Human Immunology 80 (2019) 53–61

Contents lists available at ScienceDirect

Human Immunology

journal homepage: www.elsevier.com/locate/humimm

Computational assessment of miRNA binding to low and high expression HLA-DPB1 allelic sequences

Mengkai Shieh[a], Nilesh Chitnis[a], Peter Clark[a,1], F. Brad Johnson[b], Malek Kamoun[b], Dimitri Monos[a,b,*]

[a] Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, United States
[b] Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, United States

ARTICLE INFO

Keywords:
HLA-DPB1
miRNA
Target prediction
DP expression
Linkage disequilibrium

ABSTRACT

Cell surface expression of HLA-DP is allele specific. SNP rs9277534 (A/G), located in the 3′UTR of the DPB1 gene, has been associated with either low (A) or high (G) expression of DP on the cell surface. Considering the role of miRNAs in the regulation of gene expression, we computationally identified the miRNAs of two BLCLs, PGF and COX, predicted to interact with their corresponding DPB1 transcripts, DPB1 * 04:01:01:01—low expression and DPB1 * 03:01:01:01—high expression. The identified target sequences are located primarily in intron 2 and the 3′UTR. We hypothesize that gene expression may be influenced first by nuclear pre-mRNA events involving intronic regions, followed by the usual 3′UTR–associated events in the cytoplasm. The low DP expression allele was found to interact *in silico* with a larger number of miRNAs than the high expression allele. This pattern holds when examining either the entire transcript unit or simply the polymorphic sites that differentiate the alleles. Interestingly, the rs9277534 A/G polymorphism appears to be in linkage disequilibrium with polymorphisms targeted by the identified miRNAs. The multiplicity of sites targeted by different miRNAs suggests that the expression of DPB1 may be a dynamic process, influenced by different miRNAs under different states of the cell.

1. Introduction

Besides the very well documented and proven influence of structural variations on the functionality of HLA molecules, their functions are also impacted by their level of expression, which appears to be critical for many immunologically relevant physiological and pathologic phenotypes. During the last few years, a number of reports have demonstrated the effects of differential HLA expression in diseases and transplantation. These include the reduced viral load of HIV patients with high expression of HLA-C [1,2], the opposite and deleterious effect in Crohn's disease [2], the adverse outcomes in hematopoietic stem cell transplantation (HSCT) when a patient's mismatched HLA-C allele has a high level of expression [3], the predicted recovery from hepatitis B virus infection depending on HLA-DP expression [4] and the higher risk for graft-versus-host disease (GVHD) among HLA-DPB1-mismatched transplants from donors with low-expression alleles to recipients with high-expression alleles [5]. Such expression differences and their biological consequences are important for developing rational diagnostic and therapeutic solutions for pathologic conditions and can also influence optimal donor selection and potentially expand the pool of acceptable unrelated donors in HSCT. It is therefore very likely that, besides the aforementioned examples, the level of expression of HLAs is relevant to a very broad range of phenotypes involving the immune response, and thus, understanding the specific mechanisms underlying the regulation of expression of the different HLA loci is likely of great importance.

Considering the established role of miRNAs in the regulation of gene expression [6] we focused on assessing the potential influence of miRNAs derived from two chromosome 6–homozygous B lymphoblastoid cell lines (BLCLs), PGF and COX, on the expression of their respective HLA-DPB1 alleles. Characteristically PGF is a low expression and COX is a high expression DP BLCL. Furthermore, the expression of DP has been reported to be associated with a SNP rs9277534 (A/G) polymorphism in the 3′UTR of the DPB1 gene [4], raising the possibility

* Corresponding author at: Abramson Research Center, Rm. 707A, The Children's Hospital of Philadelphia, 3615 Civic Center Blvd, Philadelphia, PA 19104, United States.
E-mail addresses: shiehm@email.chop.edu (M. Shieh), chitnisn@email.chop.edu (N. Chitnis), clarkpe@pennmedicine.upenn.edu (P. Clark), johnsonb@pennmedicine.upenn.edu (F.B. Johnson), malekkam@pennmedicine.upenn.edu (M. Kamoun), monosd@email.chop.edu (D. Monos).
[1] Gene Therapy & Orphan Disease Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104, United States.

https://doi.org/10.1016/j.humimm.2018.09.002
Received 21 July 2018; Received in revised form 27 August 2018; Accepted 12 September 2018
Available online 15 September 2018

that miRNAs may target this or other polymorphic sites in linkage disequilibrium with rs9277534 and by doing so regulate gene expression for DPB1. Our approach involved the computational assessment of targeting by miRNAs encoded by PGF and COX cells on the two respective DPB1 genomic sequences of DPB1 * 04:01:01:01 (PGF) and DPB1 * 03:01:01:01 (COX). Our findings identify specific miRNAs and their respective target sequences on the two DPB1 alleles. The resulting patterns potentially reveal important insights and partially explain the observed differences in expression of the two DP molecules. The results of this computational approach can help us design experiments, whereby direct relationships discovered between specific miRNAs and their targeted sequences on the different DPB1 alleles could help explain differences in DP cell surface expression. Furthermore, this type of knowledge will contribute to our understanding of genomic interrelationships between HLAs and other genomic sequences that are involved in and contribute to disease processes.

2. Materials and methods

2.1. Cell culture and cell surface expression of HLA-DP

Details of the COX and PGF cell cultures have been described previously [7].

Cell surface expression of DP was assessed by flow cytometry [8]. FITC bound anti-human HLA-DP monomorphic monoclonal antibody B7/21 (Leinco Technologies) and anti-CD19-PC5 (Beckman Coulter Inc.) were added to COX or PGF cells (6–10 $\times$ 10^6/ml). Cells were analyzed on a Flow cytometer (Beckman Coulter Cytomics FC500) using TB crossmatch settings.

2.2. RNA extraction and small RNA-seq

Total RNA was extracted from PGF and COX cells using the Qiagen miRNeasy kit (Cat #217084) per manufacturer's protocol. Cell density was 1 $\times$ 10^6/ml, and 5 $\times$ 10^6 cells were collected for RNA extraction from each cell line. RNA was quantified on a Nanodrop ND-100 spectrophotometer, followed by RNA quality assessment on an Agilent 2200 TapeStation (Agilent Technologies, Palo Alto, CA). Library construction, workflow analysis and sequencing runs were performed following standard Illumina TruSeq Small RNA protocol (15004197 Revision G). 50-base-pair single-end reads were generated on the Illumina NextSeq 500 sequencing platform and stored in FASTQ format.

miRNA expression can be impacted by the physiology of the cells. These COX and PGF cells were cultured for some limited period of time under the same conditions. No synchronization of growth or identification of developmental state was attempted. It is expected that each cell within the cultures would be at different growth phases of their cell cycle.

2.3. Identifying miRNAs expressed in PGF and COX cells

FastQC [9] (Version 0.11.5) was run on the raw sequencing data for quality control checks. A patched [10] version of miRDeep2 [11,12] was used to analyze the small RNA-seq data of each cell type separately for miRNA detection and analysis. The GRCh38 primary assembly was combined with the COX MHC alternate locus as the genome reference for miRDeep2 analysis, so as to include the complete MHC sequences of both COX and PGF haplotypes. The following parameters were used for the miRDeep2 data preprocessing script, mapper.pl: -e -h -i -j -k TGG AATTCTCGGGTGCCAAGG -l 18 -m -p hg38_prim + COX_MHC. These parameters specify a FASTQ input file, parsing to FASTA format, conversion from RNA to DNA alphabet, removal of entries containing impermissible characters, clipping of adapter sequence specified in '-k', discarding reads shorter than 18 nt, collapsing of reads, and utilizing the aforementioned GRCh38 "primary assembly + COX_MHC alternate locus" as the genome reference. For running the subsequent miRDeep2.pl wrapper script, the human subset of miRBase [13] release 21 was used as the dataset for known, annotated human miRNAs, and the mouse subset of miRBase 21 was used as the known miRNA dataset of a related species.

The following filters were applied to the miRNA candidates identified by miRDeep2, in order to provide a high-confidence subset for subsequent HLA-DPB1 target prediction. A threshold of at least 78% for the estimated probability of the candidate miRNA being a true positive—as calculated by miRDeep2—was required, as well as removal of all entries overlapping known non-miRNA Rfam [14] elements, followed by the requirement of a RANDFOLD [15] p-value of less than or equal to 0.05. The RANDFOLD p-value reflects the statistical significance of the difference between the minimum free energy of an original RNA sequence and randomly shuffled versions of the sequence, found by Bonnet *et al.* to be a highly discriminating metric for identifying authentic precursor miRNAs, with an analogous effect absent in ribosomal and transfer RNAs. Mature miRNAs with non-zero read count were retained, followed by removal of duplicates having identical sequence, retaining the entry with the maximum read count. A single miRNA sequence can be mapped to multiple genomic loci by miRDeep2; therefore, we randomly choose one of the possible loci of origin to associate with the sequence, while also retaining information for all possible loci of origin, which is included in the final results tables and alignment-annotation diagrams.

2.4. Computational HLA-DPB1 target prediction for identified miRNAs

The subset of miRNAs which passed the aforementioned filters was passed on to RNA22 [16] v2 and TargetScan [17] v7 for target prediction. COX and PGF HLA-DPB1 gene sequences and annotations were obtained from the IPD-IMGT/HLA Database [18,19] (Release 3.31.0). miRNA target prediction was performed using RNA22 on the complete HLA-DPB1 gene sequences and TargetScan on the 3′UTR sequences. The parameters used for RNA22v2 target prediction were as follows: "Sensitivity of 21%, Specificity of 92%," seed size of 7 with no seed mismatches, minimum of 14 base pairs in heteroduplex, minimum heteroduplex energy < -20 kcal/mol, and no G:U wobbles allowed in the seed. RNA22 was selected as a target prediction tool because it was the only known prediction tool designed to handle any genomic region, with others designed for 3′UTR or 3′UTR and CDS regions only. However, given the relatively low performance metrics of an earlier benchmarked version of RNA22 (v1) compared to other tools [20–23], we set the RNA22v2 parameters to be rather stringent, in order to attain higher precision. TargetScan v7 was selected because it is one of the best-performing [21] target prediction tools for 3′UTRs, as well as the only one currently accepting non-miRBase-annotated miRNAs (i.e., in the standalone version, with current versions of DIANA microT and PicTar not accepting custom miRNAs). Targetscan was run using only human data, given the relatively low conservation of the MHC across species, miRNA sequences were represented in all uppercase letters (the program was found to omit parts of the scoring procedure when lowercase letters were used), and a context + + score of -0.7 was used as a threshold for retaining candidates, in order to remain on the higher-precision side of the precision-recall curve [21] ($\sim$90% of candidate targets—93.1% for COX and 89.2% for PGF—were removed using this threshold, partially to compensate for lack of the conservation measure in this analysis). Evaluations of various target prediction tools have indicated that the relatively different classes of tools to which TargetScan and RNA22 belong (i.e., those focused on the seed region, utilizing cross-species conservation filters vs. those characterized by novel analysis of sequence patterns/other features) allow them to find disparate sets of targets, and thus, they may better serve as complementary parts of a combined target prediction approach [22,23], rather than as competing tools. Our analysis therefore employed a complementary target prediction approach.

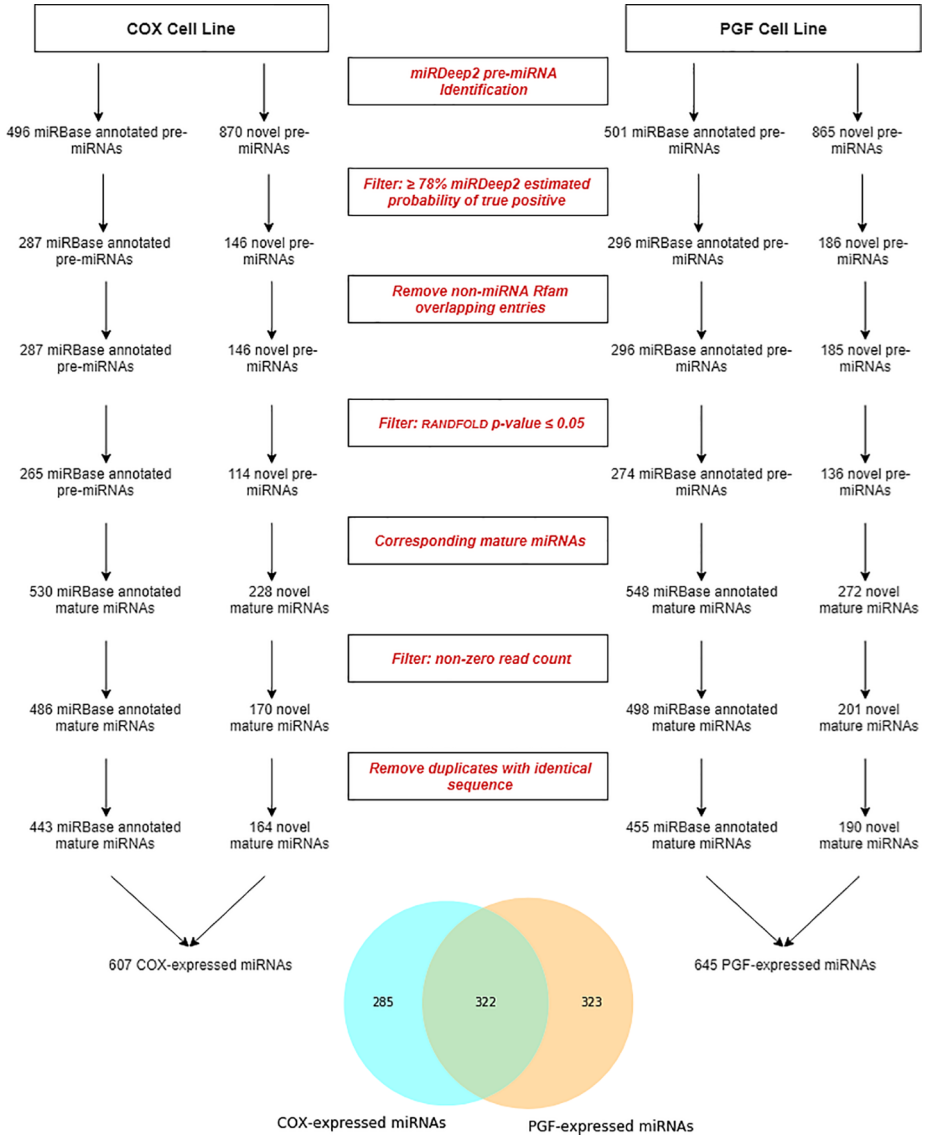

Fig. 1. Flowchart of filtering steps applied to COX and PGF miRNA candidates resulting in sets of high confidence miRNAs for each cell line.

2.5. Alignment-Annotation of miRNAs and predicted targets

The Alignment-Annotator [24] web service was used for multiple sequence alignment of the COX and PGF HLA-DPB1 genes, with mapping of the miRNAs to their predicted targets depicted as stacking annotations located immediately beneath corresponding HLA-DPB1 allele sequences. We have included pop-up bubbles for each miRNA sequence within the alignment-annotator diagrams that include such information as: associated miRBase name, genomic locus of origin name, maximum associated read count for each miRNA sequence, all possible mapped genomic loci of origin, all associated near-duplicate clusters for all possible loci of origin, and all associated targets for the particular miRNA sequence. All such details are also contained in the detailed supplemental tables contained in each supplemental zip archive.

2.6. Identification of near-duplicate miRNAs found within separate runs of miRDeep2

Since the output of the separate miRDeep2 analyses of the COX and PGF samples was found to contain an unknown number of near-duplicates that differed by only 1 or a few nucleotides, we sought to characterize and quantify these near-duplicates within our data. These near-duplicates seem to result from output of slightly different miRNA candidates, which contain differentially predicted borders upon processing of separate samples. These similar miRNA candidates may reflect an identical underlying miRNA across samples or they could represent different isomiRs.

Our procedure for identifying these near-duplicates was as follows ("position" in the following description refers to the central position of the mapped genomic locus of origin for each miRNA, in which we considered all possible genomic loci of origin for each miRNA): we iterated through our sets of miRNAs and compared the position of each miRNA to that of the previously iterated miRNAs, along with the previously identified clusters, as defined next, to identify those miRNAs whose positions fell within a threshold distance of 1 nt from the other miRNAs and clusters. The current miRNA being examined would either be joined to the nearest miRNA to form a cluster or else be added to the nearest cluster (its position being the average of its members' positions), whichever was closer. The list of near-duplicates is available in supplemental file near_duplicates.csv.

Table 1
MHC-encoded miRNAs within our original pool of 930 miRNAs.

miRBase name	Genomic origin name	Read count	Sequence	Source
hsa-miR-219a-1-3p	chr6:33207896-33207917_+	985	AGAGUUGAGUCUGGACGUCCCG	COX miRBase-annotated miRs
hsa-miR-219a-1-5p	chr6:33207855-33207874_+	60	UGAUUGUCCAAACGCAAUUC	COX miRBase-annotated miRs
	chr6:33207857-33207878_-	66	UCGAGAAUUGCGUUUGGACAAU	PGF Novel miRs
	chr6:33207892-33207913_-	20	ACGUCCAGACUCAACUCUCGGC	PGF Novel miRs
	chr6:32621716-32621735_-	44	AAUUUCUGCAUAGUCCACCU	PGF Novel miRs
	chr6:32621749-32621770_-	12	GUGGAUUAUGCAGAAAUUUCUA	PGF Novel miRs
hsa-miR-6891-5p	chr6:31355295-31355316_-	124	GUAAGGAGGGGGAUGAGGGGUC	PGF miRBase-annotated miRs
hsa-miR-6891-3p	chr6:31355229-31355248_-	11	CCCUCAUCUUCCCCUCCUUU	PGF miRBase-annotated miRs

3. Results

3.1. Identifying miRNAs expressed in PGF and COX cells

Deep sequencing of the miRNA transcriptome was performed on two chromosome 6–homozygous B-lymphoblastoid cell lines (BLCLs) that contained opposing rs9277534 SNP variants. PGF contains the A variant, which is associated with low DP expression, whereas COX contains the G variant, which is associated with high DP expression. To confirm the differential cell surface expression of DP between the two cell lines, cell surface expression of DP was assessed using flow cytometry in three different experiments. The relative expression of DP on the two cell lines was on average 1.0 (PGF) : 1.8 (COX), with separate measurements of 1: 1.22, 1: 1.95, and 1: 2.34.

Four technical replicates for each cell type were sequenced, generating a total of eight RNA-seq datasets. No aberrant results were found using FastQC for quality control checks. The reads from each sample were mapped to the GRCh38 reference genome, containing only the primary assembly sequence and the COX MHC alternate locus. The miRDeep2 mapping procedure reported that 56.5% of the raw reads from the PGF samples and 62.3% of the raw reads from the COX samples aligned to the genome reference.

These aligned reads were processed by miRDeep2 to identify novel and miRBase-annotated miRNA transcripts. The miRDeep2 output was aggregated according to cell type, with the application of a series of filters to obtain only the highest confidence candidates for further target prediction. Following the applied filters, the COX-expressed miRNA candidates were reduced from an initial miRDeep2 output of 496 miRBase-annotated miRNA precursors and 870 novel miRNA precursors to a final total of 443 unique miRBase-annotated mature miRNAs and 164 unique novel mature miRNAs (Fig. 1). The PGF-expressed miRNA candidates were reduced from an initial miRDeep2 output of 501 miRBase annotated miRNA precursors and 865 novel miRNA precursors to a final total of 455 unique miRBase-annotated mature miRNAs and 190 unique novel mature miRNAs (Fig. 1). A final total of 607 distinct, high-confidence miRNAs were identified in the COX cells, with 285 (46.9%) of them identified only in the COX cells. A total of 645 distinct, high-confidence miRNAs were identified in the PGF cells, 323 (50.1%) of which were only identified in the PGF cells. The number of these high-confidence miRNAs identified in common in the two cell lines was 322 (Fig. 1).

3.2. Computational HLA-DPB1 target prediction for identified miRNAs

The set of miRNAs expressed in each cell line that passed the filtering process was then used for subsequent target prediction, using the DPB1 * 03:01:01:01 (COX) and HLA-DPB1 * 04:01:01:01 (PGF) sequences as targets. Our target prediction strategy utilized relatively stringent parameter and threshold settings in order to attain heightened precision (positive predictive value). The 5′UTR, intronic, and exonic subsequences were evaluated by RNA22 for target prediction, while 3′UTR predicted targets were the result of both RNA22 and TargetScan predictions (although there were no 3′UTR targets produced by RNA22 using our parameter settings).

Our approach was to assess the theoretical binding of the collection of miRNAs from each cell line to both DPB1 allelic sequences. The rationale for this approach was that the vast majority of miRNAs expressed in each BLCL are not encoded by chromosome 6 and therefore should be independent of the HLA profile/haplotype of the cell. Therefore, in order to avoid any bias, we have chosen to perform the target prediction analysis for each separate set of miRNAs from each BLCL as well as for a unified set of miRNAs, encompassing both BLCLs, against both DPB1 alleles. In addition, we have chosen to present not simply the full set of miRNAs that are predicted to interact with sequences within the entire length of each of the two alleles but also specifically the set of miRNAs that target polymorphic regions that differentiate the two alleles. The intent here was to demonstrate the different patterns of miRNAs that interact only with polymorphic sites, thereby revealing allele specific patterns of miRNA interactions.

Of the 930 (285 + 322 + 323) total distinct, high-confidence miRNAs found to be expressed in the two BLCL cells and used for target prediction (Fig. 1), we found 8 of these miRNAs to be derived from the MHC region, 4 of which represented novel miRNAs (Table 1).

The first of our target analyses examined the set of all COX-expressed miRNAs and assessed predicted targeting of these miRNAs against the DPB1 * 03:01:01:01 (COX) and DPB1 * 04:01:01:01 (PGF) alleles. Forty-two (of 607) COX-expressed miRNAs were found to target the COX HLA-DPB1 * 03:01:01:01 allele, and forty-five (of 607) COX-expressed miRNAs were found to target the PGF HLA-DPB1 * 04:01:01:01 allele. Within these two sets, 32 miRNAs were found to target both HLA-DPB1 alleles (Fig. 2A).

Some miRNAs have multiple targets on a single HLA-DPB1 allele, such that the total number of targets, indicated in the bar plot (Fig. 2B), is larger than the number of miRNAs in the corresponding Venn diagram (Fig. 2A).

The bar plot in Fig. 2B shows the number of COX-expressed miRNA targets in different segments of the DPB1 * 03:01:01:01 (COX) and DPB1 * 04:01:01:01 (PGF) alleles. One miRNA was predicted to interact with the 5′UTR of both DPB1 alleles. Intron 1 of both alleles had 11 predicted targets by a common set of miRNAs. The DPB1 * 03:01:01:01 (COX) allele had 2 additional intron 1 targets which were not present on the DPB1 * 04:01:01:01 (PGF) allele, by miRNAs that targeted specifically the DPB1 * 03:01:01:01 (COX) allele. Intron 2 of the DPB1 * 04:01:01:01 (PGF) allele had 10 predicted miRNA targets, 6 of which were also present at the corresponding site on the DPB1 * 03:01:01:01 (COX) allele and involved the same miRNAs. The DPB1 * 03:01:01:01 (COX) allele had 6 additional intron 2 targets involving miRNAs that targeted the DPB1 * 03:01:01:01 (COX) allele only. The 3′UTR of the DPB1 * 04:01:01:01 (PGF) allele had 27 predicted targets, 16 of which were also present at the corresponding site on the DPB1 * 03:01:01:01 (COX) allele and involved the same miRNAs, and one of which was an additional DPB1 * 04:01:01:01 (PGF) target—not found on DPB1 * 03:01:01:01 (COX)—of a miRNA that targeted both alleles. The 3′UTR of the DPB1 * 03:01:01:01 (COX) allele had 3 additional targets by miRNAs that targeted the DPB1 * 03:01:01:01 (COX) allele only.

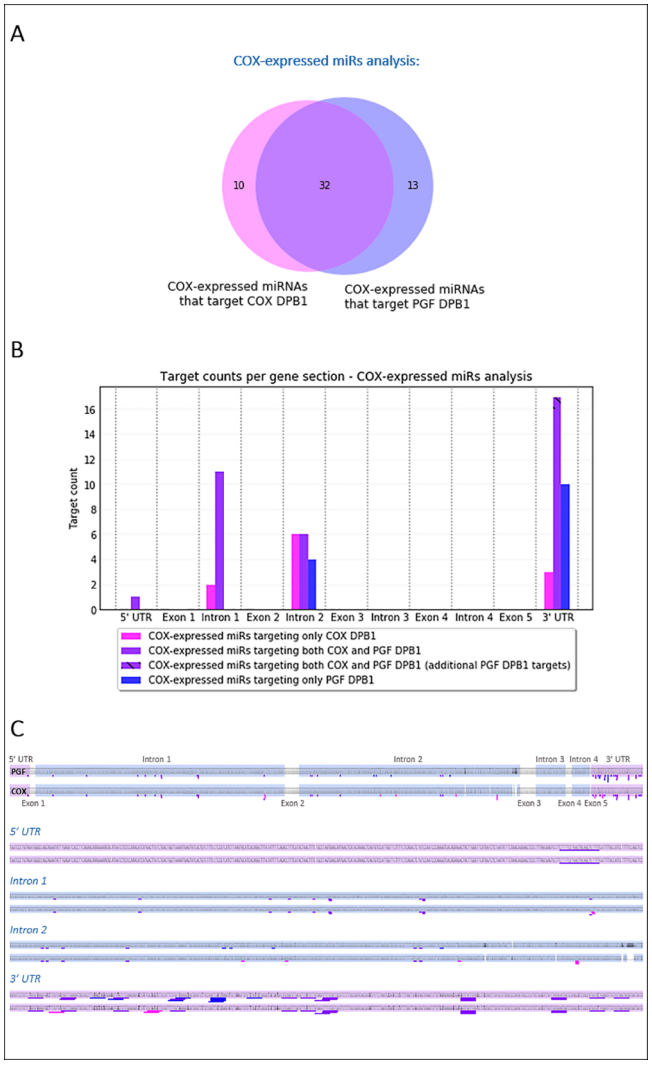

Fig. 2. A: The numbers of COX-expressed miRNAs that have predicted targets on DPB1 * 03:01:01:01 (COX) and DPB1 * 04:01:01:01 (PGF) alleles. B: Bar plot displaying the target counts per gene section on the different alleles. C: Gene representations depicting miRNAs at their target locations on the two DPB1 allelic sequences.

We have also included gene representations (Fig. 2C) which depict exact miRNA targeting along the full gene as well as on individual genetic sub-sections. The full-size, interactive versions of these diagrams are available in supplemental file COX-expressed.zip.

Fig. 3 contains a simplified version of the predicted DPB1 * 03:01:01:01 (COX) and DPB1 * 04:01:01:01 (PGF) targets, miRNAs, and associated miRNA read counts, with color-coding of tables to match the categories in Fig. 2. Full targeting results tables are available in supplemental file COX-expressed.zip.

The following results were obtained when we considered only sites containing a polymorphism that differs between the two DPB1 alleles, denoted as a “polymorphic region” or “PMR” (Fig. 4). A total of 20 of 607 COX-expressed miRNAs targeted the DPB1 * 03:01:01:01 (COX) allele at PMRs. A total of 23 of 607 COX-expressed miRNAs targeted the DPB1 * 04:01:01:01 (PGF) allele at PMRs. Ten of these miRNAs targeted both DPB1 alleles at PMRs (Fig. 4A).

Fig. 4B summarizes the PMR target counts per genetic subsection. Intron 1 of the DPB1 * 03:01:01:01 (COX) allele had 2 predicted PMR targets associated with miRNAs that target PMRs of the DPB1 * 03:01:01:01 (COX) allele only, while the DPB1 * 04:01:01:01 (PGF) allele had no intron 1 PMR targets. Intron 2 of the DPB1 * 04:01:01:01 (PGF) allele had 6 predicted PMR targets, 2 of which were also present on the DPB1 * 03:01:01:01 (COX) allele, associated with the same miRNAs. The DPB1 * 03:01:01:01 (COX) DPB1 allele had 6 additional intron 2 PMR targets associated with miRNAs that targeted PMRs of the DPB1 * 03:01:01:01 (COX) allele only. The 3′UTR of the DPB1 * 04:01:01:01 (PGF) allele had 19 predicted PMR targets, 8 of which were also present on the DPB1 * 03:01:01:01 (COX) allele, associated with the same miRNAs, and one of which was an additional DPB1 * 04:01:01:01 (PGF) PMR target—not found on DPB1 * 03:01:01:01 (COX)—of a miRNA that targeted both alleles at polymorphic sites. The 3′UTR of the DPB1 * 03:01:01:01 (COX) allele had 3 additional PMR targets by miRNAs that targeted the DPB1 * 03:01:01:01 (COX) allele only.

Fig. 4C depicts the exact locations of miRNA PMR targeting on the full gene, as well as on individual genetic sub-sections. The full-size, interactive versions of these diagrams are available in supplemental file COX-expressed.zip.

Fig. 5 contains a simplified version of the predicted DPB1 * 03:01:01:01 (COX) PMR and DPB1 * 04:01:01:01 (PGF) PMR targets, miRNAs, and associated miRNA read counts, with color-coding of tables to match the categories in Fig. 4. Full targeting results tables are available in supplemental file COX-expressed.zip.

The same target prediction procedure (evaluating all targets, as well as only those coinciding with polymorphic sites) was repeated using different subsets or supersets of the originally expressed miRNAs from the two types of cell lines (i.e. PGF, COX intersect PGF, exclusively PGF or COX, PGF plus COX). The results of the target predictions on the DPB1 * 04:01:01:01 (PGF) and DPB1 * 03:01:01:01 (COX) alleles using these different sets of expressed miRNAs can be found in the supplemental materials. The miRNA subsets/supersets utilized for target prediction on the separate DPB1 alleles include the miRNAs (1) expressed in COX cells (Figs. 2–5 and supplemental file COX-expressed.zip), (2) expressed in PGF cells (supplemental file PGF-expressed.zip), (3) expressed in common between PGF and COX cells (intersection set of COX and PGF miRNAs) (Supplemental file common_set.zip), (4) expressed exclusively in COX cells (Supplemental file COX-exclusive.zip), (5) expressed exclusively in PGF cells (Supplemental file PGF-exclusive.zip) and (6) expressed in total in the two cell lines (Union set of COX and PGF miRNAs) (Supplemental file Union_set.zip).

Characteristically, all miRNA–DPB1 allele interactions analyzed demonstrated an increased number of miRNAs interacting with the 3′UTR of DPB1 * 04:01:01:01 (PGF) as compared to DPB1 * 03:01:01:01 (COX) (Fig. 6). This difference supports the notion that the rs9277534 A alleles may simply contain many more actionable miRNA target sites, resulting in greater miRNA-induced suppression. Additionally, when considering only polymorphic site targets, intron 2 retained the majority of targets from among the intronic subsections, suggesting that intron 2 targets may more likely contribute to differential expression if intronic miRNA targets indeed play a role in differential DPB1 expression.

3.3. Alignment-annotation of miRNAs and predicted targets

Included in each supplemental zip archive are the Alignment-Annotator [24,25] diagrams that display the sequence alignments of the two HLA-DPB1 alleles, together with stacking annotations corresponding to miRNAs mapping to their predicted targets, with targeting details viewable in pop-up bubbles or upon right-clicking of targeted sequences. These and further details are also contained in the detailed supplemental tables contained in each supplemental zip archive.

4. Discussion

We have reproduced the reported HLA-DPB1 allele-specific expression of DP dimer molecule on BLCLs [4,5], based on two

Gene section	miRNA	miR sequence	Read count
['Intron 1']	chr13_33189393-33189413_-	CAGGAGGAAAGGGAAGGGACU	32
['Intron 1']	hsa-miR-5090	CGGGGCAGAUUGGUGUAGGGUGCA	20
['Intron 2']	hsa-miR-5090	CGGGGCAGAUUGGUGUAGGGUGCA	20
['Intron 2']	chr1_182391716-182391736_-	CCUCUCCUAACCUCGCUCUCG	1166
['Intron 2']	chr3_45118121-45118142_+	ACAGUACCUGGGUUCUUCUCCU	232
['Intron 2']	hsa-miR-17-5p	CAAAGUGCUUACAGUGCAGGUAG	133829
['Intron 2']	hsa-miR-20b-5p	CAAAGUGCUCAUAGUGCAGGUAG	221
['Intron 2']	hsa-miR-93-5p	CAAAGUGCUGUUCGUGCAGGU	119898
['3' UTR']	hsa-miR-16-3p	CCAGUAUUAACUGUGCUGCUGAA	329
['3' UTR']	chr5_34244924-34244948_+	AGGCUGCAGGUUCCGAGCCCUGCCC	26
['3' UTR']	chr4_1249938-1249955_+	GGAGAGGCCGCACUGGGC	5

Gene section	miRNA	miR sequence	Read count
['Intron 2']	chr6_37561767-37561789_-	UGGGCUGUGUGUCUCUGUUUUGG	15
['Intron 2']	chr5_108748670-108748690_-	GCGGGCCCUGGGAGUGGAGAC	17
['Intron 2']	chr19_5978306-5978326_+	AUCUCCCGCCUCCUUUUCCCG	24
['Intron 2']	hsa-miR-769-5p	UGAGACCUCUGGGUUCUGAGCU	65050
['3' UTR']	hsa-miR-550a-5p	GUGCCUGAGGGAGUAAGAGCC	126
['3' UTR']	hsa-miR-550a-5p	GUGCCUGAGGGAGUAAGAG	10
['3' UTR']	hsa-miR-505-5p	GGGAGCCAGGAAGUAUUGAUGUU	605
['3' UTR']	hsa-miR-3939_star	UUCCUGUAUGUGGGCGUGCA	10
['3' UTR']	hsa-miR-3194-5p	GGCCAGCCACCAGGAGGGCUGC	271
['3' UTR']	chr4_10078620-10078640_-	ACAGUGAGGUAGAGGGAGUGC	32
['3' UTR']	hsa-miR-30e-3p	CUUUCAGUCGGAUGUUUACAGC	54160
['3' UTR']	hsa-miR-30d-3p	CUUUCAGUCAGAUGUUUGCUGC	14501
['3' UTR']	hsa-miR-30a-3p	CUUUCAGUCGGAUGUUUGCAGCU	1429
['3' UTR']	hsa-miR-505-5p	GGGAGCCAGGAAGUAUUGAUGUU	605

Gene section	miRNA	miR sequence	Read count
['5' UTR']	chr10_101902637-101902657_-	CAAAGAGCUGUGGUAGAAAGU	123
['Intron 1']	hsa-miR-6840-3p	GCCCAGGACUUUGUGCGGGGUG	19
['Intron 1']	hsa-miR-6737-5p	UUGGGGUGGUCGGCCCUGGAGGG	6
['Intron 1']	chr17_82594048-82594070_+	CUGCUCCGGCGUCUCUGUCCACA	13
['Intron 1']	chr19_54148127-54148146_+	UCUGCUCUCUCCCACCCGCA	50
['Intron 1']	hsa-miR-3918_star	UCUCCAGCUGGGACCCUGCAC	14
['Intron 1']	chr6_41733589-41733610_-	UGGCUCCCUUCUCUCCGUCUGC	33
['Intron 1']	hsa-miR-1254	AGCCUGGAAGCUGGAGCCUGCAGU	2172
['Intron 1']	chr16_30533050-30533067_-	CUGGGCUGAAGGCUCCCC	11
['Intron 1']	chr15_86879110-86879133_+	AAGAGGAUGUCUGGCAGGUGGCAG	16
['Intron 1']	hsa-miR-3679-5p	UGAGGAUAUGGCAGGGAAGGGGA	119
['Intron 1']	chr17_40122407-40122428_-	GUGGAGGCCGCGGCUGAGGCCC	112
['Intron 2']	chr19_54148127-54148146_+	UCUGCUCUCUCCCACCCGCA	50
['Intron 2']	hsa-miR-4755-3p	AGCCAGGCUCUGAAGGGAAAGU	47
['Intron 2']	hsa-miR-942-5p	UCUUCUCUGUUUUGGCCAUGUGU	6489
['Intron 2']	hsa-miR-3155a_star	CCUCCCACUGCAGAGCCUGGGGA	33
['Intron 2']	chr8_98393702-98393725_-	CCCAGCCUACUGGAGGAUAAGAGG	11
['Intron 2']	hsa-miR-1304-5p	UUUGAGGCUACAGUGAGAUGUG	106
['3' UTR']	hsa-miR-660-3p	ACCUCCUGUGUGCAUGGAUUA	54
['3' UTR']	chr13_33189449-33189470_-	UCCCUUCCCUUUCCUCCUGCUU	10
['3' UTR']	hsa-miR-545-3p	AUCAGCAAACAUUUAUUGUGU	3
['3' UTR']	hsa-miR-3667-3p	ACCUUCCUCUCCAUGGGUCUUU	233
['3' UTR']	hsa-miR-6501-3p	CCAGAGCAGCCUGCGGUAACAGU	12
['3' UTR']	chr10_101902602-101902623_+	ACAAAGAGCUGUGGUAGAAAGU	94
['3' UTR']	chr16_28822850-28822870_-	GGAGAGAAUCAAGUCGGUGAA	27
['3' UTR']	chr20_36050800-36050823_-	CGGAGGCCUCUGCUGUGGCUGCCC	12
['3' UTR']	chr17_40122407-40122428_-	GUGGAGGCCGCGGCUGAGGCCC	112
['3' UTR']	hsa-let-7a-3p	CUAUACAAUCUACUGUCUUUCC	2888
['3' UTR']	hsa-let-7f-3p	CUAUACAAUCUAUUGCCUUCCC	2051
['3' UTR']	hsa-miR-98-3p	CUAUACAACUUACUACUUUCCC	1722
['3' UTR']	hsa-miR-6516-5p	UUUGCAGUAACAGGUGUGAGC	88
['3' UTR']	chr9_97514749-97514770_+	AUUGCAGUCGCUGUGCCAGGAC	22
['3' UTR']	hsa-miR-545-5p	CCUCAGUAAAUGUUUAUUAGAUG	237
['3' UTR']	chr7_92682470-92682491_-	AUCAUUUAUGCUUGCGGAGGAC	17

Gene section	miRNA	miR sequence	Read count
['5' UTR']	chr10_101902637-101902657_-	CAAAGAGCUGUGGUAGAAAGU	123
['Intron 1']	hsa-miR-6840-3p	GCCCAGGACUUUGUGCGGGGUG	19
['Intron 1']	hsa-miR-6737-5p	UUGGGGUGGUCGGCCCUGGAGGG	6
['Intron 1']	chr17_82594048-82594070_+	CUGCUCCGGCGUCUCUGUCCACA	13
['Intron 1']	chr19_54148127-54148146_+	UCUGCUCUCUCCCACCCGCA	50
['Intron 1']	hsa-miR-3918_star	UCUCCAGCUGGGACCCUGCAC	14
['Intron 1']	chr6_41733589-41733610_-	UGGCUCCCUUCUCUCCGUCUGC	33
['Intron 1']	hsa-miR-1254	AGCCUGGAAGCUGGAGCCUGCAGU	2172
['Intron 1']	chr16_30533050-30533067_-	CUGGGCUGAAGGCUCCCC	11
['Intron 1']	chr15_86879110-86879133_+	AAGAGGAUGUCUGGCAGGUGGCAG	16
['Intron 1']	hsa-miR-3679-5p	UGAGGAUAUGGCAGGGAAGGGGA	119
['Intron 1']	chr17_40122407-40122428_-	GUGGAGGCCGCGGCUGAGGCCC	112
['Intron 2']	chr19_54148127-54148146_+	UCUGCUCUCUCCCACCCGCA	50
['Intron 2']	hsa-miR-4755-3p	AGCCAGGCUCUGAAGGGAAAGU	47
['Intron 2']	hsa-miR-942-5p	UCUUCUCUGUUUUGGCCAUGUGU	6489
['Intron 2']	hsa-miR-3155a_star	CCUCCCACUGCAGAGCCUGGGGA	33
['Intron 2']	chr8_98393702-98393725_-	CCCAGCCUACUGGAGGAUAAGAGG	11
['Intron 2']	hsa-miR-1304-5p	UUUGAGGCUACAGUGAGAUGUG	106
['3' UTR']	hsa-miR-660-3p	ACCUCCUGUGUGCAUGGAUUA	54
['3' UTR']	hsa-miR-3667-3p	ACCUUCCUCUCCAUGGGUCUUU	233
['3' UTR']	chr13_33189449-33189470_-	UCCCUUCCCUUUCCUCCUGCUU	10
['3' UTR']	hsa-miR-545-3p	AUCAGCAAACAUUUAUUGUGU	3
['3' UTR']	hsa-miR-3667-3p	ACCUUCCUCUCCAUGGGUCUUU	233
['3' UTR']	hsa-miR-6501-3p	CCAGAGCAGCCUGCGGUAACAGU	12
['3' UTR']	chr10_101902602-101902623_+	ACAAAGAGCUGUGGUAGAAAGU	94
['3' UTR']	chr16_28822850-28822870_-	GGAGAGAAUCAAGUCGGUGAA	27
['3' UTR']	chr20_36050800-36050823_-	CGGAGGCCUCUGCUGUGGCUGCCC	12
['3' UTR']	chr17_40122407-40122428_-	GUGGAGGCCGCGGCUGAGGCCC	112
['3' UTR']	hsa-let-7a-3p	CUAUACAAUCUACUGUCUUUCC	2888
['3' UTR']	hsa-let-7f-3p	CUAUACAAUCUAUUGCCUUCCC	2051
['3' UTR']	hsa-miR-98-3p	CUAUACAACUUACUACUUUCCC	1722
['3' UTR']	hsa-miR-6516-5p	UUUGCAGUAACAGGUGUGAGC	88
['3' UTR']	chr9_97514749-97514770_+	AUUGCAGUCGCUGUGCCAGGAC	22
['3' UTR']	hsa-miR-545-5p	CCUCAGUAAAUGUUUAUUAGAUG	237
['3' UTR']	chr7_92682470-92682491_-	AUCAUUUAUGCUUGCGGAGGAC	17

Fig. 3. Predicted DPB1 * 03:01:01:01 (COX) and DPB1*04:01:01:01 (PGF) targets and their associated miRNAs, with color-coding of tables to match the categories in Fig. 2.

representative MHC haplotypes, PGF and COX, which contain the A and G alleles of rs9277534, respectively. The expression of DP dimer molecules on COX is almost twice as much as the DP expression on PGF cells. We have therefore utilized the two BLCLs to investigate the specifics of the interactions of miRNAs derived from the two cell lines with the genomic sequences of these two DP alleles. This assessment was performed computationally and pertinent target sequences on the two DPB1 alleles have been identified. The underlying assumption is that miRNA binding to DPB1 transcripts influences cell surface expression of DP molecules.

miRNAs can potentially bind and interact at many different sites on each mRNA—or associated hnRNA or pre-mRNA—transcript [26], raising the possibility that the resulting overall miRNA repression of the gene product may be the conglomerate effect of disparate miRNAs interacting with the transcript in its different forms.

PGF contains the rs9277534 A allele, which is associated with lower DP expression [4] and correlates with the greater number of predicted miRNA targets seen in our results, while COX contains the rs9277534 G allele, which is associated with higher DP expression [4] and correlates with the lower number of predicted miRNA targets seen in our results. We have performed an analogous targeting procedure as above, using our COX-PGF union-set miRNAs, as well as miRNAs from a public dataset of 10 normal primary B cells in various level of development [27], on 23 HLA-DPB1 alleles representing all the distinct 3′UTRs of HLA-DPB1 alleles contained in IMGT 3.32.0 that contain a full 3′UTR sequence (see supplemental file "multiple_DPB1_alleles_analysis.zip"). Due to the significant linkage disequilibrium shared within rs9277534 allele classes in the 3′UTR, the DPB1 * 03:01:01:01 and DPB1 * 04:01:01:01 alleles that we have thoroughly interrogated serve as illustrative representatives of their respective rs9277534 allele classes and reflect the general trend of miRNA targeting for each rs9277534 class type. We plan to further analyze these data, with adjustment of various parameters, as well as subgroup analyses of miRNAs from the 10 primary B cells in different levels of development to more fully characterize their differential targeting profiles.

The 3′UTR and CDS are not the only regions of possible miRNA interaction. One Argonaute HITS-CLIP study found 12% of Ago-mRNA HITS-CLIP tags to be located in intronic sequences in mouse brain [28]. In a HITS-CLIP study of human brain, 15% of Ago2 binding sites were found within intronic regions [29], and 8% of Ago2 binding sites were localized to introns in a HITS-CLIPS study of human myocardium [30]. In plants (Arabidopsis thaliana and Oryza sativa), evidence from degradome data supports miRNA-mediated pre-mRNA cleavage at intronic sites, with targeting specificity of transcript isoforms demonstrated, consistent with alternative splicing out of intronic targets [31]. One caveat for all of these studies is that there remains a possibility that the identified interactions could arise from retained introns in the cytoplasm. Despite the questions remaining, what we can state is the following: We observe many predicted targets within our interrogated DPB1 alleles in introns 1 and 2, the presence of which may translate to interactions of miRNAs with these regions when the relevant transcripts are localized in the nucleus; such types of potential nuclear interactions, which are generally understudied but have recently begun to receive increasing attention [32–35], may contribute to resultant levels of corresponding mRNA transcripts. Possible mechanisms include regulation of pre-miRNA stability/turnover or splicing efficiency via intronic and/or DPB1-associated lncRNA targeting.

Even though the miRNA pools that are endogenous to each cell type

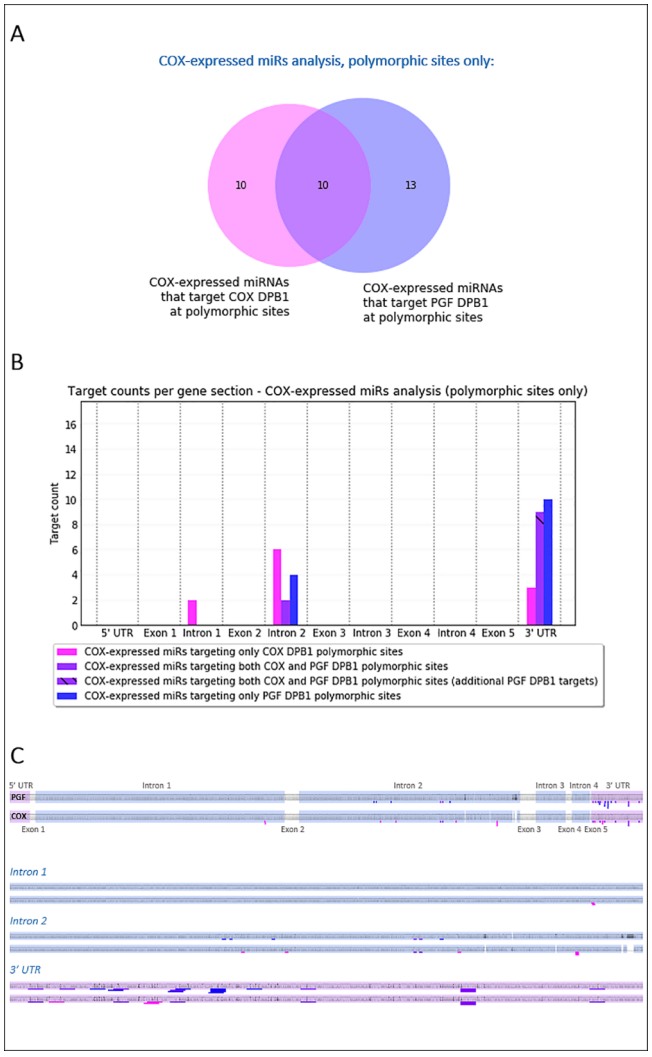

Fig. 4. A: The numbers of COX-expressed miRNAs that have predicted PMR targets on DPB1 * 03:01:01:01 (COX) and DPB1 * 04:01:01:01 (PGF) alleles. B: Bar plot displaying the PMR target counts per gene section on the different alleles. C: Gene representations depicting miRNAs at their PMR target locations on the two DPB1 allelic sequences.

seem to be unique for the cell types tested, we have performed target prediction on each allele using different miRNA pools in order to better understand whether patterns of miRNA interactions differ depending on the miRNAs involved and the target DPB1 allelic sequences. We have also separated each analysis into “all targets” and “polymorphic site only targets” in order to better identify miRNAs that seem to be relevant to particular haplotypes. While the union set of miRNAs and their potential interactions with the two DPB1 alleles can be argued to only represent an imaginary and theoretical scenario, the common set of miRNAs, as being surely present within both cell lines, is a reliable and real set of miRNAs for identifying potential target sequences that differentiate the two alleles.

Some of the miRNAs predicted to differentially target the interrogated DPB1 alleles possess high RNA-seq read counts, demonstrating that these are potentially high-yield candidates to test in further confirmatory biological experiments.

The computational nature of our approach entails that we present our data in a broad and at times hypothesis-agnostic way; since there is still much that is unknown about nuclear miRNA actions and effects, we provide details that may only gain relevance in the future.

Intron 2 binding appears to follow the patterns of nucleotide divergence between rs9277534 A/G allele clades that have been elucidated by Klasberg *et al.* [36]. The first third of intron 2 sequences corresponds to the so-called homogenization zone, which has low divergence between clades and which corresponds to miRNA binding sites that are common between our COX & PGF DPB1 alleles. The next two intron 2 zones, which fall into the transition and differentiation zones [36], are characterized in our analysis by an increasing number of miRNAs that target only one DPB1 allele or the other, which follows the markedly increased inter-clade diversity pattern noted in those two zones [36] . We also notice that the majority of the 3′UTR miRNA binding sites specific to one allele or the other localize to the first half of the 3′UTR, reflective of the higher divergence observed between the clades in that region [36].

We hypothesize that there may be an initial level of allele-specific differential miRNA-associated control of DPB1 transcripts before they leave the nucleus, perhaps relating to intron 2 binding sites; after leaving the nucleus, the well-understood cytoplasmic miRNA effects on the transcripts would then commence, which are known to mostly involve the 3′UTR.

When examining a publically available dataset of miRNAs derived from 10 normal primary B cell samples in various levels of development [27], we found a set of 44 miRNAs that were expressed in common in all these samples as well as in our COX and PGF cells. Targeting results

Gene section	miRNA	miR sequence	Read count
['Intron 1']	chr13_33189393-33189413_-	CAGGAGGAAAGGGAAGGGACU	32
['Intron 1']	hsa-miR-5090	CGGGGCAGAUUGGUGUAGGGUGCA	20
['Intron 2']	hsa-miR-5090	CGGGGCAGAUUGGUGUAGGGUGCA	20
['Intron 2']	chr1_182391716-182391736_-	CCUCUCCUAACCUCGCUCUCG	1166
['Intron 2']	chr3_45118121-45118142_+	ACAGUACCUGGGUUCUUCUCCU	232
['Intron 2']	hsa-miR-17-5p	CAAAGUGCUUACAGUGCAGGUAG	133829
['Intron 2']	hsa-miR-20b-5p	CAAAGUGCUCAUAGUGCAGGUAG	221
['Intron 2']	hsa-miR-93-5p	CAAAGUGCUGUUCGUGCAGGU	119898
['3' UTR"]	hsa-miR-16-3p	CCAGUAUUAACUGUGCUGCUGAA	329
['3' UTR"]	chr5_34244924-34244948_+	AGGCUGCAGGUUCCGAGCCCUGCCC	26
['3' UTR"]	chr4_1249938-1249955_+	GGAGAGGCCGCACUGGGC	5

Gene section	miRNA	miR sequence	Read count
['Intron 2']	chr6_37561767-37561789_-	UGGGCUGUGUGUCUCUGUUUUGG	15
['Intron 2']	chr5_108748670-108748690_-	GCGGGCCCUGGGAGUGGAGAC	17
['Intron 2']	chr19_5978306-5978326_+	AUCUCCCGCCUCCUUUUCCCG	24
['Intron 2']	hsa-miR-769-5p	UGAGACCUCUGGGUUCUGAGCU	65050
['3' UTR"]	hsa-miR-550a-5p	GUGCCUGAGGGAGUAAGAGCC	126
['3' UTR"]	hsa-miR-550a-5p	GUGCCUGAGGGAGUAAGAG	10
['3' UTR"]	hsa-miR-505-5p	GGGAGCCAGGAAGUAUUGAUGUU	605
['3' UTR"]	hsa-miR-3939_star	UUCCUGUAUGUGGGCGUGCA	10
['3' UTR"]	hsa-miR-3194-5p	GGCCAGCCACCAGGAGGGCUGC	271
['3' UTR"]	chr4_10078620-10078640_-	ACAGUGAGGUAGAGGGAGUGC	32
['3' UTR"]	hsa-miR-30e-3p	CUUUCAGUCGGAUGUUUACAGC	54160
['3' UTR"]	hsa-miR-30d-3p	CUUUCAGUCAGAUGUUUGCUGC	14501
['3' UTR"]	hsa-miR-30a-3p	CUUUCAGUCGGAUGUUUGCAGCU	1429
['3' UTR"]	hsa-miR-505-5p	GGGAGCCAGGAAGUAUUGAUGUU	605

Gene section	miRNA	miR sequence	Read count
['Intron 2']	chr8_98393702-98393725_-	CCCAGCCUACUGGAGGAUAAGAGG	11
['Intron 2']	hsa-miR-1304-5p	UUUGAGGCUACAGUGAGAUGUG	106
['3' UTR"]	hsa-miR-660-3p	ACCUCCUGUGUGCAUGGAUUA	54
['3' UTR"]	hsa-miR-545-3p	AUCAGCAAACAUUUAUUGUGU	3
['3' UTR"]	hsa-miR-3667-3p	ACCUUCCUCUCCAUGGGUCUUU	233
['3' UTR"]	chr10_101902602-101902623_+	ACAAAGAGCUGUGGUAGAAAGU	94
['3' UTR"]	hsa-let-7a-3p	CUAUACAAUCUACUGUCUUUCC	2888
['3' UTR"]	hsa-let-7f-3p	CUAUACAAUCUAUUGCCUUCCC	2051
['3' UTR"]	hsa-miR-98-3p	CUAUACAACUUACUACUUUCCC	1722
['3' UTR"]	hsa-miR-545-5p	CCUCAGUAAAUGUUUAUUAGAUG	237

Gene section	miRNA	miR sequence	Read count
['Intron 2']	chr8_98393702-98393725_-	CCCAGCCUACUGGAGGAUAAGAGG	11
['Intron 2']	hsa-miR-1304-5p	UUUGAGGCUACAGUGAGAUGUG	106
['3' UTR"]	hsa-miR-660-3p	ACCUCCUGUGUGCAUGGAUUA	54
['3' UTR"]	hsa-miR-3667-3p	ACCUUCCUCUCCAUGGGUCUUU	233
['3' UTR"]	hsa-miR-545-3p	AUCAGCAAACAUUUAUUGUGU	3
['3' UTR"]	hsa-miR-3667-3p	ACCUUCCUCUCCAUGGGUCUUU	233
['3' UTR"]	chr10_101902602-101902623_+	ACAAAGAGCUGUGGUAGAAAGU	94
['3' UTR"]	hsa-let-7a-3p	CUAUACAAUCUACUGUCUUUCC	2888
['3' UTR"]	hsa-let-7f-3p	CUAUACAAUCUAUUGCCUUCCC	2051
['3' UTR"]	hsa-miR-98-3p	CUAUACAACUUACUACUUUCCC	1722
['3' UTR"]	hsa-miR-545-5p	CCUCAGUAAAUGUUUAUUAGAUG	237

Fig. 5. Predicted DPB1 * 03:01:01:01 (COX) and DPB1 * 04:01:01:01 (PGF) PMR targets and their associated miRNAs, with color-coding of tables to match the categories in Fig. 4.

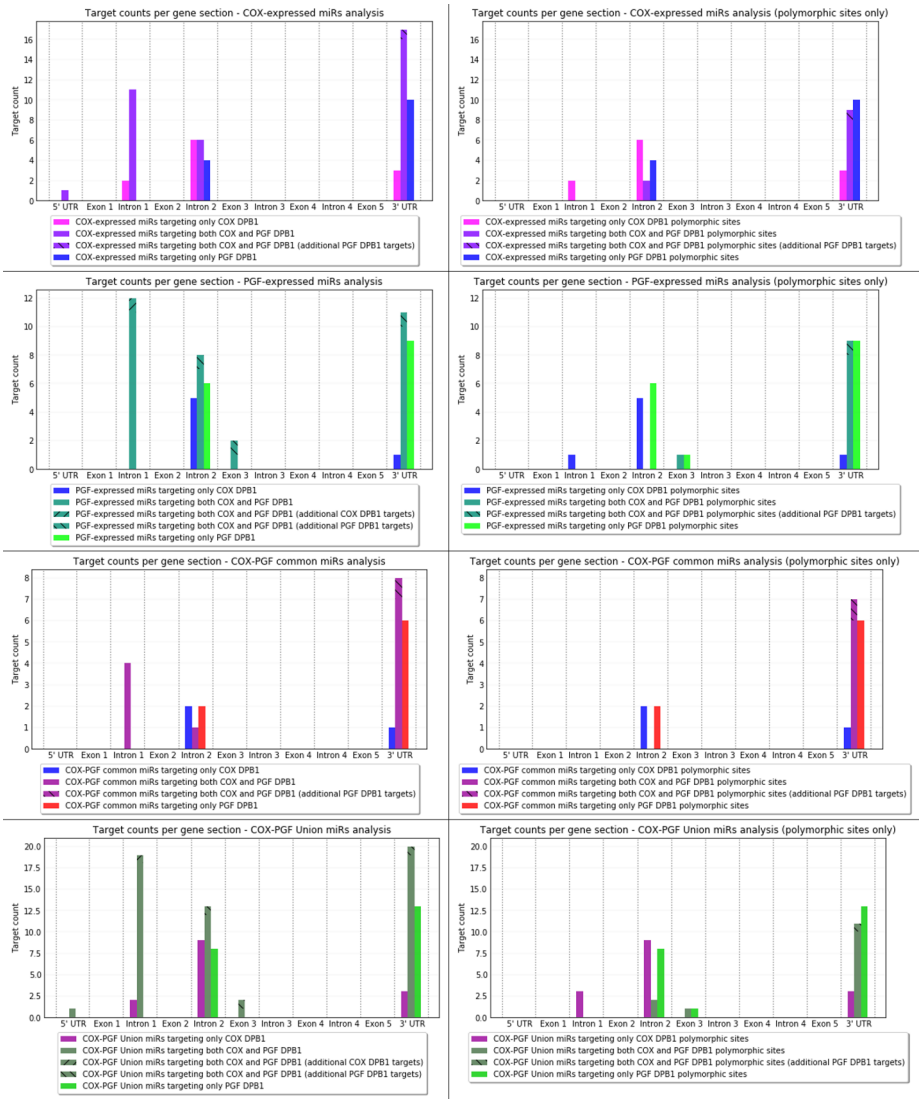

Fig. 6. Bar plots for each miRNA analysis category displaying target counts per gene section for each allele. Bar plots representing "all targets" are displayed on the left, while corresponding PMR-only target bar plots are shown on the right.

from this 44 miRNA subset re-emphasized a top candidate of interest in our results—hsa-miR-30e-3p—which is amply expressed and differentially targets the DPB1 * 04:01:01:01 (PGF) allele while sparing the DPB1 * 03:01:01:01 (COX) allele, even when widely varying levels of different targeting parameters. We have begun efforts to experimentally verify a selection of our *in silico* results, with this particular miRNA as a top candidate for explaining the differential expression between rs9277534 A/G DPB1 alleles.

Computational miRNA target prediction programs can have a high rate of false positive predictions [21]. In order to minimize this weakness, we utilized stringent parameters in our analysis to achieve higher precision/positive predictive value, at the expense of possibly omitting a large number of true predictions. RNA22 is known to underperform when tested head-to-head with other prediction programs [20,21]; however, when considering that RNA22 is in a special class of programs that is able to identify a unique class of miRNAs [22,23], we felt justified in utilizing the program, as long as we utilized stringent parameters. One problem with utilizing stringent parameters with RNA22, however, is that such a process may undermine its very ability to pick up some of the unique miRNA-target pairs in the first place (e.g. it normally has the ability to identify various levels of incomplete seed matches or G-U wobbles within the seed region). We felt that at this stage, it is more important to first identify higher-confidence candidates for experimental verification, whereas in the future, it may prove to be worthwhile to explore further lower-confidence candidates as well.

While this computational assessment addresses the possible influence of the miRNAs in the post-transcriptional control of DPB1 expression in the two BLCLs, we should be mindful that, as previously described [37], there may be inter-individual or inter-allele variability of HLA cell surface expression that is not due to differences of mRNA expression levels but may be due to mechanisms involved in post-transcriptional regulation beyond that of miRNAs. This possibility adds an additional level of complexity not accounted for in the approach we have taken.

Experimental verification of the effect that the identified miRNAs may have on both mRNA levels and cell surface expression will further our understanding of DP expression and most likely of other HLA genes as well.

5. Declarations of interest

The authors declare no conflict of interest.

Acknowledgements

We thank Emmanouil Maragkakis, You Cai, and Kai Wang for their

valuable comments and discussions on portions of the manuscript.

This work was supported by institutional funds from The Children's Hospital of Philadelphia to D.M.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humimm.2018.09.002.

References

[1] R. Thomas, R. Apps, Y. Qi, X. Gao, V. Male, C. O'hUigin, G. O'Connor, D. Ge, J. Fellay, J.N. Martin, J. Margolick, J.J. Goedert, S. Buchbinder, G.D. Kirk, M.P. Martin, A. Telenti, S.G. Deeks, B.D. Walker, D. Goldstein, D.W. McVicar, A. Moffett, M. Carrington, HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C, Nat. Genet. 41 (2009) 1290–1294.

[2] R. Apps, Y. Qi, J.M. Carlson, H. Chen, X. Gao, R. Thomas, Y. Yuki, G.Q. Del Prete, P. Goulder, Z.L. Brumme, C.J. Brumme, M. John, S. Mallal, G. Nelson, R. Bosch, D. Heckerman, J.L. Stein, K.A. Soderberg, M.A. Moody, T.N. Denny, X. Zeng, J. Fang, A. Moffett, J.D. Lifson, J.J. Goedert, S. Buchbinder, G.D. Kirk, J. Fellay, P. McLaren, S.G. Deeks, F. Pereyra, B. Walker, N.L. Michael, A. Weintrob, S. Wolinsky, W. Liao, M. Carrington, Influence of HLA-C expression level on HIV control, Science 340 (2013) 87–91.

[3] E.W. Petersdorf, T.A. Gooley, M. Malkki, A.P. Bacigalupo, A. Cesbron, E. Du Toit, G. Ehninger, T. Egeland, G.F. Fischer, T. Gervais, M.D. Haagenson, M.M. Horowitz, K. Hsu, P. Jindra, A. Madrigal, M. Oudshoorn, O. Ringden, M.L. Schroeder, S.R. Spellman, J.-M. Tiercy, A. Velardi, C.S. Witt, C. O'Huigin, R. Apps, M. Carrington, for the International Histocompatibility Working Group in Hematopoietic Cell Transplantation, HLA-C expression levels define permissible mismatches in hematopoietic cell transplantation, Blood 124 (2014) 3996–4003.

[4] R. Thomas, C.L. Thio, R. Apps, Y. Qi, X. Gao, D. Marti, J.L. Stein, K.A. Soderberg, M.A. Moody, J.J. Goedert, G.D. Kirk, W.K. Hoots, S. Wolinsky, M. Carrington, A novel variant marking HLA-DP Expression levels predicts recovery from hepatitis B virus infection, J. Virol. 86 (2012) 6979–6985.

[5] E.W. Petersdorf, M. Malkki, C. O'hUigin, M. Carrington, T. Gooley, M.D. Haagenson, M.M. Horowitz, S.R. Spellman, T. Wang, P. Stevenson, High HLA-DP expression and graft-versus-host disease, N. Engl. J. Med. 373 (2015) 599–609.

[6] D.P. Bartel, MicroRNAs: target recognition and regulatory functions, Cell 136 (2009) 215–233.

[7] P.M. Clark, N. Chitnis, M. Shieh, M. Kamoun, F.B. Johnson, D. Monos, Novel and haplotype specific micrornas encoded by the major histocompatibility complex, Sci. Rep. 8 (2018), https://doi.org/10.1038/s41598-018-19427-6.

[8] D.S. Monos, E. Czanky, S.J. Ono, S.F. Radka, D. Kappes, J.L. Strominger, L cells expressing DQ molecules of the DR3 and DR4 haplotypes: reactivity patterns with mAbs, Immunogenetics 42 (1995) 172–180.

[9] S. Andrews, FastQC: A Quality Control tool for High Throughput Sequence Data, (n.d.). https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed July 13, 2018).

[10] S. Mackowiak, mirdeep2_patch: Discovering known and novel miRNAs from small RNA sequencing data – updated scripts before release, 2017. https://github.com/Drmirdeep/mirdeep2_patch (accessed March 14, 2018).

[11] M.R. Friedländer, S.D. Mackowiak, N. Li, W. Chen, N. Rajewsky, miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades, Nucleic Acids Res. 40 (2012) 37–52.

[12] mirdeep2: Discovering known and novel miRNAs from small RNA sequencing data, Systems Biology of Gene Regulatory Elements (N. Rajewsky lab), 2018. https://github.com/rajewsky-lab/mirdeep2 (accessed March 14, 2018).

[13] A. Kozomara, S. Griffiths-Jones, miRBase: annotating high confidence microRNAs using deep sequencing data, Nucleic Acids Res. 42 (2014) D68–D73.

[14] I. Kalvari, J. Argasinska, N. Quinones-Olvera, E.P. Nawrocki, E. Rivas, S.R. Eddy, A. Bateman, R.D. Finn, A.I. Petrov, Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families, Nucleic Acids Res. 46 (2018) D335–D342.

[15] E. Bonnet, J. Wuyts, P. Rouze, Y. Van de Peer, Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences, Bioinformatics 20 (2004) 2911–2917.

[16] K.C. Miranda, T. Huynh, Y. Tay, Y.-S. Ang, W.-L. Tam, A.M. Thomson, B. Lim, I. Rigoutsos, A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes, Cell 126 (2006) 1203–1217.

[17] V. Agarwal, G.W. Bell, J.-W. Nam, D.P. Bartel, Predicting effective microRNA target sites in mammalian mRNAs, ELife 4 (2015), https://doi.org/10.7554/eLife.05005.

[18] J. Robinson, A. Malik, P. Parham, J.G. Bodmer, S.G. Marsh, IMGT/HLA database–a sequence database for the human major histocompatibility complex, Tissue Antigens 55 (2000) 280–287.

[19] J. Robinson, J.A. Halliwell, J.D. Hayhurst, P. Flicek, P. Parham, S.G.E. Marsh, The IPD and IMGT/HLA database: allele variant databases, Nucleic Acids Res. 43 (2015) D423–D431.

[20] M. Selbach, B. Schwanhäusser, N. Thierfelder, Z. Fang, R. Khanin, N. Rajewsky, Widespread changes in protein synthesis induced by microRNAs, Nature 455 (2008) 58–63.

[21] P. Alexiou, M. Maragkakis, G.L. Papadopoulos, M. Reczko, A.G. Hatzigeorgiou, Lost in translation: an assessment and perspective for computational microRNA target identification, Bioinformatics 25 (2009) 3049–3055.

[22] A.C. Oliveira, L.A. Bovolenta, P.G. Nachtigall, M.E. Herkenhoff, N. Lemke, D. Pinhal, Combining results from distinct MicroRNA Target prediction tools enhances the performance of analyses, Front. Genet. 8 (2017), https://doi.org/10.3389/fgene.2017.00059.

[23] T.M. Witkos, E. Koscianska, W.J. Krzyzosiak, Practical Aspects of microRNA Target Prediction, (n.d.) 17.

[24] C. Gille, M. Fähling, B. Weyand, T. Wieland, A. Gille, Alignment-Annotator web server: rendering and annotating sequence alignments, Nucleic Acids Res. 42 (2014) W3–W6.

[25] C. Gille, W. Birgit, A. Gille, Sequence alignment visualization in HTML5 without Java, Bioinformatics 30 (2014) 121–122.

[26] S. Wu, S. Huang, J. Ding, Y. Zhao, L. Liang, T. Liu, R. Zhan, X. He, Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3′ untranslated region, Oncogene 29 (2010) 2302–2308.

[27] D.D. Jima, J. Zhang, C. Jacobs, K.L. Richards, C.H. Dunphy, W.W.L. Choi, W. Yan Au, G. Srivastava, M.B. Czader, D.A. Rizzieri, A.S. Lagoo, P.L. Lugar, K.P. Mann, C.R. Flowers, L. Bernal-Mizrachi, K.N. Naresh, A.M. Evens, L.I. Gordon, M. Luftig, D.R. Friedman, J.B. Weinberg, M.A. Thompson, J.I. Gill, Q. Liu, T. How, V. Grubor, Y. Gao, A. Patel, H. Wu, J. Zhu, G.C. Blobe, P.E. Lipsky, A. Chadburn, S.S. Davefor the Hematologic Malignancies Research Consortium (HMRC), Deep sequencing of the small RNA transcriptome of normal and malignant human B cells identifies hundreds of novel microRNAs, Blood 116 (2010) e118–e127.

[28] S.W. Chi, J.B. Zang, A. Mele, R.B. Darnell, Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps, Nature 460 (2009) 479.

[29] R.L. Boudreau, P. Jiang, B.L. Gilmore, R.M. Spengler, R. Tirabassi, J.A. Nelson, C.A. Ross, Y. Xing, B.L. Davidson, Transcriptome-wide discovery of microRNA binding sites in human brain, Neuron 81 (2014) 294–305.

[30] R.M. Spengler, X. Zhang, C. Cheng, J.M. McLendon, J.M. Skeie, F.L. Johnson, B.L. Davidson, R.L. Boudreau, Elucidation of transcriptome-wide microRNA binding sites in human cardiac tissues by Ago2 HITS-CLIP, Nucleic Acids Res. (2016) gkw640.

[31] Y. Meng, C. Shao, X. Ma, H. Wang, Introns targeted by plant microRNAs: a possible novel mechanism of gene regulation, Rice 6 (2013) 8.

[32] C. Catalanotto, C. Cogoni, G. Zardo, MicroRNA in control of gene expression: an overview of nuclear functions, Int. J. Mol. Sci. 17 (2016) 1712.

[33] A.K.L. Leung, The whereabouts of microRNA actions: cytoplasm and beyond, Trends Cell Biol. 25 (2015) 601–610.

[34] S. Pitchiaya, L.A. Heinicke, J.I. Park, E.L. Cameron, N.G. Walter, Resolving subcellular miRNA trafficking and turnover at single-molecule resolution, Cell Rep. 19 (2017) 630–642.

[35] J.E.J. Rasko, J.J.-L. Wong, Nuclear microRNAs in normal hemopoiesis and cancer, J. Hematol. Oncol. 10 (2017), https://doi.org/10.1186/s13045-016-0375-x.

[36] S. Klasberg, K. Lang, M. Günther, G. Schober, C. Massalski, A.H. Schmidt, V. Lange, G. Schöfl, Patterns of non-ARD variation in more than 300 full-length HLA-DPB1 alleles, Hum. Immunol. (2018), https://doi.org/10.1016/j.humimm.2018.05.006.

[37] C. René, C. Lozano, M. Villalba, J.-F. Eliaou, 5′ and 3′ untranslated regions contribute to the differential expression of specific HLA-A alleles: Molecular immunology, Eur. J. Immunol. 45 (2015) 3454–3463.