


NEW ALLELE ALERT

Genomic Full-Length Sequence of the HLA-DQB1*06:47 Allele Was Identified by PacBio Sequencing

Ning Xie¹ | Yi Duan¹ | Yuping Yang¹ | Zhihui Feng² | Xiuxiang Liu¹ ¹NICU, Qingdao Women and Children's Hospital, Qingdao, China | ²HLA Typing Laboratory, Qingdao Blood Center, Qingdao, China**Correspondence:** Xiuxiang Liu (liuxiuxiang99@163.com)**Received:** 9 June 2025 | **Revised:** 11 June 2025 | **Accepted:** 16 June 2025**Funding:** This work was supported by Qingdao Medical and Health Research Project, 2024-WJKY107, 2024-WJKY114.**Keywords:** genomic full length | HLA-DQB1*06:47 | PacBio sequencing

ABSTRACT

The complete genomic sequence of *HLA-DQB1*06:47* was determined using a PacBio sequencing approach conducted in China.

We present the first complete genomic sequencing of the HLA-DQB1*06:47 allele, identified in a regular platelet donor from Qingdao, Shandong Province. The full HLA typing profile of this individual was shown to be: *HLA-A*02:01, 11:01; -B*35:01, 48:01; -C*08:01:01, 15:02; -DRB1*14:05, 15:01:01; -DQB1*05:03, 06:47*. The HLA system, recognised as the most polymorphic in the human genome, currently consists of 2858 identified HLA-DQB1 alleles, according to the IPD-IMGT/HLA Database Release 3.60 (April 9, 2025) [1]. The *HLA-DQB1*06:47* allele was officially named by the WHO Nomenclature Committee for Factors of the HLA System in December 2011. Here, we report the 7026 base pair full-length genomic sequence of *HLA-DQB1*06:47*, obtained using PacBio sequencing, a third-generation sequencing platform. This test was performed with the donor's informed consent.

The variant was initially identified by high-resolution HLA typing using the SBT kit (Genome Diagnostics BV, Utrecht, The Netherlands). Genomic DNA was extracted from peripheral blood leukocytes following the manufacturer's instructions for DNA isolation kits (Life Technologies), and subsequent molecular testing was conducted. Briefly, group-specific amplification was performed using primers targeting the 5' untranslated (UT) and 3' UT regions. This was followed by Sanger sequencing with generic primers in both forward and reverse directions using cycle sequencing [2].

The sequencing results for HLA-DQB1 in this sample did not match any known allele combination. To investigate the cause of this discrepancy, third-generation sequencing (PacBio) was utilised. The sequencing library was loaded onto a chip and sequenced using the PacBio Sequel II high-throughput gene sequencer. The instrument's excitation wavelength prompted the group to emit fluorescence, which was recorded as a fluorescence signal. The resulting data were processed using PacBio's official software, Circular Consensus Sequencing (CCS), to generate HiFi data. Barcodes were then applied, followed by backposting, mutation detection, clustering, typing, and haplotype verification of the results [3]. Finally, the full-length sequence of the *HLA-DQB1*06:47* allele was confirmed.

The full-length sequence of the *HLA-DQB1*06:47* allele, totaling 7026 base pairs, includes all exons, introns, as well as the 5'-upstream and 3'-downstream regions, which are reported here for the first time. When compared to the *DQB1*06:02:01:01* allele, the novel *DQB1*06:47* exhibits a single nucleotide change at position 4598 in exon 3, where C is replaced by A (codon CAC → CAA), as shown in Figure 1. This non-synonymous change results in the substitution of Histidine at position 206 with Glutamine.

The complete nucleotide sequence of this allele has been submitted to the GenBank nucleotide sequence database under

DQB1*06:02:01:01	TGACAGCAAT	TTTCTCTCCC	CTGAAATATG	ATCCCCACCT	AATTTGCTTT	ATTA AAAAATC	CCAAGTATAA	TAACAACCTGG	-369
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TTTTTAACAA	TATTACAGAG	ATGTTTACTG	TGAAATCGCA	TTTTTTTCTT	TGCGCTCTCA	AAATCCCCGA	GGCATTCTGTT	-289
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CTTCAGCTCT	TCTATAATCG	AGAGGAAATT	TTCACCTCAG	ATGTTTCATCC	AGTACAATTT	GAAGACGTCA	CAGTGCCAGG	-209
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CACTGGATTG	AGAACCTTCA	CAAAAAAAT	GTCTGCCAG	AGACAGATGA	GGTCCTTCAG	CTCCAGTGCT	GATTGGTTCT	-129
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TTTCCGAGGG	ACCATCCAAT	CCTGCCACTC	ACGGAACAT	CCACAGGTTT	TTATTCTTTC	TGCCAGGTAC	ATCAGATCCA	-49
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TCAGGTCCGA	GCTGTGTTGA	CTACCACTGC	TTTTCCCTTC	GTCTCAGTTA	Exon1 TGCTCTGGAA	GAAGGCTTTG	CGGATCCCCG	31
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GAGACCTTCG	GGTAGCAACT	GTCACCTTGA	TGCTGGCGAT	GCTGAGCTCC	CTACTGGCTG	AGGGCAGAGA	Intron1 CTCTCCCGGT	111
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	AAGTGCAGGG	CAGCTGCTCT	CCAGAGCCGC	CACTCTGGGA	ACAGGCTCTC	CTTGGGCTGG	GGTATGGGGG	ATGGTGATCT	191
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CTATGATCTT	GGACACAATC	TTTCATCAAC	ATTTCTCTC	TCTTGGGAAA	GAAAGCTATG	TTGCATTCCC	ATTTATCTTT	271
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TAGTGATGAG	ATGAGGACAA	TCCAATCCCC	ATCCTACAGG	CTTAAGCCTG	GAAGAGGAGG	AGAGAGGAGA	GAAAGAGGAG	351
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	ATAAAGTGTA	CATTACTAC	CAGCGATAGG	ACAAAGTGAG	CATGGGGTTA	TTTTTGAAGA	TATGAATTC	TCCAAAGACA	431
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CAGCAGGATT	TGTCATTTAG	GCGTGCCCCA	AGACTGCCTG	GACTTAGGAT	TATGAGATCC	TGCATTGGAA	AATGTGAGGC	511
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	AGCAATGGTG	TCTGTAGTCT	CTGTATTGA	GGAAAATTG	TCTCTATTCC	TGACCGAGTG	GAGCATTGTG	GGAGGCCAAA	591
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TCTTGGTGCT	GAGGGAAGCT	GACTGGCTGA	CCACAGAGAG	GGAGTCTTCG	GGTTTCACTG	ATTTATGGGC	AAATGGTGAC	671
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TTGAGTGGGA	TTCAGAGACC	CGAGTTAATG	GTGGACTGAA	TTAGTAGGA	AGGAGGATGT	AAAGAAGGGA	AATAACACAT	751
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	ATTGTGAAAC	CACTCACTTC	AGACAGAACA	ATACTTTATA	TAACTTCTCT	CTCACTCCTT	CTAACATCCT	GTGTGTAGAT	831
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	ATCATGATTT	TCITTTACAC	AATTATACTT	GTGATATGGA	TATTCTGTTA	CATAAECTGC	CCAAACTGGT	GACTGCCACA	911
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GTTTAATTGG	AATCTAGTTT	ATCAAATCCA	AAAGCTTGTG	CTCTTTCCAT	GAATAAATGT	CTCTATCTAG	GACTCAGAGG	991
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TGTAGGTCCT	TTCCAACATA	GAAGGGAGTG	AACCTCAACG	GGACTTGGGA	GGGTAATATC	AGGCATGGGA	AGGAAGGTAT	1071
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TTTACCAGG	GACCAAGAGA	ATACCGGTGT	CAGAACGAGG	CCAGGCTTAA	TTCTGGACC	TATCTCGTCA	TTCCGTTGAA	1151
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CTCTCAGATT	TATGTGGATA	ACTTTATCTC	TGAGGTATCC	AGGAGCTTCA	TGAAAAATGG	GATTTATGTC	GAGAACGCCC	1231
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TGATCCCCTC	AAGTGCAGAG	GTGCATGTAA	AATCAGCCCG	ACTGCCTCTT	CGCTGGGTTT	ACAGGCTCAG	GCAGGGACAG	1311
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GGTTTCCTC	CCTTTCCTGG	ATGTAGGAAG	GCAGATTCCA	GAAGCCCGCA	AAGAAGCGCG	GCAGAGCTGG	GCAGAGCCGC	1391
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CGGGAGGATC	CCAGGTCTGG	AGCGCCAGGC	ACGGGCGGGC	GGGAECTGGA	GGTCGCGCGG	GCGGTTCCAC	AGCTCCAGGC	1471
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CGGGTCAGGG	CGGCGGCTGC	GGGGCGGCC	GGGCTGGGGC	CTGACTGACC	GGCCGGTGAT	TCCCCGCAGA	GGATTTCGTG	1551
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TTCCAGTTTA	AGGGCATGTG	CTACTTCACC	AACGGGACGG	AGCGCGTGCG	TCTTGTGACC	AGATACATCT	ATAACCGAGA	1631
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GGAGTACGCG	CGTTCGACA	GCGACGTGGG	GGTGTACCGC	GCGGTGACGC	CGCAGGGGCG	GCCTGATGCC	GAGTACTGGA	1711
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	ACAGCCAGAA	GGAAGTCTTG	GAGGGGACCC	GGGCGGAGTT	GGACACGGTG	TGCAGACACA	ACTACGAGGT	GGCGTTCCGC	1791
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GGGATCTTGC	AGAGGAGAGG	Intron2 TGAGCGTCTG	CGCCCTCCG	TGAGGCCAC	CCTTGGCCGA	GACCCCGAGT	TTCTGTGCCG	1871
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GGAGGGGCGA	TGGGGGCGAG	GTCTCTGAAA	TCTGAGCCCA	GTTTCATTCCA	CCCCAGGGGA	AAGGAGGCGG	CGCGGGGGTG	1951
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GTGGGGGCGA	GTGCATCGGA	GGGGCGGGGA	CCTAGGGCAG	AGCAGGGGGA	CAAGCAGAGT	TGGCCAGGCT	GCCTAGTGTC	2031
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	

FIGURE 1 | Alignment of full-length sequences of the *DQB1*06:02:01:01* and *DQB1*06:47* alleles. Dashes indicate identity with the *DQB1*06:02:01:01* sequence. The pipes indicate the exon and intron boundaries. The single nucleotide change between *DQB1*06:02:01:01* and *DQB1*06:47* at position 4598 in exon 3 is shown. The nucleotide positions are numbered according to the genomic full-length alignment in the IPD-IMGT/HLA Database.

DQB1*06:02:01:01	CCCCCAGCC	TCCCCGTCCG	TCGGCCTCGT	CCTCCGCTCT	GCACGTTTCT	CGCCTCGTGC	CTTATGTGTT	TGCCTCTCG	2111
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TACCTTACCT	TCGCTAAGCA	GTTCTCTCTG	CCCCCAGTGC	CCACCCTCTT	CCCCTGCCCC	CGGGCCTCGC	TAGCACTGCC	2191
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CCACCCAGCA	AGGCCACCGT	GCGCAGCTTG	CGCCGAGGA	AGCTTCAGGC	TTGGCCTGGT	GGAGTTAGGG	CTGCCCCACA	2271
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GCTGCGCGCA	GGGCATCCAG	CAATTACAGT	TGCGAAATAA	GATATTTTGA	CTTTTGGCT	TCAAATCATT	ATTCATCGTA	2351
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	ATTCTGTTTT	CTTAAATGGC	TACCACTAAT	GCGGAGATC	TTTGAGGCGA	GAGACTGTTT	AATTATTGCA	TGCCTGGTAC	2431
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CTGACACATT	GACTGGTATG	TGGTGTGAGC	TCAATGATCT	TCTGTAAAT	TCATGAATAA	ATGTACTION	CTGCCAATCC	2511
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	AGTTAGGTTT	AAGAAAAAAA	CAGAGGCTTA	CAATGAACT	TTATCAATTA	TTTCTAACA	TTTTGCTTAA	TGCTTTCAAG	2591
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TAAACTCTTA	CTGACTTGGG	TCTTAATAGA	ATTTGTGAAT	ACAAAGTCTG	AGGAAAAATG	GTTTGCTAAA	AATAAAAAACA	2671
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	ATGCTTGAAT	GATGTTGTAA	GGCAGAGTTT	TAATTTCTTA	AAAAAGCTGA	ACAAATAGCA	CTATGCAAAG	ATCAGAAGTT	2751
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TGGGAATAAA	TAGGTGAAG	CTATTAATTT	ATTGGATAAA	AATAGTTTCA	GGTTGCTTTT	GGCCTAGGTT	ATCCCTCC	2831
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CCCCATCACT	ATCCACTTCA	GGAAATAACA	TTCTGAAAGT	CAATTTTACC	CATTCAGGAA	GCACCTTATTT	CTAGACAGTT	2911
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GCCTTATCAA	ATACCATCTA	TGTTATGTCA	TTAATCTCA	CAGCTACCTG	TGCATTAGAG	ATTAGCATCA	CCACTTTATA	2991
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TATCCTAATA	TTGGTACATG	ATAAACACTT	TAAGTAATCA	GCCCACAATT	ACTCACCAAG	ATCTTAAGCC	TCCCAAAGTA	3071
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CACAACATTC	TTATGTTCTT	CACTACACAT	ATGTAGAGTC	TAAGGGATGT	AAGGCCTTGT	TAAAGCCAGT	TTTGACCAGA	3151
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	AGCAGCAATG	AGTCTATTCC	TGTTTGTCT	GTGTGTTAAT	GCGGCATTGA	AAATTCATGA	TAAATCAATC	CCTAGTCTGA	3231
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CCCCGGTGT	ATCTATGCAG	GTTCAAAAA	CTTTTAGTTT	GCTTTACACT	CCCTTGCTT	CTTTGACTC	ACATCCTAAT	3311
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GCCAGCAAGT	ACTTATATTT	TTGCTATTTC	AGTTTTATTT	CCATAAAATT	TATTTTATCA	TCTTTTCTCA	TAAATTTGTG	3391
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CCCTCTATTT	TACTCCCAG	TCTGTGAAG	ATGAACAAAT	CTTATAAGGC	CACATAGCTG	ACTCTGATTT	CAGTTGGCCT	3471
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CCAGGAAGGA	GAACCTAAAG	AAAAGTTCAA	GTCCAAGCAG	AAACCATGAT	TCCTTCCGGA	TGATGGCTCA	AGAGTGCCAT	3551
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TTAATTGGGG	TGCCACCTGC	TGACCTCAGC	AAATCCCAGT	TATATGTATA	TGTTACACATT	ACAGGCTCAT	TAGCCACAGC	3631
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CGACCTCTGC	ACGGATCTCA	GAATATTTTC	TGTGGAGAAC	ATACACGATA	ATGCTGATT	TCAGAACAAAG	AAAGTAATTTG	3711
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TCAATAGCAA	GGAAATGGAG	TAGGGTAGAC	AGCTAGTAAT	TAAACTCACT	TGTGCGTTAA	AAAGAAATTA	AGGAAAAAAA	3791
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GAAAATGAGA	GAACATATTA	CTAAATAAAG	AAAGCACACA	TAAATATTTT	GCTATAGTTT	CACACTAAGA	GAATAAAGGA	3871
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	AATGTGATAA	AGTGGCTTAA	AAGGAAAGGA	TGAGATGTGT	AAAAGAGGCG	GGGAAAGATG	TGTCATTTTT	TTTACTATGA	3951
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GCAGCAATCT	GAGAAGATAA	AGGAATTGAG	TTGTGGACAA	ACATGATGTT	TGATCAAGGG	TTATTTGTTT	TGAAGGCTG	4031
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CCTACTTTTT	TTTCAAATAT	TACAAACTTT	TGAAATAACA	TTCTACTTTT	TTTTCAAATA	TTACAAACTT	TTGAAATAAC	4111
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	ATTCTTTTTG	TTTTTTGCTG	TCTGTTACTA	GATTGCACAT	TCTATAAATG	CAGGGACCAT	AGTATGTTAT	TTATCTTTGA	4191
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	ATTCTCAGTG	ATTGTCAAAT	TTATATTTGT	TGAATGAACC	TAAATCCAAG	ACTTCGACTC	CAGGTATCTT	TCCACTCTGG	4271
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TCCAAGGAGG	GATCCTTCTT	CATGGTAGAC	GTGGTGTGTG	TTCTCACATC	TCACTCTTAT	GTCTTTTCTT	GTCTGTTACT	4351
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	GCCCTCAGT	Exon3 GAGCCACAG	TGACCATCTC	CCCATCCAGG	ACAGAGGCC	TCAACCACCA	CAACCTGCTG	GTCTGCTCGG	4431
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	TGACAGATTT	CTATCCAGGC	CAGATCAAAG	TCCGGTGTT	TCGGAATGAT	CAGGAGGAGA	CAGCCGGCGT	TGTGTCCACC	4511
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CCCCTTATTA	GGAAATGGTGA	CTGGACTTTC	CAGATCTCTG	TGATGCTGGA	AATGACTCCC	CAGCGTGGAG	ATGTCTACAC	4591
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	
DQB1*06:02:01:01	CTGCCACGTG	GAGCACCCCA	GCCTCCAGAG	CCCCATCACC	GTGGAGTGGC	Intron3 GTAAGGGGAT	ATTGAGTTTC	TGTTACTATG	4671
DQB1*06:47	-----	-----	-----	-----	-----	-----	-----	-----	

FIGURE 1 | (Continued)

DQB1*06:02:01:01 DQB1*06:47	GGCCACAAG	ACAAAGAGCT	CCTTCTGACC	CATTCTTCC	CATCTCTTAT	CCCTGATGTC	ACTACTGAGC	TGGGAATCAC	4751
DQB1*06:02:01:01 DQB1*06:47	AGGAGACTAG	AGCACCTCTT	GCTCCATGGC	AAGTGATCA	GAAGAATCCT	GGTCTCATCA	CCTTCCAGA	TGCTAGGGAA	4831
DQB1*06:02:01:01 DQB1*06:47	ATTACTCTAC	ATACTGTTGC	TCTGAATCCC	AGTCCTGATA	GCTCTGAGGG	ACTGATTCTT	AGGGCTGGTG	ACTGGGATCT	4911
DQB1*06:02:01:01 DQB1*06:47	TAGGGTCTAA	GGTTATGGAT	GAGTTCTTGA	AGAGCAGAGA	TTTGCTTCCC	CACTCTCTCA	CCTATTCACT	GTATCCAAGG	4991
DQB1*06:02:01:01 DQB1*06:47	ACCTATTGGC	TGGTCTTTCC	CCTCCTTAGG	GGTGGTCTGA	ATGGAGAACT	AGGTTCCTTT	GATGCCTTCA	CCTCTGCAC	5071
DQB1*06:02:01:01 DQB1*06:47	CTCAGACTGG	ACTTCAACTC	CTCAGCAGGG	ATGCTATGGG	GTGGGGGAC	AAACACAGAC	ACTCAGTTCT	GCTCTTAGG	5151 Exon4
DQB1*06:02:01:01 DQB1*06:47	GGCTCAGTCT	GAATCTGCC	AGAGCAAGAT	GCTGAGTGGC	GTTGGAGGCT	TCGTGCTGGG	GCTGATCTTC	CTTGGGCTGG	5231
DQB1*06:02:01:01 DQB1*06:47	GCCTTATCAT	CCGTCAAAGG	AGTCAGAAAG	GTGAGGAACC	CCAGGGGAAA	AGGGGAAGAT	GGGCTGTGAC	CCAGACCCTC	5311 Intron4
DQB1*06:02:01:01 DQB1*06:47	TGTTCAGAGT	GGTCCTGTCT	GTAGATTAGC	TCTTCTCTCC	TGACCCTGAG	AAGAAGTGGC	AGGAGACAGG	ACAAGATGGG	5391
DQB1*06:02:01:01 DQB1*06:47	AGGAGGCATT	GGAATCTGAT	TTTACTGGTT	GAAAGGTAGC	GCTGTCACAG	AGCTGACTGA	TAGAGCTTAT	TCCAGGCAT	5471
DQB1*06:02:01:01 DQB1*06:47	CCTTACCATT	CATCATTGTC	TCACTGGCTC	CTTCCAAAA	GCTTCTCCA	TTAAGAGGGT	CAGAGCCTCA	GCTTCTTTTC	5551
DQB1*06:02:01:01 DQB1*06:47	TTTCTAGTGA	CAATTCCTT	TGTTTTAGGG	GATTTTAAAT	TAGGGTGTCTG	AAGGCCTGGA	AGAACATGGG	TGGGAAGAGA	5631
DQB1*06:02:01:01 DQB1*06:47	ATGTAECTCT	AAGTCATGTG	TGTCATTTTC	CTTTGGGGTG	AGAGAGTGGC	TGTTTGTGTA	ATGAGACCTT	TCTCTGCATA	5711
DQB1*06:02:01:01 DQB1*06:47	ACTTCTTTT	GTAAGACCTC	AAGGGCCTCC	ACCAGCAGGT	AATATTTTCAG	CCATGATCCA	GTGTGGGGAG	GCACAGGTAT	5791 Exon5 Intron6
DQB1*06:02:01:01 DQB1*06:47	AAGAGGGAAAG	AGCATGAGCT	GAGTGTACTT	GACCACAGTG	GTCCATGTTC	ATGGCCTATT	TGCTGCTATG	AGGATCAAGA	5871
DQB1*06:02:01:01 DQB1*06:47	CTTAGGGGAG	AAGTTTGCCA	GTTTCTAGGA	ATCTCCAGAC	ATTGTTCCCC	AGAACCAAGC	CTTAACTTTG	GTGGCATCTT	5951
DQB1*06:02:01:01 DQB1*06:47	CTTGTGAAAT	GTGGAGCCAG	AACCACAGCT	TAAATGTTAG	ACACTAGGAT	GATGCCCACT	TTGTGCCACA	TGATGGTGGC	6031
DQB1*06:02:01:01 DQB1*06:47	TACTGCCTGT	AAGCATTTTC	CAGTACTGA	AAGAGGCTGC	TAGTGGTGGG	GAAGAGATAT	CATCCAATTT	CCTAAAAAGA	6111
DQB1*06:02:01:01 DQB1*06:47	CTGAACCCTT	CATATTCCCC	AGAAGAATAA	CAGCTGTTC	CCACCTCCCA	CACATCTGCA	TCAAGCTGAA	GTTCTGTGTC	6191
DQB1*06:02:01:01 DQB1*06:47	CTCATGAGCT	GATTTACCTT	TTGCACAGAT	CTTGGGGGAG	GTGATGACAA	TACACCCCTG	ACCTCAACTT	TCTCTGTCTG	6271
DQB1*06:02:01:01 DQB1*06:47	AAGTGCAGG	GGGCCCTGA	AGGGTGGGG	AGATGGCAGG	CCCACCAGGA	TACCTGTGTC	TGATCAATGC	TCTTCTCTCT	6351
DQB1*06:02:01:01 DQB1*06:47	TCTCCAGGGC	TTCTGCACTG	ACTCCTGAGA	CTATTTTAAAC	TAGGATTGGT	TATCACTCTT	CTGTGATGCC	TGCTTATGCC	6431 Exon6
DQB1*06:02:01:01 DQB1*06:47	TGCCAGAAT	TCCAGCTGC	CTGTGTCAGC	TTGTCCCCT	GAGATCAAAG	TCCTACAGTG	GCTGTCACGC	AGCCACCAGG	6511
DQB1*06:02:01:01 DQB1*06:47	TCATCTCCTT	TCATCCCCAC	CCCAAGGCCT	TGGCTGTGAC	TCTGCTTCTC	GCACTGACCC	AGAGCC		6577

FIGURE 1 | (Continued)

accession number PQ790246, and to the IPD-IMGT/HLA Database under accession number HWS1010052. The name *HLA-DQB1*06:47* was officially assigned by the World Health Organisation (WHO) Nomenclature Committee for Factors of the HLA System in 2011 [4], with the sequence extension submitted in December 2024.

Author Contributions

Ning Xie: performed full-length sequencing test; wrote the paper. **Yi Duan** and **Yuping Yang:** provided the SBT-typing of *HLA-A*, *HLA-B*, *HLA-C*, *HLA-DRB1* and *HLA-DQB1*. **Zhihui Feng:** analysed sequencing data and developed the figure. **Xiuxiang Liu:** got the fund support, designed the test scheme, and reviewed this paper.

Acknowledgements

All the experiments comply with the regulations in China. This work was supported by the Qingdao Medical and Health Research Project (2024-WJKY107) and Qingdao Medical and Health Research Project (2024-WJKY114).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The full-length nucleotide sequence of this allele is available in European Nucleotide Archive GenBank under accession number PQ790246 and in the IPD-IMGT/HLA Database (submission number

HWS10100052). The data that support the findings of this study are openly available in IPD-IMGT/HLA at <https://www.ebi.ac.uk/ipd/imgt/hla/>.

References

1. D. J. Barker, G. Maccari, X. Georgiou, et al., “IPD-IMGT/HLA Database,” *Nucleic Acids Research* 51 (2023): D1053–D1060.
2. C. E. Voorter, F. Palusci, and M. G. Tilanus, “Sequence-Based Typing of HLA: An Improved Group-Specific Full-Length Gene Sequencing Approach,” in *Methods in Molecular Biology*, ed. M. Beksac (Humana Press, 2014), 101–114.
3. N. P. Mayor, J. Robinson, A. J. McWhinnie, et al., “HLA Typing for the Next Generation,” *PLoS One* 10, no. 5 (2015): e0127153, <https://doi.org/10.1371/journal.pone.0127153>.
4. S. G. E. Marsh, E. D. Albert, W. F. Bodmer, et al., “Nomenclature for Factors of the HLA System,” *Tissue Antigens* 75 (2010): 291–455.