

## BRIEF COMMUNICATION

# Comprehensive and high-resolution typing of swine leukocyte antigen DQA from genomic DNA and determination of 25 new SLA class II haplotypes

M. T. Le<sup>1,†</sup>, H. Choi<sup>1,†</sup>, M.-K. Choi<sup>1</sup>, D. T. Nguyen<sup>1</sup>, J.-H. Kim<sup>1</sup>, H. G. Seo<sup>1</sup>, S.-Y. Cha<sup>2</sup>, K. Seo<sup>3</sup>, T. Chun<sup>4</sup>, L. B. Schook<sup>5</sup> & C. Park<sup>1</sup>

1 Department of Animal Biotechnology, Konkuk University, Seoul, South Korea

2 College of Veterinary Medicine, Chonbuk National University, Jeonju, South Korea

3 College of Veterinary Medicine, Konkuk University, Seoul, South Korea

4 College of Life Sciences and Biotechnology, Korea University, Seoul, South Korea

5 Department of Animal Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA

## Key words

direct sequencing; haplotype; major histocompatibility complex; polymerase chain reaction; swine leukocyte antigen; SLA-DQA; swine

## Correspondence

Chankyu Park  
Department of Animal Biotechnology  
Konkuk University  
Hwayang-dong  
Kwangjin-gu  
Seoul 143-701  
Korea  
Tel: +82 2 450 3697  
Fax: +82 2 457 8488  
e-mail: chankyu@konkuk.ac.kr

Received 10 April 2012; revised 16

September 2012; accepted 28 September 2012

doi: 10.1111/tan.12017

## Abstract

We previously reported the development of genomic-DNA-based high-resolution genotyping methods for SLA-DQB1 and DRB1. Here, we report the successful typing of SLA-DQA using similar methodological principles. We designed a method for comprehensive genotyping of SLA-DQA using intronic sequence information of SLA-DQA exon 2 that we had obtained from 12 animals with different SLA-DQB1 genotypes. We expanded our typing to 76 selected animals with diverse DQB1 and DRB1 genotypes, 140 random animals from 7 pig breeds, and 3 wild boars. This resulted in the identification of 17 DQA alleles with 49 genotypes. Two new alleles were identified from wild boars. Combine with SLA-DQB1, and DRB1 typing results, we identified 34 SLA class II haplotypes including 25 that were previously unreported.

Swine leukocyte antigen (SLA) genes, the most polymorphic genes of the pig genome, code for molecules that present self and non-self antigens to T cells and trigger specific immune responses, thereby playing a crucial role in the immune system (1). Because they are located in a gene-dense area on chromosome 7 and are closely linked to many other immune response genes (2), SLA genes might also be considered as important markers associated with immune responses (3), disease resistance and susceptibility (4), and performance and production traits (5, 6). We previously reported the development of comprehensive high-resolution typing methods for SLA-DQB1 (7) and DRB1 (8). DQA is an alpha chain component of the hetero-dimeric complex of the SLA-DQ molecule that also has relatively high sequence

diversity, with 20 alleles (<http://www.ebi.ac.uk/ipd/mhc/sla>). Our main purpose was to develop a systemic method for obtaining comprehensive and accurate genotyping results for a major SLA gene, SLA-DQA. This can be achieved by identifying the nucleotide polymorphisms of SLA-DQA exon 2. SLA-DQA exon 2 harbors most of the polymorphisms that are critical for determination of the epitope specificity of SLA-DQA molecules and is consistent with the ISAG Swine Leukocyte Antigen Nomenclature.

The comprehensive amplification of a highly polymorphic region in the genome, such as the major histocompatibility complex (MHC), requires a large amount of preliminary information about the nucleotide variations. This requires information about the sequence variation among at least a few hundred animals. A total of 231 animals from 14 different breeds or genetic backgrounds were used in this study, including 1 American Guinea Hog (AGH), 20 Berkshire pigs, 24

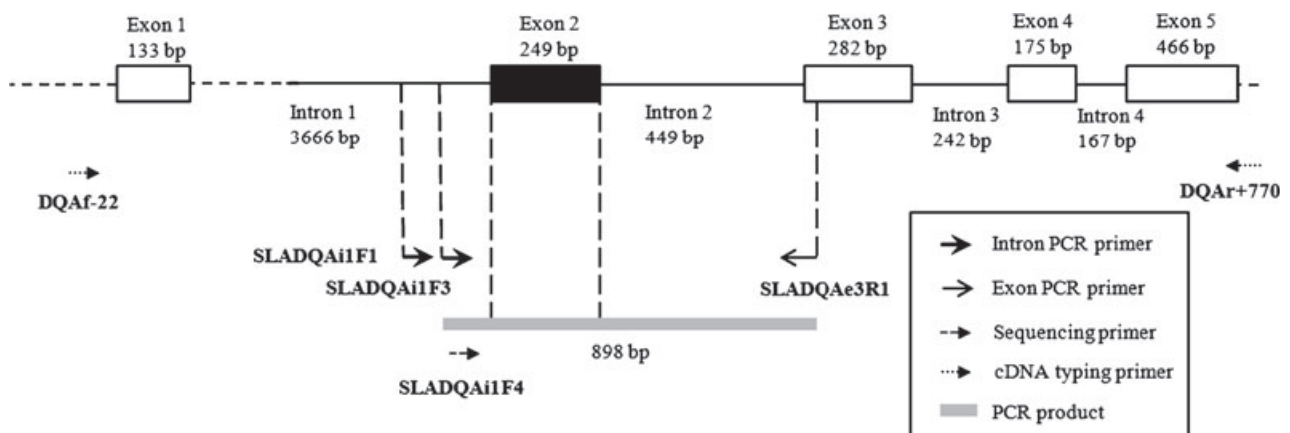
<sup>†</sup>These authors contributed equally to this work.

Duroc pigs, 3 PWG micro-pigs<sup>®</sup> (PWG Genetics, Pyoungtaik, Korea), 26 Korean native pigs (KNP), 24 Landrace pigs, 4 Lanyu pigs, 2 Meishan pigs, 24 National Institutes of Health (NIH) miniature pigs, 5 Ossabaw Island Hogs (OSW) (9), 30 Seoul National University (SNU) miniature pigs derived from Chicago Medical University stock (10), 35 Yorkshire pigs, 3 wild boars and 30 Yorkshire-Landrace crosses. Because SLA-DQB1 is tightly linked to SLA-DQA, we took an advantage of the SLA-DQB1 genotyping information from our previous study to minimize the number of animals required to determine the sequence variations of diverse SLA-DQA alleles (7). We selected 12 animals with 12 different SLA-DQB1 alleles, including 9 homozygous (*DQB1*\*0101, 0201, 0301, 040101, 04sk51, 0501, 0503, 0601 and 0701) and 3 heterozygous samples (0101/0303, 0701/0901 and 0201/C07) that together represented 8 SLA-DQB1 groups in the two-digit nomenclature system (11).

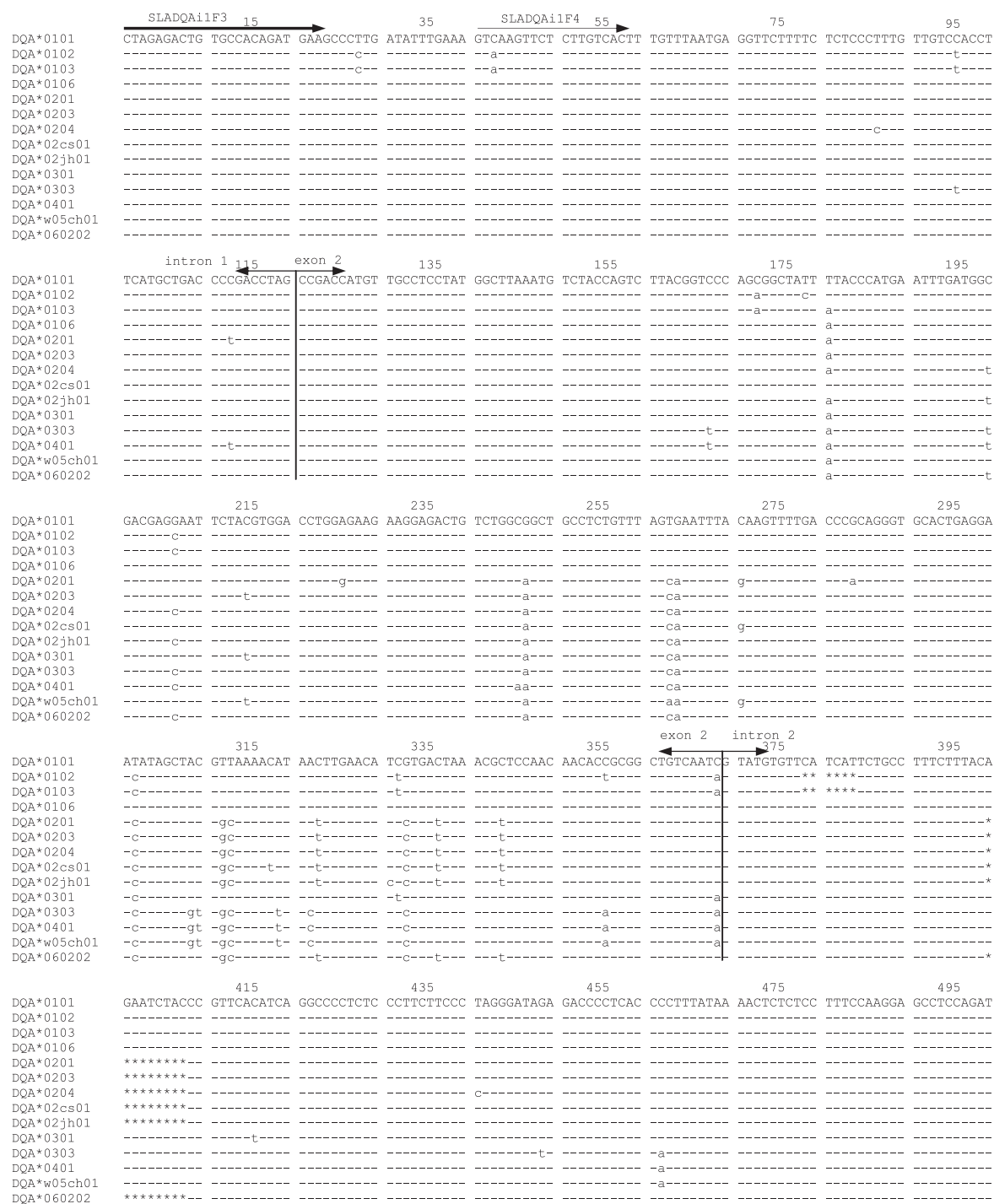
To obtain sequences of SLA-DQA introns 1 and 2, we initially designed several forward primers against the conserved regions between two previously reported genomic sequences (accession numbers AY303988 and BX088590). The reverse primer SLADQAE3R1 (nucleotide position 494–516) was designed against the conserved regions of SLA-DQA exon 3 by comparing 18 full-length cDNA sequences (<http://www.ebi.ac.uk/ipd/mhc/sla>) and taking an advantage of the small size of intron 2 (449 bp; Figures 1, 2). We had limited success in our initial attempt to amplify the intronic sequences around the SLA-DQA exon 2 using the primer set SLADQAI1F1 (5'-GCTGGCAGCCACTGCTCCAA-3') and SLADQAE3R1 (5'-ACAGATGAGGGTGTGGGCTGA-3'). The area we attempted to amplify was a 1099-bp amplicon that contained 334 bp of SLA-DQA intron 1 and 249 bp of complete exon 2 and 448 bp of intron 2 and 68 bp of exon 3 (data not shown). To improve the results, we subjected all amplified fragments to cloning and subsequent sequence analysis to obtain more information about the sequence

diversity at the primer design candidate sites in SLA-DQA introns 1 and 2. After several rounds of the iterative process of designing and testing primers, we developed a universal primer set for the genotyping PCR of SLA-DQA, SLADQAI1F3 (5'-CTAGAGACTGTGCCACAGATGAAG-3') and SLADQAE3R1 that successfully amplified 898 bp of SLA-DQA-specific amplicons from all 12 test animals (Figure S1, Supporting Information) under the following conditions: an initial denaturation at 95°C for 5 min followed by 35 cycles of 1 min at 94°C, 1 min at 58°C, 1 min at 72°C, and a 5-min final extension at 72°C. The sequence information for SLA-DQA partial intron 1, exon 2, intron 2 and partial exon 3 from 14 alleles, including information from 76 additional animals (for a total of 88 genotyped animals), was deposited into Genbank under accession numbers JN859596–JN859605 and JN859592–JN859595 (Figure 3). Therefore, the use of genotype information from a linked gene allowed us to obtain the information necessary to develop a successful genotyping method for SLA-DQA by analyzing only 12 animals.

Although SLADQAI1F3 and SLADQAE3R1 work efficiently for the comprehensive amplification of SLA-DQA, their use as sequence analysis primers for the amplicons did not yield successful results because the presence of frequent nucleotide deletions in intron 2 (nucleotide positions 379–384, 400–408, 632–638, 667, 797–801) produces unreadable chromatograms from SLA-DQA heterozygotes that consist of both deletion and non-deletion alleles and the sequencing results had strong noise peaks in the sequencing chromatograms as a result of the use of the same primers for both amplification and sequencing (Figure 2). To overcome these problems, we designed several sequencing primers. We achieved successful sequencing results for 14 alleles from 36 genotypes using SLADQAI1F4 (5'-GTMAAGTTCTCTTGTCAC-3') (Figures 1, 2; Table S1). SLADQAI1F4 was designed against the nucleotide positions 41–58 with degeneration at nucleotide position 43. The



**Figure 1** The general strategy of genomic sequence-based genotyping for SLA-DQA. The diagram shows the location of each primer for PCR amplification and sequencing. The size (bp) of each intron, exon, and PCR product is indicated.



**Figure 2** Analysis of nucleotide polymorphisms of the SLA-DQA exon 2 and flanking regions for 14 detected alleles. Nucleotide sequences of SLA-DQA exon 2 and flanking regions, including partial and full sequences from intron 1, intron 2 and part of exon 3, were shown. Allele names are indicated in the left-hand column. Identical nucleotides are shown as slashes. Asterisks indicate nucleotide deletion. Exon/intron boundaries are shown by vertical lines with short symmetrical arrows. Thin arrows show the locations of sequencing primers and bold arrows denote the polymerase chain reaction (PCR) primers. For the sequence analysis of cloned inserts, PCR products were gel-purified and ligated into pGEM-T Easy Vector. The ligation products were transformed into DH10B cells. At least five positive colonies were selected from each ligation, and additional colonies were analyzed to resolve discrepancies if necessary. The cloned inserts were bi-directionally sequenced using universal primers T7 (5'-GTAATACGACTCACTATAGGGC-3') and SP6 (5'-ATTTAGGTGACACTATAG-3'). Sequencing results were imported into CLC Workbench (CLC Bio) to assemble sequencing results from the 5' and 3' directions.

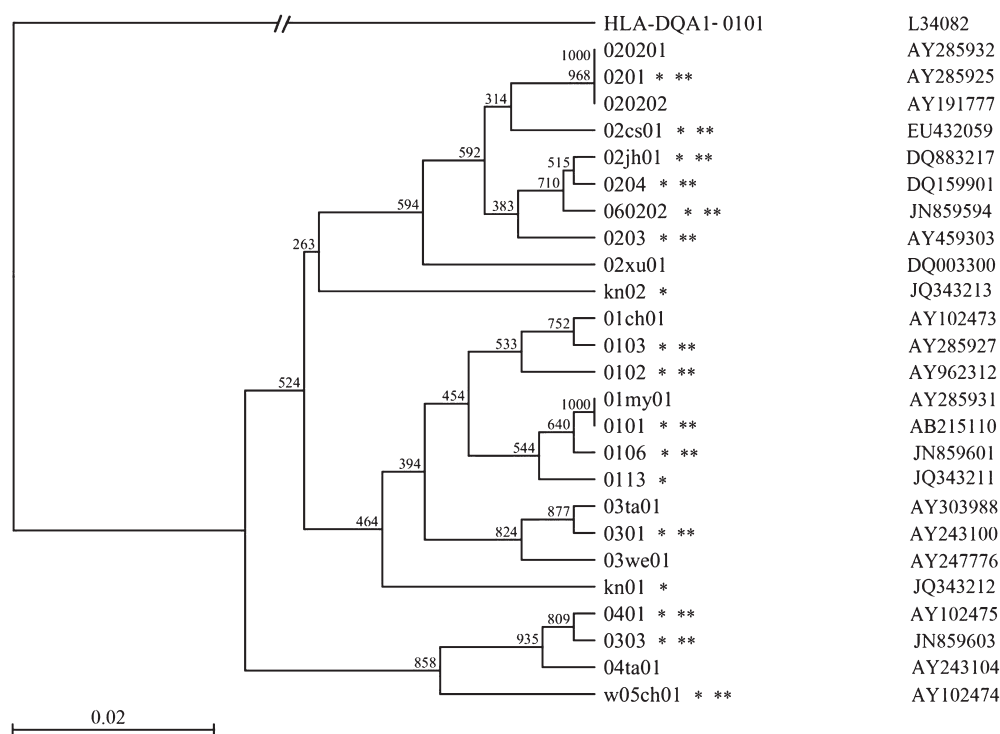
	515	535	555	575	595
DQA*0101	TTTCCCATGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*0102	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*0103	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*0106	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*0201	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*0203	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*0204	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*02cs01	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*02jh01	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*0301	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*0303	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*0401	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*w05ch01	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
DQA*060202	AGATTGCTGG	ACCTTCATCC	TCTCCCGTCT	TACCCATCAC	GTATCTCCAT
	615	635	655	675	695
DQA*0101	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*0102	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*0103	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*0106	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*0201	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*0203	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*0204	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*02cs01	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*02jh01	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*0301	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*0303	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*0401	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*w05ch01	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
DQA*060202	ATTTTGAAT	GAGGTCCCGG	G*****TC	CTCAACAGCC	GGGTCCTCAA
	715	735	755	775	795
DQA*0101	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*0102	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*0103	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*0106	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*0201	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*0203	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*0204	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*02cs01	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*02jh01	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*0301	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*0303	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*0401	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*w05ch01	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
DQA*060202	AGGGAAACAA	GAGCTCTTGC	GCCATCAGAC	TGGGGAACAT	TGGTGGGAGG
	815	intron 2	835	exon 3	855
DQA*0101	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*0102	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*0103	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*0106	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*0201	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*0203	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*0204	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*02cs01	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*02jh01	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*0301	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*0303	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*0401	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*w05ch01	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
DQA*060202	*TTCAAATCA	GTGCTGTTTT	CTCACCACAG	AGGTTCCCTGA	GGTGACTGTG
	875	SLADQae3R1	895		
DQA*0101	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*0102	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*0103	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*0106	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*0201	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*0203	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*0204	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*02cs01	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*02jh01	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*0301	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*0303	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*0401	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*w05ch01	ACTGGGTCAG	CCCAACACCC	TCATCTGT		
DQA*060202	ACTGGGTCAG	CCCAACACCC	TCATCTGT		

Figure 2 Continued

methodology for the direct sequencing of genotyping PCR products was previously described (8). When we obtained sequencing results of new alleles or results that were variations of previously known sequences, independent PCR products were sequenced at least twice at the first occurrence.

To assess the accuracy and reliability of the DQA-genomic sequence-based typing (GSBT) method, we compared the results of SLA-DQA GSBT to those from a cDNA cloning-based method. We isolated both genomic DNA and total RNA from five animals of three breeds. The cDNA-based typing was performed using the previously reported primers, DQAF-22 (5'-ACCTTCCCTTCTGGAGTGTG-3') and DQAR+770 (5'-TGACTACCACTTCTTCGT-3') (12), to obtain 830 bp full-length cDNA amplicons of

SLA-DQA. The genotyping results from five animals showed four different genotypes that consisted of five alleles. There was complete agreement between the two methods (Table S2). We also compared the results of SLA-DQA GSBT to those of a previously reported low-resolution PCR-sequence-specific primer (SSP) for 105 Korean native pigs (13). The results were essentially identical between the two methods, except that our new method was able to further clarify ambiguous or unresolved results of the PCR-SSP typing (Table S3). Another critical issue in the PCR-based genotyping of highly polymorphic loci is the possibility of allele dropout that can occur as a result of preferential amplification of one heterozygote allele. The successful genotyping of 39 heterozygotes (Tables S1, S4 and S5)



**Figure 3** Phylogenetic analysis with 25 SLA-DQA alleles. A phylogenetic tree was constructed using the complete nucleotide sequences of SLA-DQA exon 2 and the UPGMA method. The numbers on the nodes indicate the bootstrap values ( $n=1000$ ), which are shown when greater than 500. '\*\*' indicates alleles that were observed in our genotyping results, '\*\*\*' indicates alleles that were used as reference sequences during SLA-DQA genotyping primer design. The GenBank accession numbers for each allele are shown at the right. HLA-DQA1 was used as an out group. Possibly ambiguous genotype: DQA\*0101 ~ 01my01; DQA\*0201 ~ 020201 ~ 020202.

suggests that there is similar amplification efficiency between the alleles in these heterozygotes. To evaluate the preferential amplification of specific alleles that were not included in the heterozygous genotyping results, we mixed an equal amount of DNA from different SLA-DQA homozygotes, thus generating five different artificial heterozygotes. The typing results were identical to expected genotypes (Figure S2; Table S6). Although experimental limitations prevented us from testing all possible allelic combinations for all known SLA-DQA alleles, our results suggests that the accuracy of our SLA-DQA typing method is comparable to or better than other available methods.

We subjected 88 animals with 17 SLA-DQB1 and 21 DRB1 alleles to SLA-DQA GSBT. We identified four more alleles, DQA\*0103, 0203, 060202 and w05ch01, in addition to the 10 alleles identified during our initial typing Tables S7 and S8). There are 20 SLA-DQA alleles reported in the pig section of the Immuno Polymorphisms Database (IPD) database, arranged in two-digit-based allelic groups. Using our new technique, we identified 14 of these alleles (Figure 3). Yorkshire- and Landrace-crossed animals (F1) had the highest observed allelic diversity with 10 alleles. SLA-DQA\*0201 appears to be the most abundant allele and is present in most of the tested breeds, while DQA\*060202 was identified in

only one Lanyu pig. The distribution of SLA-DQA alleles from animals of different genetic backgrounds is described in Table S8. Because the genotyping results of three SLA class II genes, DQB1, DRB1, and DQA were available for all 88 animals, we compared the allelic diversity among the three loci. The heterozygosity level for SLA-DQA and DRB1 were the same at 0.67, while the level was 0.54 for DQB1 (Table S7).

To assess the genetic diversity of SLA-DQA among different pig breeds in the absence of sample selection, we conducted DQA-GSBT for 20 animals from each of 7 pig breeds for which we had genotyping information (Table S4). The allele frequency and the level of heterozygosity were estimated using the POPGENE software package (<http://www.ualberta.ca/~fyeh/popgene.html>). The pig breeds included Birkshire, Landrace, NIH, Yorkshire, Duroc, SNU, and KNP. We identified 11 alleles (Table 1) with a heterozygosity of 0.65 (Table S9). This result is similar to that from the original 88 pigs. After excluding monomorphic NIH miniature pigs, the heterozygosity was 0.76. The SLA-DQA heterozygosity of Landrace and Yorkshire pigs were very high, at 0.95 and 0.8, respectively. The most frequent allele was SLA-DQA\*0201 (0.364) and the least frequent allele was DQA\*02cs01 (0.011; Table 1). SLA-DQA\*0201 appeared in all analyzed breeds, and DQA\*0101 and DQA\*0303 were

**Table 1** Comparison of the allele frequency of SLA-DQA among seven breeds of pigs using high-resolution genomic sequence-based typing (GSBT)

Alleles	Birkshire	Duroc	Landrace	Yorkshire	NIH	SNU	KNP	All (n=140)
0101 <sup>a</sup>	0.05	0	0.125	0.45	0	0	0.28	0.129
0102	0	0	0	0	0	0.475	0.00	0.068
0106	0.2	0	0.3	0	0	0	0	0.071
0201 <sup>a</sup>	0.2	0.65	0.15	0.3	1	0.125	0.13	0.364
0203	0	0.23	0	0	0	0	0	0.032
0204	0	0.03	0.15	0	0	0	0.15	0.046
02jh01	0	0	0	0	0	0	0.45	0.064
02cs01	0	0	0	0.075	0	0	0.00	0.011
0301	0.425	0	0	0.15	0	0.4	0	0.139
0303	0.075	0.1	0.05	0.025	0	0	0.00	0.036
0401	0.05	0	0.225	0	0	0	0	0.039

KNP, Korean native pigs; NIH, National Institutes of Health; SNU, Seoul National University. Twenty animals of each breed were used.

<sup>a</sup>Possibly ambiguous genotype: DQA\*0101 ~ 01my01; DQA\*0201 ~ 020201 ~ 020202.

the next-most common alleles; each appearing in four breeds (Table 1). SLA-DQA\*0303 appeared only in the four European breeds (Birkshire, Duroc, Landrace, and Yorkshire). SLA-DQA\*0102, \*0203, 02cs01 and \*02jh01 were unique to SNU miniature pigs, Duroc, Yorkshire and KNP, respectively. While this may be due to the limited sample size, these patterns were also observed from the typing results of the 88 selected samples (Table S8). The allelic diversity of SLA-DQA was the highest in Birkshire and Landrace pigs (6 alleles) followed by Yorkshire pigs (5 alleles), Duroc pigs (4 alleles), KNP (4 alleles), SNU miniature pigs (3 alleles), and NIH miniature pigs (monomorphic).

The sequence information for the complete exon 2 satisfies the criteria of the Swine Leukocyte Antigen Nomenclature Committee for allelic discrimination of SLA-DQA. However, sequence comparison among 19 SLA-DQA alleles to full-length cDNA sequences in the IPD database showed that the exon 2 sequences of DQA\*0101 and 01my01, 0201 and 020201 and 020202 were identical. This indicates that, in these cases, polymorphisms outside the DQA exon 2 could not be differentiated using only exon 2. However, such differences may not be functionally critical for determining the antigenic specificity of MHC class II molecules because the peptide binding groove of these molecules are determined primarily by sequence variations in exon 2 (14). We found two new alleles, DQA\*kn01 and DQA\*kn02, from genotyping three wild boars (accession numbers JQ343212 and JQ343213) (Table S5; Figure 3). A phylogenetic analysis showed that DQA\*kn01 was clustered with both DQA\*01 and DQA\*03, and DQA\*kn02 was clustered with DQA\*02 (Figure 3). However, because we only detected each of these alleles in a single individual and were unable to confirm them by cDNA typing due to lack of RNA extractable tissues, further confirmation is needed. We also identified SLA-DQA\*0113 (accession numbers JQ343211) from the typing results of three wild boars, bringing the total number of alleles identified using SLA-DQA GSBT to 17. Although 51 unique SLA-DQA sequences were identified from NCBI searches (Table S10),

only 20 were included in the pig section of the IPD database, which requires validation and assignment of official names (Table S11). Our GSBT method can be used as a useful tool to validate officially unconfirmed alleles.

The combined analysis of SLA-DQA genotyping results with those of SLA-DQB1 and DRB1 allowed us to analyze the MHC class II haplotypes consisting of DQA, DQB1 and DRB1. Although the genotype information of SLA-DQA may not define new haplotypes, it should improve the information content of the SLA class II haplotype. A total of 34 haplotypes were identified in this study including 25 that have not been previously reported (Table 2). Because the genotyping results were obtained from animals of diverse origins, tracing the segregation of all observed haplotypes using family analysis was not always possible. Therefore, we established acceptance criteria for the haplotypes observed in this study. We considered haplotypes confirmed when genotyping results of all three loci, DQB1, DRB1 and DQA, were homozygous and the haplotypes were shared by more than one animal, or when the separation of haplotypes was possible among heterozygotes after removing one known haplotype from more than one animal. The results of haplotype phase analysis were identical between our manual analysis and that of the Pypop program (15) and described together (Table 2). Among 34 haplotypes, 25 were new haplotypes, including 19 confirmed and 6 putative haplotypes. The six haplotypes that were not confirmed were each observed in only one animal. In addition, we extended loci information to SLA-DQA in three previously reported haplotypes; Hp0.11, Hp0.13, and Hp0.25 (Table 2). Hp0.2 (DQB1\*0201/DRB1\*0201/DQA\*0201) was the most common haplotype, with 21 observations in 7 of 12 breeds. This haplotype accounted for 11.93% of all observed haplotypes. Some of the new haplotypes, such as Hp8, Hp17 and Hp21, occurred with high frequency. These results indicate that more effort is required to obtain sufficient haplotype information of SLA class II including the confirmation of the putative haplotypes reported in this study.

**Table 2** Identified high-resolution SLA class II haplotypes

No.	SLA class II haplotype (Hp-)	Locus			No. of observed haplotype												Total (frequency) <sup>b</sup>
		DQB1	DRB1	DQA	AGH <i>r</i> <sup>a</sup> = 1	Du 4	Ka 3	KNP 6	Ld 4	YxL 30	La 4	Me 2	NIH 8	Os 5	SNU 10	Yo 15	
1	0.1 <sup>c</sup>	0101	0101	0101 <sup>d</sup>				2		11							13 (0.07386)
2	0.2	0201	0201	0201 <sup>d</sup>		2				1	5					6	21 (0.11932)
3	0.15a	0201	0401	0203							1					1	1 (0.00568)
4	new <sup>e</sup> Hp 1	0201	0401	0301	1											1	2 (0.01136)
5	new Hp 2	0201	0501	0103										4			4 (0.02273)
6	new Hp 3	0201	0501	0203						2	1					1	7 (0.03977)
7	new Hp 4	0201	0701	0201	1					1						1	3 (0.01705)
8	new Hp 5	0201	0901	0301								3					3 (0.01705)
9	new Hp 6 <sup>f</sup>	0201	kn05	0204		1											1 (0.00568)
10	new Hp 7	0202	0402	0201					1							3	4 (0.02273)
11	new Hp 8	0203	0801	0201							2						2 (0.01136)
12	0.3	0301	0301	0102												13	13 (0.07386)
13	new Hp 9	0301	0401	0301												5	5 (0.02841)
14	new Hp 10	0301	kn05	0204		2											2 (0.01136)
15	0.13 <sup>g</sup>	0303	0403	0204			3		1	1	3					2	10 (0.05682)
16	0.4	040101	0201	020201							2			5			7 (0.03977)
17	new Hp 11	040101	kn04	0201					2								2 (0.01136)
18	new Hp 12 <sup>f</sup>	0402	0201	0301												2	2 (0.01136)
19	0.11 <sup>g</sup>	0402	0901	0301								1					1 (0.00568)
20	new Hp 13	04sk51	1102	02cs01							1					4	5 (0.02841)
21	new Hp 14	0501	0102	0401							5	4					9 (0.05114)
22	new Hp 15	0502	0701	w05ch01			3										3 (0.01705)
23	0.30	0503	1101	02jh01					7								7 (0.03977)
24	new Hp 16	0601	0502	0101							3						3 (0.01705)
25	new Hp 17	0601	1001	0101							12					2	14 (0.07955)
26	new Hp 18 <sup>f</sup>	0601	10jh01	0101							1						1 (0.00568)
27	new Hp 19	0701	0404	0301							2					5	7 (0.03977)
28	new Hp 20 <sup>f</sup>	0701	0501	0103							1						1 (0.00568)
29	new Hp 21	0701	0603Q	0106							9					2	11 (0.06250)
30	new Hp 22 <sup>f</sup>	0701	1102	0106							1						1 (0.00568)
31	new Hp 23	0901	0901	0303		3					1	2					6 (0.03409)
32	0.25 <sup>g</sup>	0901	1301	0303							2						2 (0.01136)
33	new Hp 24	C07	1102	02cs01												2	2 (0.01136)
34	new Hp 25 <sup>f</sup>	Lu02	0201	060202								1					1 (0.00568)

AGH, American Guinea Hog; Du, Duroc; KNP, Korean native pig; Ld, Landrace; La, Lanyu; Me, Meishan; NIH, NIH miniature pig; Os, Ossabaw pig; PWG, PWG micro pig; SNU, Seoul National University miniature pig; Yo, Yorkshire; YxL, Yorkshire cross Landrace.

<sup>a</sup>Number of animals for each breed.

<sup>b</sup>The haplotype phase and frequency were estimated using Pypop program (27).

<sup>c</sup>Reported haplotype from IPD data.

<sup>d</sup>Possibly ambiguous genotype: *DQA\*0101* ~ 01my01; *DQA\*0201* ~ 020201 ~ 020202.

<sup>e</sup>New haplotypes identified in the study.

<sup>f</sup>Putative haplotypes.

<sup>g</sup>Reported haplotype but undefined DQA locus.

Development of methods for MHC typing has significantly improved during last five decades, and there are inherent advantages and disadvantages to each method (16). cDNA sequence based and SSP typing have been more frequently used than other methods. However, the former method requires RNA samples, and the latter one has limited accuracy for unknown alleles. More recently, next generation sequencing (NGS) technologies emerged as a potential method for MHC genotyping. These methods have the advantage of new sequencing technologies that enable the resolution of

individual haplotypes in one sequencing run (17). However, the accuracy of the results, efficiency, and cost effectiveness are still unclear, especially when these methods are applied to species with many unidentified alleles, such as livestock species. The benefits of SLA-GSBT, such as the ability to genotype new alleles with high resolution, no requirement for DNA cloning steps, and the use of genomic DNA as a sample source, may make it the most appropriate method for studies requiring high-resolution genotyping results of SLA-DQA.

## Acknowledgments

This work was supported by a grant from Next-Generation BioGreen 21 (No. PJ009071) Program and the support of 'Cooperative Research Program for Agriculture Science & Technology Development (PJ009103)', the Rural Development Administration, Republic of Korea.

## Conflict of interest

The authors have declared no conflicting interests.

## References

1. Lechler R, Batchelor R, Lombardi G. The relationship between MHC restricted and allospecific T cell recognition. *Immunol Lett* 1991; **29**: 41–50.
2. Smith TP, Rohrer GA, Alexander LJ et al. Directed integration of the physical and genetic linkage maps of swine chromosome 7 reveals that the SLA spans the centromere. *Genome Res* 1995; **5**: 259–71.
3. Mallard BA, Wilkie BN, Kennedy BW. Genetic and other effects on antibody and cell mediated immune response in swine leucocyte antigen (SLA)-defined miniature pigs. *Anim Genet* 1989; **20**: 167–78.
4. Lacey C, Wilkie BN, Kennedy BW, Mallard BA. Genetic and other effects on bacterial phagocytosis and killing by cultured peripheral blood monocytes of SLA-defined miniature pigs. *Anim Genet* 1989; **20**: 371–81.
5. Chardon P, Renard C, Vaiman M. The major histocompatibility complex in swine. *Immunol Rev* 1999; **167**: 179–92.
6. Lunney JK, Ho CS, Wysocki M, Smith DM. Molecular genetics of the swine major histocompatibility complex, the SLA complex. *Dev Comp Immunol* 2009; **33**: 362–74.
7. Park K, Choi H, Thong LM et al. Simple and comprehensive SLA-DQB1 genotyping using genomic PCR and direct sequencing. *Tissue Antigens* 2010; **76**: 301–10.
8. Thong LM, Choi H, Kwon OJ et al. Systematic analysis of swine leukocyte antigen-DRB1 nucleotide polymorphisms using genomic DNA-based high-resolution genotyping and identification of new alleles. *Tissue Antigens* 2011; **77**: 572–83.
9. Swindle MM. *Swine in the Laboratory: Surgery, Anesthesia, Imaging, and Experimental Techniques*, 2nd edn. Boca Raton: CRC Press, 2007.
10. Setcavage TM, Kim YB. Variability of the immunological state of germfree colostrum-deprived Minnesota miniature piglets. *Infect Immun* 1976; **13**: 600–7.
11. Ho CS, Lunney JK, Ando A et al. Nomenclature for factors of the SLA system, update 2008. *Tissue Antigens* 2009; **73**: 307–15.
12. Lee JH, Simond D, Hawthorne WJ et al. Characterization of the swine major histocompatibility complex alleles at eight loci in Westran pigs. *Xenotransplantation* 2005; **12**: 303–7.
13. Cho H-O, Ho C-S, Lee Y-J et al. Establishment of a resource population of SLA haplotype-defined Korean native pigs. *Mol Cells* 2010; **29**: 493–9.
14. Hirsch F, Germana S, Gustafsson K, Pratt K, Sachs DH, Leguern C. Structure and expression of class II alpha genes in miniature swine. *J Immunol* 1992; **149**: 841–6.
15. Lancaster AK, Single RM, Solberg OD, Nelson MP, Thomson G. PyPop update – a software pipeline for large-scale multilocus population genomics. *Tissue Antigens* 2007; **69** (Suppl 1): 192–7.
16. Kuwata S, Maeda H, Juji T. [DNA typing of HLA]. *Rinsho Byori. Jpn J Clin Pathol* 1991; **39**: 1079–86.
17. Erlich RL, Jia X, Anderson S et al. Next-generation sequencing for HLA typing of class I loci. *BMC Genomics* 2011; **12**: 42.

## Supporting Information

The following supporting information is available for this article:

Figure S1. PCR amplification results using SLA-DQA-specific primers for 12 different alleles. An 898-bp segment of the SLA-DQA-specific band was consistently amplified from all alleles tested. Slight nonspecific bands did not affect the sequencing results, as the sequencing primers were different from the PCR primers. Allele names are provided along the top. 'L' indicates a 100 bp DNA ladder.

Figure S2. Example electropherograms from a simulated heterozygote typing. The DNA from SLADQA\* 0101 (top) and \*0401 (middle) homozygotes were pooled, and SLA-DQA genotyping was performed on the pooled DNA (bottom).

Table S1. SLA-DQA genotyping results of 88 pigs from 12 breeds using GSBT method.

Table S2. Comparison of genotyping results between cDNA and gDNA sequence-based typing for SLA-DQA.

Table S3. Comparison of the results from our new method (GSBT) to those from low-resolution typing for SLA-DQA.

Table S4. SLA-DQA genotyping results of 140 pigs from 7 breeds using GSBT typing.

Table S5. Result of genomic sequenced-based SLA-DQA typing of 3 wild boars.

Table S6. Confirmation of genotyping results using pooled DNA from 2 homozygote alleles.

Table S7. Genotyping results and level of heterozygosity for SLA-DQB1, DRB1 and DQA from 88 pigs.

Table S8. Allele distribution of SLA-DQA detected from the typing of 88 pigs from 12 breeds.

Table S9. Differences in SLA-DQA heterozygosity among 7 breeds of pigs.

Table S10. Summary of the number of SLA class II alleles contained in Genbank (as of Jan. 2012).

Table S11. Number of reported alleles for MHC class II genes in the IPD database (as of Feb. 2012).