecting domains (Prakash et al., 2000). The large G domain (residues 1–317) retains the main biochemical properties of the full-length protein and two residues Arg48 and Ser73 located in this domain were found to be crucial for nucleotide-dependent homodimerization and cleavage of GTP or GDP (Ghosh et al., 2006). The availability of structural information, together with the well-characterized functions in immunity, makes the *GBP1* gene a good candidate locus for identifying functional genetic variants that potentially underlie disease resistance, in human populations and in economically important livestock species.

Little is so far known about functional genetic variation harbored in the *GBP1* gene of major livestock species. Recent analysis of the cDNA sequences of porcine *GBP1* gene discovered four single nucleotide polymorphisms (SNPs), including two synonymous mutations each in the exon 8 and the exon 11 and two non-synonymous mutations (in the same codon) resulted in an amino acid change from Lysine to Glycine in the exon 2 (Ma et al., 2008). Additionally, association analysis revealed that those two non-synonymous SNPs in the exon 2 had significant associations with red blood cell traits (Ma et al., 2008). Nevertheless, whether there are quantitative trait loci (QTLs) related to immunity that are co-localized with the *GBP1* gene in pig breeds, warrants further study.

Previous genetic studies revealed no indication of an overall reduction in genetic variability of international commercial versus European local pig breeds, at two candidate genes (*FABP4* and *IGF2*) (Ojeda et al., 2006; Ojeda et al., 2008). In contrast, a genome-wide diversity study of worldwide chicken using SNPs demonstrated significant genetic diversity reduction in commercial breeds (Muir et al., 2008). As a result, whether the pattern observed at the *FABP4* and *IGF2* in domestic pigs can be extrapolated to other candidate genes (e.g., *GBP1*), remains an open question. In this study, we re-sequenced a 587-bp fragment spanning the exon 2 of the *GBP1* gene in a sample panel consisting of wild boars and local domestic pigs from three geographic regions (China, Iberian Peninsula, and Central Europe) and three commercial breeds, to further explore novel genetic variants harbored in this gene.

2. Materials and methods

The sample panel was composed of 105 individuals, including 59 domestic pigs, 34 wild boars, and 12 individuals

of three commercial breeds (Table 1). Chinese local pigs were collected from northeastern (five individuals) and northwestern (two individuals) Yunnan, China. Effort was made to collect samples from unrelated individuals. These samples covered three geographic regions – China, Iberian Peninsula, and Central Europe (Hungary and Romania), representing Asian and European lineages or two areas of pig domestication (Giuffra et al., 2000; Larson et al., 2005; Megens et al., 2008).

Genomic DNA was extracted from ear skin tissues using DNeasy Blood & Tissue Kit (Qiagen GmbH, Hilden, Germany). A 587-bp fragment spanning the exon 2 of the *GBP1* gene was amplified by using PCR primer pair Exon 2-SNP-F (5' GGA TAA CAC TTC GGT AAC TTG C 3') and Exon 2-SNP-R (5' GAA GGG GAA ACT GAG ACA CAA T 3'), as previously described (Ma et al., 2008). PCR products were purified and sequenced for both strands at the High-Throughput Genomics Unit (HTGU), Department of Genome Sciences, University of Washington (<http://www.htseq.org/>). The raw sequence trace files were aligned and checked using software package DNASTAR v7.1 (DNASTAR Inc., Madison, WI, USA).

Haplotypes were reconstructed from unphased genotypic sequences by PHASE v2.1.1 algorithm (Stephens and Donnelly, 2003; Stephens et al., 2001), implemented in DnaSP v5.10 (Librado and Rozas, 2009). Diversity measures were calculated in Arlequin v3.5.1.2 (Excoffier and Lischer, 2010). The phylogenetic relations among haplotypes were built by the program Network v4.600 (<http://www.fluxus-engineering.com/sharenet.htm>), using median joining algorithm (Bandelt et al., 1999).

3. Results and discussion

After trimming primer sequences, we finally obtained a 543-bp sequence fragment, corresponding to nucleotide positions between 2118 and 2660 of the porcine *GBP1* reference sequence NC_010446.4 (15,728 bp, Sscrofa10.2 Primary Assembly). This sequence fragment included 164 bp of the intron 1, 200 bp of the exon 2, and 179 bp of the intron 2. There were 14 SNPs detected among all sequences

Table 1
Samples of domestic pigs and wild boars analyzed in this study.

Sample name	Country	<i>n</i>	Sample group
Portuguese Alentejano	Portugal	26	dIB
Portuguese wild boar	Portugal	5	wIB
Spanish wild boar	Spain	12	wIB
Black Slavonian	Hungary	6	dHG
Swallow-bellied Mangalica	Hungary	5	dHG
Red Mangalica	Hungary	5	dHG
Blonde Mangalica	Hungary	5	dHG
Romanian Mangalica	Romania	5	dHG
Hungarian wild boar	Hungary	14	wHG
Chinese local pig	China	7	dCN
Chinese wild boar	China	3	wCN
Pietrain	Hungary	5	CB
Landrace	Hungary	3	CB
Large White	Hungary	4	CB
Total		105	

Abbreviations: *n*, the number of individuals; dIB, Iberian domestic pigs; wIB, Iberian wild boars; dHG, Hungarian and Romanian domestic pigs; wHG, Hungarian wild boars; dCN, Chinese domestic pigs; wCN, Chinese wild boars; CB, commercial breeds.

Table 2
The frequency and distribution of 19 haplotypes across sample groups.

Haplotype	SNPs	N	Frequency	dIB	wIB	dHG	wHG	dCN	wCN	CB
H01	TCTAAATCCGGCC	1	0.0048	1						
H02	TCTAACCCCGGCC	2	0.0095	2						
H03	TCTAACTCCGGCC	4	0.0190	1					3	
H04	TCTAAATGCTACC	2	0.0095	1			1			
H05	CTGGCTGCTACC	4	0.0190	1		1		2		
H06	TCTAACCCCGGCC	8	0.0381	2	3		3			
H07	TCTGGCCCGCGCC	10	0.0476	5	5					
H08	TCTAACCGCTGCC	106	0.5048	20	12	39	12	3		20
H09	TATAACCGCGGCC	27	0.1286	15	3		9			
H10	TCTGGCCGTTACC	3	0.0143	1			2			
H11	TATAACCGCGGCT	1	0.0048	1						

Nucleotide positions of these 14 SNPs corresponding to 2123, 2161, 2216, 2301, 2302, 2314, 2315, 2356, 2502, 2516, 2519, 2526, 2628, and 2641 of the porcine *GBP1* gene reference sequence NC_010446.4; Coding SNPs are shaded in gray; N, the number of sequences ($N=2n$, where n is the number of individuals); Abbreviations of sample groups are provided in Table 1.

(Table 2), of which three, five, and six were located in the intron 1, the coding part of the exon 2 (190 bp, corresponding to 1–190 bp of the *GBP1* whole coding region sequence), and the intron 2, respectively. Among those five coding SNPs, two non-synonymous mutations 2301A>G and 2302A>G that were equal to ns10A>G and ns11A>G reported in a previous study (Ma et al., 2008), occurred at the first and second positions of the 4th codon, causing a replacement from Lysine to Glycine (K4G); Two novel mutations 2314C>A (non-synonymous) and 2315C>A (synonymous) occurred at the second and third positions of the 8th codon, resulting in a change from Proline to Glutamine (P8Q); The final one novel mutation 2356C>T (non-synonymous) occurred at the second position of the 22th codon, leading to a substitution from Alanine to Valine (A22V). Notably, high level of genetic variability observed in the *GBP1* gene (14 SNPs in 543 bp sequences, $14/543=0.026$) from this study is comparable to that found in the *FABP4* gene (134 SNPs in 6400 bp sequences, $134/6400=0.021$) (Ojeda et al., 2006).

A total of 19 haplotypes were reconstructed by PHASE analysis from these 14 SNPs (Table 2). The predominant haplotype H08 (identical to the reference sequence NC_010446.4) with a frequency of 50.48% was distributed in all sample groups except for Chinese wild boars (wCN); while the second common haplotype H09 had a frequency of 12.86%, found in three sample groups (dIB, wIB, and dHG) (Table 2 and Fig. 1). Most of the haplotypes were shared by two or more sample groups, except for seven unique haplotypes (Table 2 and Fig. 1). Those shared haplotypes revealed a clear signature of genetic introgression from Chinese domestic pigs (dCN) into European domestic pigs (dIB and dHG). For instance, the haplotypes H05 and H08 were shared by dCN, dIB, and dHG, while the haplotypes H16 and H17 were shared by dCN and dHG. The genetic contribution of Chinese domestic pigs to European domestic pigs and perhaps indirectly to commercial breeds detected here, was congruent with historical records and previous genetic studies (e.g., Giuffra et al., 2000).

As expected, Chinese domestic pigs (dCN) displayed highest values of both haplotype diversity and nucleotide diversity, whereas the three commercial breeds pooled together (CB) showed lowest value of haplotype diversity (Table 3). Within Europe, both domestic pigs (dIB) and wild boars (wIB) from the Iberian Peninsula exhibited higher levels of genetic variability than those (dHG and wHG) from Central Europe (Hungary and Romania). This pattern of diversity distribution observed here was congruent with previous studies that revealed larger overall genetic variability in Asian (or Chinese) domestic pigs and wild boars than their European counterparts (Megens et al., 2008; Ojeda et al., 2008).

In summary, we detected three novel coding SNPs (two non-synonymous and one synonymous) in the coding part of the *GBP1* exon 2 (residues 1–63), which belongs to the large G domain (residues 1–317) of the *GBP1* protein (Ghosh et al., 2006; Prakash et al., 2000). Although our sequenced coding fragment contains seven GTP/Mg²⁺ binding sites (residues 47–53), both newly discovered non-synonymous mutations (P8Q and A22V) are located at upper positions of this binding domain. Thus, whether the two mutations would have structural and functional effects, needs to be further explored. In addition, we found six

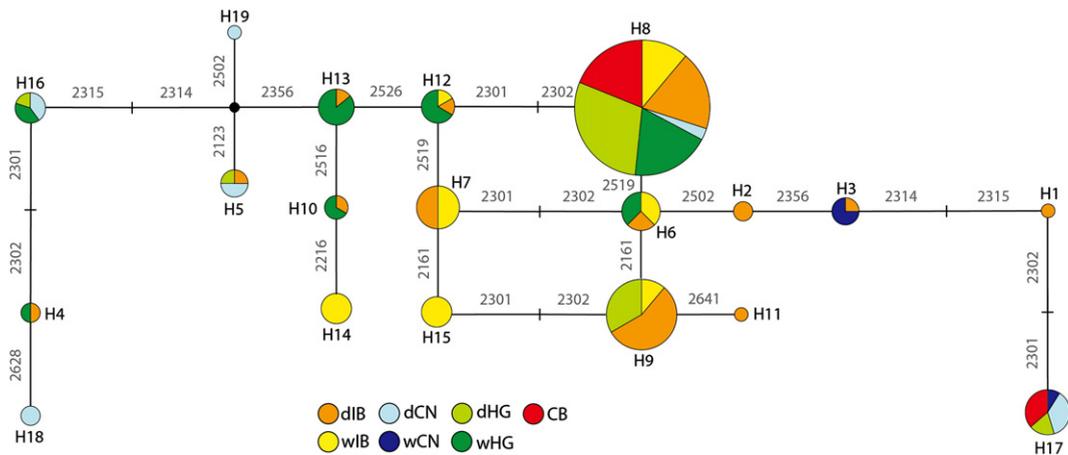


Fig. 1. The median-joining network of the 543-bp *GBP1* haplotypes. The size of each circle is proportional to its frequency in the number of sequences. Abbreviations for sample groups are provided in Table 1.

Table 3
Diversity measures across sample groups.

Sample group	<i>N</i>	<i>S</i>	<i>h</i>	π
dIB	52	12	0.7685	0.0041
wIB	34	7	0.8182	0.0049
dHG	52	10	0.4133	0.0027
wHG	28	8	0.7593	0.0037
dCN	14	10	0.8681	0.0081
wCN	6	7	0.7333	0.0074
CB	24	7	0.2899	0.0037
All	210	14	0.7208	0.0050

N, the number of sequences; *S*, number of polymorphic sites; *h*, haplotype diversity; π , nucleotide diversity per site. Abbreviations of sample groups are provided in Table 1.

haplotypes with frequencies below 1%, but none of them were present in the three commercial breeds (Pietrain, Landrace, and Large White). Although a limited number of individuals and breeds were analyzed, the absence of rare alleles (or haplotypes) in the commercial breeds is an indication that a significant proportion of genetic diversity in domestic species is not present in commercial breeds (Muir et al., 2008). Furthermore, this study demonstrated the potential to find sufficient genetic variation for population genetic analyses of demography versus selection, in functional candidate genes of domestic pigs and wild boars worldwide. This study also highlighted the urgent needs of applying re-sequencing techniques such as next-generation sequencing technology to investigate genome-wide functional genetic variation in wild boars and local domestic pigs worldwide.

Conflict of interest statement

The authors have no conflict of interests regarding the research reported in this manuscript.

Acknowledgments

We thank two anonymous reviewers for their constructive comments. This work was funded by Fundação para a Ciência e Tecnologia (FCT) project PTDC/CVT/68907/2006 (AB-P). SC is supported by a FCT individual fellowship grant SFRH/BPD/46082/2008. We thank Vlado Margeta for his help in sampling Hungary Black Slavonian breed.

References

Anderson, S.L., Carton, J.M., Lou, J., Xing, L., Rubin, B.Y., 1999. Interferon-induced guanylate binding protein-1 (GBP-1) mediates an antiviral effect against vesicular stomatitis virus and encephalomyocarditis virus. *Virology* 256, 8–14.

Bandelt, H.J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48.

Borden, E.C., Sen, G.C., Uze, G., Silverman, R.H., Ransohoff, R.M., Foster, G.R., Stark, G.R., 2007. Interferons at age 50: past, current and future impact on biomedicine. *Nat. Rev. Drug Discov.* 6, 975–990.

Cheng, Y.S., Colonna, R.J., Yin, F.H., 1983. Interferon induction of fibroblast proteins with guanylate binding activity. *J. Biol. Chem.* 258, 7746–7750.

Degradandi, D., Konermann, C., Beuter-Gunia, C., Kresse, A., Würthner, J., Kurig, S., Beer, S., Pfeffer, K., 2007. Extensive characterization of IFN-induced GTPases mGBP1 to mGBP10 involved in host defense. *J. Immunol.* 179, 7729–7740.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567.

Ghosh, A., Praefcke, G.J.K., Renault, L., Wittinghofer, A., Herrmann, C., 2006. How guanylate-binding proteins achieve assembly-stimulated processive cleavage of GTP to GMP. *Nature* 440, 101–104.

Giuffra, E., Kijas, J.M., Amarger, V., Carlborg, O., Jeon, J.T., Andersson, L., 2000. The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* 154, 1785–1791.

Kim, B.-H., Shenoy, A.R., Kumar, P., Das, R., Tiwari, S., MacMicking, J.D., 2011. A family of IFN- γ -inducible 65-kD GTPases protects against bacterial infection. *Science* 332, 717–721.

Kresse, A., Konermann, C., Degrandi, D., Beuter-Gunia, C., Würthner, J., Pfeffer, K., Beer, S., 2008. Analyses of murine GBP homology clusters based on in silico, in vitro and in vivo studies. *BMC Genomics* 9, 158.

Larson, G., Dobney, K., Albarella, U., Fang, M., Matisoo-Smith, E., Robins, J., Lowden, S., Finlayson, H., Brand, T., Willerslev, E., Rowley-Conwy, P., Andersson, L., Cooper, A., 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science* 307, 1618–1621.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452.

Ma, G., Huang, J., Sun, N., Liu, X., Zhu, M., Wu, Z., Zhao, S., 2008. Molecular characterization of the porcine GBP1 and GBP2 genes. *Mol. Immunol.* 45, 2797–2807.

MacMicking, J.D., 2004. IFN-inducible GTPases and immunity to intracellular pathogens. *Trends Immunol.* 25, 601–609.

Martens, S., Howard, J., 2006. The interferon-inducible GTPases. *Annu. Rev. Cell Dev. Biol.* 22, 559–589.

Megens, H.-J., Crooijmans, R.P.M.A., San Cristobal, M., Hui, X., Li, N., Groenen, M.A.M., 2008. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. *Genet. Sel. Evol.* 40, 103–128.

Muir, W.M., Wong, G.K.-S., Zhang, Y., Wang, J., Groenen, M.A.M., Crooijmans, R.P.M.A., Megens, H.-J., Zhang, H., Okimoto, R., Vereijken, A., Jungerius, A., Albers, G.A.A., Lawley, C.T., Delany, M.E., MacEachern, S., Cheng, H.H., 2008. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proc. Natl. Acad. Sci.* 105, 17312–17317.

Ojeda, A., Rozas, J., Folch, J.M., Pérez-Enciso, M., 2006. Unexpected high polymorphism at the FABP4 gene unveils a complex history for pig populations. *Genetics* 174, 2119–2127.

Ojeda, A., Huang, L.-S., Ren, J., Angiolillo, A., Cho, I.-C., Soto, H., Lemús-Flores, C., Makuza, S.M., Folch, J.M., Pérez-Enciso, M., 2008. Selection in the making: a worldwide survey of haplotypic diversity around a causative mutation in porcine IGF2. *Genetics* 178, 1639–1652.

Olszewski, M.A., Gray, J., Vestal, D.J., 2006. In silico genomic analysis of the human and murine Guanylate-Binding Protein (GBP) gene clusters. *J. Interferon Cytokine Res.* 26, 328–352.

Prakash, B., Praefcke, G.J.K., Renault, L., Wittinghofer, A., Herrmann, C., 2000. Structure of human guanylate-binding protein 1 representing a unique class of GTP-binding proteins. *Nature* 403, 567–571.

Schnoor, M., Betanzos, A., Weber, D.A., Parkos, C.A., 2009. Guanylate-binding protein-1 is expressed at tight junctions of intestinal epithelial cells in response to interferon-gamma and regulates barrier function through effects on apoptosis. *Mucosal Immunol.* 2, 33–42.

Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.* 73, 1162–1169.

Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68, 978–989.

Vestal, D.J., Jeyaratnam, J.A., 2010. The guanylate-binding proteins: emerging insights into the biochemical properties and functions of this family of large interferon-induced guanosine triphosphatase. *J. Interferon Cytokine Res.* 31, 89–97.