CHARLES UNIVERSITY
THIRD FACULTY OF MEDICINE

Summary of the Doctoral Thesis

Regulation of HLA class II genes expression

Marta Zajacová

Prague, 2018

Doctoral study programmes in Biomedicine

Charles University and Czech Academy of Sciences

Discipline: Molecular and Cell Biology, Genetics and Virology

Head of the Division Council: Prof. RNDr. Stanislav Zadražil, DrSc.

Place of work: Department of medical genetics, Third Faculty of Medicine, Charles University, Prague

Author: Marta Zajacová, M.Sc.

Supervisor: Prof. MUDr. Marie Černá, DrSc

Examiners: prof. RNDr. Jan Černý, Ph.D., doc. MUDr. Antonij Slavčev, CSc.

Date of distribution of the Summary of the Ph.D. thesis: 11. 5. 2018

Thesis defense is held on 26. 6. 2018 at 15:00 at the Department of Genetics and Microbiology of the Faculty of Science, Charles University (Viničná 5, Prague 2).

The full text of the Ph.D. thesis is available at the dean's office at the Third Faculty of Medicine, Charles University.

Contents

Shrnutí

Geny HLA II. třídy jsou vysoce polymorfní, a to také v regulačních nekódujících genových oblastech. Polymorfismus v promotorové oblasti vytváří silný potenciál pro rozdílnou expresi jednotlivých alel. I když je známo, že množství molekul HLA II. třídy na povrchu buněk má významnou úlohu při utváření imunitní odpovědi, polymorfismus v expresi genů HLA třídy II dosud nebyl podrobně analyzován.

Cílem této práce bylo zkoumat expresi mRNA a metylaci promotorové DNA u alel genů HLA II. třídy. Byly provedeny dvě studie, které se zabývaly různými aspekty regulace.

V **části A** byla zkoumána metylace DNA u 10 alel promotoru *DQA1* a její vliv na expresi mRNA tohoto genu. DNA metylace v plné krvi byla stanovena pomocí bisulfitového sekvenování a exprese mRNA byla měřena za použití RT-qPCR. I když byly pozorovány mezialelické rozdíly v celkové metylaci (nejvíce metylované alely byly *DQA1*02:01* a **04:01*), nebyla pozorována očekávaná negativní korelace mezi methylační hustotou DNA promotoru *DQA1* a expresí alely. Mezialelické rozdíly v metylaci jednotlivých CpG pozic prokázány nebyly, ale domníváme se, že genetický polymorfismus v oblasti (zejména oblast upstream od pozice -400, která je téměř kompletně metylovaná ve všech alelách) může vést k různým interpretacím 5meCpG na stejné pozici v různých alelických kontextech.

V **části B** byla analyzována úroveň exprese mRNA alel genů *HLA-DQA1* a *-DQB1* v plné krvi a B lymfocytech a monocytech pomocí RT-qPCR. Transkripce genů II. třídy byla vyšší (s výjimkou alel *DQB1*06*) a vykazovala nižší mezialelickou variaci v B lymfocytech ve srovnání s monocyty. Hierarchii exprese *DQB1* lze zobecnit vztahem *DQB1** 06 > **03*, **05* > **02* v monocytech a *DQB1*06* > **02*, **03*, **05* v B lymfocytech. Hierarchie exprese *DQA1* je *DQA1*03* > **01* ≥ **02* > **05* v monocytech a *DQA1 *03*, **05* > **01*, **02* v B lymfocytech. Vzhledem k nízkému počtu vzorků pro určité kombinace typu alel a buněk byly pouze některé z těchto výsledků statisticky významné. Isoformy CIITA jsou diskutovány jako faktory, které mohou být odpovědné za rozdílnou expresi alel HLA II. třídy v specifických buněčných typech. Nakonec jsme se pokusili objasnit vztah mezi expresí DQ dimerů a rizikem rozvoje autoimunitního diabetu.

Závěrem, tato práce potvrzuje a popisuje polymorfismus v expresi mRNA genů *HLA-DQA1* a *-DQB1* a stanovuje expresní hierarchii těchto alel v B lymfocytech, monocytech a plné krvi. Přináší profil metylace DNA promotoru genu *DQA1* a ukazuje, že metylace DNA není sama o sobě schopna vysvětlit mezialelické rozdíly v expresi. V našem zjednodušeném modelu není úroveň povrchové exprese DQ dimerů sama o sobě schopna vysvětlit jejich asociaci s T1D.

Summary

HLA class II genes are known to be highly polymorphic, even in the regulatory non-coding gene regions. Polymorphism in the promoter region potentially forms a strong basis for an uneven allele-specific expression. Even though it is known that the amount of HLA class II molecules on the cell surface has a significant role in shaping immune response, HLA class II expression polymorphism has not yet been thoroughly measured.

The thesis aims to shed light into allele-specific mRNA expression and promoter DNA methylation of HLA class II genes. Two studies, each addressing different aspects of the HLA class II allele expression regulation, were conducted.

Study A examines the DNA methylation of 10 *DQA1* promoter and its effect on the *DQA1* mRNA expression. DNA methylation in whole blood cells was determined with bisulfite sequencing and mRNA expression was measured using RT-qPCR. Even though inter-allelic differences in overall methylation were observed (the most methylated alleles were *DQA1*02:01* and **04:01*), the expected negative correlation between the *DQA1* promoter DNA methylation density and the allele expression was not observed. We suggest that the genetic polymorphism in the region (especially region upstream of position -400, which is almost completely methylated in all alleles) may lead to different interpretation of the same 5meCpG mark in different allelic contexts.

Study B analyses *DQA1* and *DQB1* mRNA expression in whole blood cells, B lymphocytes and monocytes using RT-qPCR. Class II transcription level is higher (with the exception of *DQB1*06* alleles) and shows lower interallelic variation in B cells compared to monocytes. The *DQB1* expression hierarchy can be generalized into *DQB1*06 > *03*,**05 > *02* pattern in monocytes and *DQB1*06 > *02*, **03*, **05* in B cells. The *DQA1* expression hierarchy is *DQA1*03 > *01 ≥ *02 > *05* in monocytes and *DQA1*03*, **05 > *01, *02* in B cells. Because of the low number of samples for certain allele-cell type combinations, only some of these results were statistically significant. CIITA isoforms are discussed as the factors that could drive the cell-type specific expression of class II alleles. Finally, we tried to relate DQ dimer expression level to the risk it carries for the development of the autoimmune diabetes.

In conclusion, this thesis confirms and describes mRNA expression polymorphism of the *HLA-DQA1* and *-DQB1* genes, and provides the expression hierarchy of these alleles in the B cells, monocytes and whole blood cells. It presents a DNA methylation profile of the *DQA1* gene promoter and shows that DNA methylation by itself is not able to explain the inter-allelic expression differences. In our simplified model of DQ dimer expression, the expression level by itself is not able to explain dimer association with T1D.

1. Introduction to Regulation of HLA Class II Genes and Their Role in Immunity

Variability in the protein structure and expression levels of HLA class II molecules is an important factor influencing the target and form of the immune response in different individuals within a human population. This thesis focuses on studying the differences in expression of alleles of HLA-*DQA1* and -*DQB1* genes and tries to identify the reasons and consequences of their expression variability.

1.1. HLA class II properties and function

1.1.1. HLA class II structure and isotypes

HLA class II molecules are dimers composed of non-covalently associated transmembrane chains α and β. Extracellular parts of protein chains together make up an antigen-binding site and short cytoplasmic tails serve for outside-in signalling (1). Three isotypes of class II molecules can form on a cell surface, HLA-DR (formed by molecules encoded by the *DRA* and *DRB1, 3, 4* and *DRB5* genes), HLA-DQ (*DQA1, DQB1*) and HLA-DP (*DPA1, DQB1*) (2). HLA-DM and -DO non-classical isotypes are not expressed on the cell surface, and instead they assist with intracellular antigen loading onto classical class II molecules (3).

1.1.2. HLA class II polymorphism

HLA class II genes exhibit an exceptionally high level of polymorphism. Tens to thousands variants are known for each of classical class II genes (with the exception of *DRA*). Variation of genes encoding for DM and DO molecules is also very limited (4).

Genetic variance affects coding region where influences the spectrum of presented peptides, and non-coding regions of the gene where it can affect efficiency of an allele transcription. Together, variation on both levels creates unique combination of properties for each allele.

1.1.3. HLA class II properties – pMHC-TCR interaction quality and quantity

Function of MHC class II molecules is to bind fragments of primarily exogenous (but also endogenous (5)) proteins and expose them on the cell surface where they can be recognized by CD4+ T cells. Specific recognition of MHC-peptide complexes by the antigen-specific T cell receptors (TCR) enables an adaptive immune response targeted against the eliciting antigen as well as establishing and maintaining the self-tolerance (6, 7).

The outcome of an encounter between an antigen presenting cell (APC) and T cell (whether the T cell will respond to the antigen and how the response will look like) is determined by several factors, like cytokines in the environment, presence of the costimulatory molecules on the APC and most importantly, by a signal delivered through TCR after interaction with a cognate pMHC complex. Both **quality (affinity) of TCR-pMHC interaction** and a **cumulative quantity of the TCR signal** are important.

1.1.3.1. pMHC-TCR affinity

A major (but not the only) function of the pMHC-TCR interaction affinity is to promote IL-2 secretion from naive cells (7). In addition, cells carrying more affine pMHC ligands form longer lasting contacts with dendritic cells (DCs), express higher levels of the IL-2 receptor chains CD25 and CD122 and produce more IL-2 upon restimulation (8).

Affinity of the pMHC-TCR interaction is relatively low, and bonds with halftimes of several seconds are sufficient for generating an immune response (9). Optimal *in vivo* T cell

effector responses (measured as production of IL-2 and IFNγ) are to ligands with intermediate TCR-pMHC half-lives (9).

The pMHC-TCR interaction is not only low affinity but also degenerate: that is, many TCRs recognize the same pMHC complex and many pMHC complexes are recognized by the same TCR (reviewed by (10)). As a result, T cells are extremely cross-reactive: individual T cell clones are able to recognize more than a 10^6 different peptides in the context of a single MHC molecule (10). This brings a potential for autoimmunity: T cells sensitivity towards an antigen can increase up to 50-fold during their development ((10) and next chapter) and T cells activated by peptides from pathogens can cross-recognize and get activated by self-derived peptides with an affinity for TCR much lower that that of the original pathogen-derived peptide (10).

1.1.3.2. pMHC-TCR affinity and dose

A pMHC-TCR interaction of a certain minimal strength (determined by both, interaction quality and quantity) is needed for a T cell survival after an antigen encounter in periphery. T cells stimulated with a weaker signal are able to secrete IL-2 (and thus can provide support to the T cells with more specific TCRs), but fail to produce lineage-specific cytokines (IL-4 and IFNγ), commit to Th1 or Th2 lineage and do not persist in the organism (7).

A comprehensive study of a T cell differentiation into IFNγ-producing effector cells in response to a range of pMHC concentrations (10^6 range) and pMHC-TCR affinities (10^6 range) showed that for all affinities, the cytokine response to increasing pMHC concentration is bell-shaped with the reduced cytokine production at low and high pMHC concentrations (11). The pMHC-TCR affinity has no effect amplitude of the response, but is inversely correlated to the pMHC concentration needed to reach the maximal response (11). Similar observations were done in study on the peripheral induction of Foxp3 regulatory T cells: for each peptide, there is an optimal concentration (peak of the bell-shaped dose-response) for the Foxp3 induction, which is inversely correlated with the pMHC-TCR affinity (12). However, in this case, the affinity and quantity of the signal were not interchangeable. The presence of TGFβ extends the TCR signal strength range over which a Foxp3 induction occurs (12).

Quantity and affinity of pMHC-TCR interactions also influences negative selection and induction of natural FoxP3+ regulatory T cells in thymus (6, 13, 14).

1.1.3.3. pMHC-TCR dose

Variation of the antigen dose and the number of MHC II molecules on the APC surface leads to preferential development of cells with different effector phenotypes due to influence on the cytokines production and Th1–Th2 equilibrium of effector cells. The effect of a particular antigen dose is not consistent in all studies, however, generally it seems that both very low and very high level favors production of IL-4 and Th2 responses, and intermediate levels favor IFNγ and Th1 development (15–19). High antigen dose is also needed for development of inflammatory Th17 cells. If naïve T cells encounter a low-dose antigen in different cytokine environments, Th17 effectors do not differentiate at antigen doses that are sufficient to drive differentiation of Th1 and Th2 effectors (20). On the other hand, very high antigen doses can drive development of Th17 cells even in the absence of exogenous polarizing cytokines (19).

The amount of antigen present during the T cell development from naïve into effector cell also sets the **avidity setpoint**—the threshold of antigen concentration to which the cell is able to respond on a secondary encounter (20).

Low MHCII expression was associated with autoimmunity through a less effective presentation of autoantigens to the protective regulatory T lymphocytes (21), or inefficient

deletion of autoreactive cells in thymus through the similar mechanism as it happens with low-stability pMHC complexes (22).

Thus, pMHC amount on the cell surface is an important regulator of the immune response. Final pMHC amount is a result of multiple factors: intrinsic stability of the MHC αβ dimer, stability of the peptide-MHC complex, production rate of the of α and β chains, antigen dose, and finally pMHC localization within a cell.

1.1.3.3.1. Protein-MHC complex stability

The p-MHC complex stability and the resulting halftime of the complex on the cell surface correlates well with the **peptide immunogenicity** (23, 24). It seems that a minimal productive complex halftime is 1 hour, as under this threshold, the peptide does not elicit immune reaction under any tested circumstances (23–25). In addition, the complex must be stable within a pH ranging from the acidic pH=5.5 of the endosomal compartments to neutral pH of the cell surface (25).

Higher p-MHC complex stability leads to higher **T cell activation** (26, 27) and weak (unstable) bonds cause reduced p-MHC half-time on the cell surface and **T cell anergy** (28) or **inability of T cells to survive** (7). p-MHC stability also influences other aspects of the immune activation: it is inversely correlated with an optimal peptide concentration (i.e., the peptide concentration needed for the optimal T cell stimulation). Furthermore, it seems that high affinity p-MHC bonds lead to more complex T cell **repertoire** (29).

The p-MHC stability correlates with a number of complexes on the cell surface (30) and changes in the stability may be able to cause difference in a cell-surface numbers and thus affect the T cell activation.

1.1.3.3.2. MHCII αβ dimer stability

One of the factors affecting the p-MHC complex stability is an intrinsic stability of the MHCII dimer, which is to a high extent determined by an ability of the α and β chains to form hydrogen bonds with each other (31). Stability of MHCII dimers differs in between haplotypes and has consequences for the repertoire and amount of peptides presented by a particular dimer.

The intrinsically **unstable** αβ dimers are expected to present more narrow repertoire of peptides, as only class II complexes with high-affinity peptides will be stable enough to last on the cell surface. Despite unstability, these dimers can be present on the cell surface in high numbers, depending on the availability of high-affinity peptides and accessory molecules (32).

Stable αβ dimers, on the other hand, are expected to present the peptides with wider affinity ranges and their expression on the cell surface is more constant and not limited by a dose of high-affinity peptide (32).

Inability to form stable pMHC complex with certain self-epitopes in thymus may enable escape of self-reactive T cells to the periphery and the instability of αβ dimer was repeatedly associated with a susceptibility to autoimmune diseases (32, 33). For example, HLA-DQ dimers formed by the autoimmunity-prone DR3-*DQA1*05:01-DQB1*02:01* and DR4-*DQA1*03:01-DQB1*03:02* haplotypes (34) are amongs the least stable ones (32), and the T1D risk correlates with the DQ dimer instability for majority of DQ haplotypes, including the most predisposing and protective ones (32). Thus, MHCII dimer unstability may be a risk factor for a subset of autoimmune diseases.

In addition to *cis* dimers, HLA-heterozygotes can form ***trans*-dimers** by assembly of α and β chains from the two haplotypes (32, 35), with a unique antigen binding abilities (36). To add to the complexity, the balance in the formation of *trans*-dimers may be affected by a widespread phenomenon of **random monoallelic expression**, where, at a given point of time, as a result of stochastic processes a cell expresses only a single allele out of a pair (37). This phenomenon was observed in HLA genes as well, where 77% of Raji B cell line cells

expressed HLA-*DRB1* from both alleles, and only 23% expressed *DQA1* from both alleles (38). This means that even in heterozygotes that are able to form stable *trans*-dimers, a fraction of cells may express only 2, or even a single type of DQ molecules.

1.1.3.3.3. Presence of HLA-DM, HLA-DO and CLIP – editing the peptide repertoire

During αβ dimer assembly in the ER, the peptide-binding site is protected from premature binding of proteins by an invariant chain (Ii) and after the Ii cleavage by its proteolytic fragment, class II-associated Ii peptide (CLIP) (39). The CLIP removal and replacement by a specific peptide is facilitated by an HLA-DM molecule, which is expressed by all class II expressing cells (3). HLA-DM also stabilizes empty HLA class II molecules, and thus prevents them from aggregation, inactivation and proteolytic degradation (40).

DM complexes with empty HLA dimers are stable and the interaction can be broken only by binding of high-affinity peptides. By ensuring preferential presentation of highly stable p-MHCII complexes (41), DM significantly narrows the repertoire of presented peptides (30) and increases the frequency of high-stability p-MHC complexes on the cell surface.

HLA-DM activity is inhibited by the HLA-DO molecule (41) which is expressed in subsets of class II expressing cells (3, 42). The effect of HLA-DM and -DO molecules in a particular cell type depends on the ratio of free (active) DM relative to the pool of inactive DM-DO complexes and corresponds with the activation state of the cell: in APCs expressing high levels of HLA-DO (immature DCs, resting B cells), inhibition of the -DM molecule enables presenting of more diverse spectrum of lower stability peptides. In this case, low stability and low frequency of individual p-MHC complexes decrease the chance of an immune response targeted against the presented peptide and promote the self-tolerance (41). The APCs expressing HLA-DM only (e.g., activated DCs, activated B cells) present the more narrow spectrum (implying also higher frequency) of the stably bound peptides, enabling efficient immune response against the peptides presented on the activated APCs (reviewed in (41)). DM presence also increases total abundance of certain subset of MHCII alleles (alleles with low CLIP affinity (see below) – DP4, DQ1, DR**04*04) on the cell surface (43).

According to model presented by Busch, the fate of the molecules that **bypass the HLA-DM editing** (even in the HLA-DM sufficient cells a certain fraction of class II molecules may bypass HLA-DM) depends on their CLIP-binding capability and their intrinsic stability in a following way: αβ dimers that strongly bind CLIP will retain this peptide; dimers with weak CLIP binding will loose this peptide and either aggregate (intrinsically unstable αβ dimers) or load other endosomal peptides (stable αβ dimers). In the latter case, the class II molecules loaded in the absence of HLA-DM carry also low-affinity peptides and this **altered peptide repertoire** may be prone to exchange on the cell surface, e.g., for extracelluler self-peptides during an inflammation (40). In this way, **low CLIP-αβ affinity may bring an autoimmune potential**, and indeed, the class II alleles conferring susceptibility to autoimmune diseases often have low affinity for CLIP (e.g., DR4 subtypes predisposing to rheumatoid arthritis)(40) and are resistant to DM editing, as is the case of T1D-predisposing DQ2 and DQ8 molecules (44).

1.1.3.3.4. Expression level

It is difficult to separate and study the effect of class II transcription rate on autoimmunity *per se*, however, the connection between them is supported by the fact that various CIITA polymorphism are associated with increased risk of autoimmune diseases (21, 45). High class II expression correlated with SLE risk carried by a haplotype in a study of Raj (46) and an increased transcription of transgenic HLA class II molecules in DC was associated with disesase severity in a murine model of an autoimmune disease (47).

1.1.3.3.5. Posttranslational mechanisms of regulating pMHC availability on the cell surface

The availability of the pMHCII complexes on the cell surface depends not only on the rate of the class II dimer production and its stability, but also on its distribution within the cell and the rate of its degradation. From the known data it seems that under normal circumstances allele identity has no major influence on rate of protein degradation or intracellular trafficking (48, 49).

1.1.4. Conclusion

The p-MHC-TCR interaction is an important factor affecting the outcome of the antigen-specific immune response. The most crucial characteristics of the interaction are the p-MHC stability and the pMHC-TCR affinity and quantity, which depends on the availability of the class II moelcules.

1.2. HLA class II expression – regulation

The following chapter describes factors that determine expression of HLA class II molecules. It presents general concepts that apply to all HLA (and MHC molecules), as well as factors that affect expression levels of certain alleles, with the special focus on the alleles of *DQA1* and *DQB1* genes.

1.2.1. HLA class II genomic region

Class II genes are located within MHC region on chromosome 6p21.3. Whole region comprises 3.6 Mb and can be divided into class I, class II and class III regions, each of them containing many genes functioning in innate and adaptive immunity. HLA class II region is located on the centromeric end and spans 0.5 Mb (50). In addition to loci coding for classical and non-classical α and β chains, most HLA haplotypes carry an additional DRB expressing locus, containing either *DRB3*, *DRB4*, or *DRB5* paralogous gene. Furthermore, there are several pseudogenes in the region: 0–3 out of 5 pseudogenes (*DRB2 and DRB6–9* (51) in DR region, the number depending on haplotype; *DQB3*; and *DPA2 and DPB2* (4). *DQA2* and *DQB2*, once thought to be pseudogenes, encode the DQX molecule expressed on specialized cell types (52).

Expression of HLA class II molecules is tightly regulated—under normal conditions, they are found only on immune cells: thymic epithelial cells, activated human T cells and antigen-presenting cells such as B lymphocytes, cells of monocyte-macrophage lineage and immature dendritic cells (constitutive expression). However, they can be induced on almost any cell by interferon-γ (inducible expression) (53) and their expression is further fine-tuned depending on several parameters, like developmental stage, activation status of the cell or extracellular stimuli (53).

1.2.2. Regulation of HLA class II expression—genetic regulation

The expression of class II molecules is regulated mainly on the transcriptional level. Several types of DNA sequences are involved in the regulation, and the most important and best defined of them is an SXY regulatory module in the proximal promoter (53), however, several other regulatory sequences—often located several kilobases from the transcription start site—play important role as well (54). These regulators will be also described in the following sections.

1.2.2.1. SXY regulatory module of proximal promoter

Proximal promoter of all class II genes, including accessory genes (Ii, HLA-DM, HLA-DO (55)) contains conserved regulatory module located 150–300 base pairs (bp) upstream of transcription initiation site. Module consists of three sequences, so called S, X (further divided into X1 and X2) and Y boxes. Sequences of SXY module are cooperatively bound by

ubiquitous, constitutively expressed protein factors, forming a multiprotein enhanceosome. The enhanceosome serves as a platform for protein-protein interaction mediated binding of major class II transcriptional regulator, class II transactivator (CIITA). CIITA recruits various transcriptional coactivators and chromatin-remodeling enzymes that trigger transcription, and thus its binding is the critical step that induces class II expression (53). In contrast to the proteins that bind the SXY module DNA, expression of CIITA is highly regulated.

An SXY module functions together as a single regulatory unit, and exact stereospecific alignment of its elements is very important for a proper arrangement of all proteins forming the enhanceosome. Therefore, not only the sequences of S, X and Y boxes, but also the distances between them are highly conserved, and polymorphisms affecting the pattern have an impact on the gene and allele expression. The important conserved characteristics of the SXY module motifs and the consequences of altering them are listed below (in the downstream direction).

1.2.2.1.1. S motif

Functional effect of S box sequence variation is not known. There is no variation in the S box sequence within any of the class II isotypes, so if a polymorphism in the S box contributes to variant regulation of the individual alleles, the effect would present only on the inter-isotype level.

1.2.2.1.2. S-X spacer region

Absolute distance between S and X boxes has a strong influence on the CIITA binding to the promoter. Native S-X distance of the *DRA* gene is 16 nucleotides, and even 1–2 bp change in the spacer length is able to strongly reduce CIITA binding to the mutated promoter (56). Increasing or decreasing the length more results in profound decrease in CIITA binding (56).

TG insertion is present at the position -179–180 in the S-X linker region of several *DQB1* alleles. It increases an absolute distance between S and X elements with a negative impact on the promoter activity; and it also serves as a binding site for a TG-specific nuclear protein, thus positively contributing to the allele expression (57). The overall effect on the promoter strength is negative, and TG insertion is responsible for 2-fold reduced activity of *DQB1*03:02* compared to **03:01* allele (57). The TG insertion is present in several *DQB1* alleles and thus can affect their promoter activity.

1.2.2.1.3. X1 box

X1 box (also referred to as X box) lies 15–17 bp downstream of the S box (58). It is bound by a heterotrimeric protein RFX (59) that functions by promoting cooperative binding interactions with other proteins that bind MHC promoters (59). In the absence of RFX, MHC class II expression is completely abrogated (59).

RFX shows *DRA* > *DPA* > *DQA* affinity gradient towards the promoter X boxes (60) and natural variation in the positions important for RFX binding can contribute to the inter- and intra-allelic variation in class II promoter activity (61).

A single nucleotide G→A change at the position –159 at the beginning of the X1 region, and A→G change at position -146 in X1/X2 region is able to decrease both basal- and agonist-induced transcription of the *DQA1* promoter constructs containing X-Y region. This could partially explain the lower transcription rate of *DQA1*05:01* allele (-159 A and -146 G genotype) compared to *DQA1*03:01* (G-A genotype) (62).

1.2.2.1.4. X2 box

X2 box partially overlaps with the X1 box and its sequence is homologous to either CRE or TRE elements (58). It is bound by the cAMP-response binding element, CREB (previously X2 binding protein, X2BP) (63), which is somewhat tolerant to the variation in its recognition

sequence (64). The X2 box sequence varies both between class II isotypes and between alleles within an isotype, but there is no study on effect of this variation on the class II promoter strength.

TRE element (TGA(C/G)TCA) is present in the *DRA* and *DPB* promoters and is recognized by the transcription factor AP-1(65), which is needed for maximal transcription of these genes (66).

1.2.2.1.5. Relative orientation of X and Y elements

X1 and Y boxes are separated by approximately 2 helical turns (20 nucleotides) of DNA and are aligned on the same side of the DNA helix (67). Requirements for exact spacing between X and Y elements are much looser compared to spacing of S and X elements, provided their relative orientation on the DNA helix is preserved (68).

1.2.2.1.6. Y box

Y box lies 19–20 bp downstream of X box (58). It is an inverted CCAAT box recognized by a trimeric protein NFY (68). Based on the A/G polymorphic site in the position -118 in the Y box, *DQA1* alleles can be divided into two groups: *DQA1*01*,**02* and **03* (G allele); and *DQA1*04*,**05* and **06* allele groups (A allele) (4, 62, 69). Y boxes of G-alleles are preferred targets of NF-Y binding, which is accompanied by higher mRNA expression of these alleles (70), both in basal and cytokine (TNFα, IFNγ) induced transcription (tested on the promoter construct in reporter gene vector (62). However, an A-allele Y box seems to be bringing an increased responsivity to the TNFα induction (62).

Y boxes of the *DQA1*01:01* and **05:01* alleles (but not *DRA* Y box) are able to bind DEK, a nuclear protein that changes the topology of chromatin (71). Importance of this observation is unknown (71) and DEK binding on the Y box and its function was not followed upon in any further study. All *DQA1* gene sequence elements are summarized in article by (72).

In conclusion, proximal promoters of HLA class II genes are highly polymorphic, and natural or artificial variation in the sequence can have a significant effect on the binding of transcription factors (70) and promoter activity. However, the high cooperativity of class II transcription factors binding to the promoter can be an effective mechanism to overcome or decrease the deleterious effect of certain sequence variants, and explain the coordinated regulation of all genes despite sequence differences (73).

1.2.2.2. CIITA

A master regulator of HLA (and MHC) class II genes transcription is the CIITA protein (53, 59, 74). Although some exceptions do exist (75), generally we can say that without CIITA, there is no class II transcription; furthermore, there is a strong positive correlation between levels of the CIITA protein or mRNA in the cell and the level of any of the HLA class II transcripts (74, 76).

1.2.2.2.1. CIITA isoforms and their function

CIITA is transcribed from 4 different promoters, each with its own role, resulting in 3 protein isoforms that differ by their N-terminal parts: the sequence of the shortest isoform IV is shared by all variants, but the isoforms I and III have an additional, isoform-specific, N-terminal part. Each isoform functions in different cell types:

Activation of CIITA promoter I leads to production of **CIITA isoform I** that regulates constitutive HLA class II expression in dendritic cells (DC) and cells of macrophage lineage (both with and without induction by IFN-ɣ) (77).

Promoter III and **CIITA isoform III** regulate constitutive expression in B lymphocytes (78) and in plasmacytoid DC; and inducible expression in activated human T cells (79). Transcription of this form is also induced by IFNγ, although to a lower extent compared to

isoform IV (78). Isoform III has the highest transactivation potential (80).This isoform is also more efficiently recruited to the HLA-*DRA* promoter *in vivo* with and shows increased interaction with components of the transcription machinery compared to forms I and IV (80).

Isoform IV, transcribed from promoter IV, is responsible for IFN-ɣ inducible CIITA expression in cells of both non-hematopoietic and hematopoietic origin (77, 78).

The usage of individual promoters is not completely exclusive, as significant fraction (around 1/3) of CIITA in DCs originates also from pIII (78), and after IFNγ induction, not only pIV, but pI is activated to small extent as well.

1.2.2.3. Inverted SXY module (XL sequences)

Several sequences highly homologous to X and Y boxes of the SXY module, but in reverse orientation, were identified within MHC class II and extended class II region. These sequences bind proteins of the class II enhanceosome including CIITA (54, 81). CIITA binding is associated with chromatin remodelling and extragenic transciption (81–83) and several authors suggested that these distal sequences could act as distal enhancers (82) or locus-control-regions (LCR) (81).

The regulatory potential is consistent with location of the XL sites at strategic positions separating the individual class II isotypes (82).

1.2.2.4. CTCF (insulator) binding sequences

The next category of sequences important for class II expression are the CTCF binding sequences. They are bound by CTCF proteins that bring together pairs of CTCF binding sites, forming functionally separated chromating loops (84, 85).

Ten functional binding sites for CTCF are located within HLA class II region, in the key positions between individual subregions (86, 87). Each of these CTCF binding sites is able to interact with several of its neighbours (86). In the presence of CIITA, 3D structure of region reorganizes and CTCF sites form interactions with CIITA, and through them with class II promoters as well (each promoter preferably interacting with its closest CTCF sites) (38, 86). This organization seems to be crucial for an efficient expression of all class II genes associated with antigen presentation (86).

1.2.3. Regulation of HLA class II expression—posttranscriptional regulation

HLA class II regulation on the posttranscriptional level involves slternative splicing of certain *DQA1* (88) and *DQB1* (89) alleles with a potential effect on cAMP signalling in B lymphocytes (90) and with production of secreted forms of HLA-DQβ (91).

The interallelic differences in mRNA stability of the *DQA1* and *DQB1* genes were described as well (92), but they do not affect all alleles (93).

1.2.4. Regulation of HLA class II expression—epigenetic regulation

Epigenetics studies heritable changes in phenotype that are not caused by an underlying change in genotype. Main epigenetic mechanisms are DNA methylation, covalent posttranslational modifications of histone proteins and RNA-mediated gene regulation. Different chromatin modifications are interconnected and influence each other.

1.2.4.1. DNA methylation

One of the epigenetic modifications in eukaryotes is DNA methylation, which involves covalent attachment of methyl group on the fifth carbon of cytosine in the CpG sequence (94).

1.2.4.2. Function of DNA methylation

Effect of DNA methylation depends on its position in genome. For regulation of transcription initiation, the most important is methylation located in regulatory regions (in a

promoter or a first intron) of the gene**.** Promoter methylation is generally associated with gene silencing and the mechanism of the association can be described by two non-exclusive models:

- Methyl group directly blocks an access of transcriptional coactivators to their cognate sequences
- 5meCpG is recognized by methyl-CpG-binding proteins (MBD) that induce repressive state of chromatin (94, 95). In this case, it is plausible that the strength of the effect could depend on the local concentration of methylated cytosines (96).

Based on a CpG content, gene promoters can be divided into three categories: low CpG promoters are enriched among tissue-specific genes, are generally methylated, and their methylation status does not preclude the gene activity. Authors assume that at low methylation density, MBD binding is probably not sufficient for active repression of expression, however, this does not exclude the role of low-density methylation in reducing the transcriptional noise (96).

Promoters with high CpG content are more frequently associated with housekeeping genes and are usually non-methylated regardless of promoter activity. Promoters with intermediate CpG content were intermediary methylated and a higher correlation between promoter methylation and activity was observed. In promoters with high and intermediate CpG content, DNA methylation, if present, represses gene expression (96).

According to criteria of the study above, all class *DQA1* and *DQB1* allele promoters belong among low-CpG promoters, which should be methylated and their methylation should not be an obstacle to the allele expression. A CpG rich region is located in the intron 1 of *DQB1* gene. As DNA methylation of regulatory element located in the first intron of a gene is able to suppress transcription of interleukin 4 gene (97), it is possible that the same mechanism could regulate the *DQB1* gene as well.

1.2.4.2.1. DNA methyltranserases and methyl-CpG binding proteins

A transfer of the methyl group from the S-adenosyl methionine donor (SAM) to the DNA is carried out by DNA-methyltransferase enzymes (DNMT). DNMT1 is a 'maintenance' methyltransferase. It adds a methyl to cytosine in hemimethylated CpG dinucleotides during replication and during a DNA repair (98). DNMT3a and DNMT3b are *de novo* methyltransferases. They recognize and methylate sequences in which none of the pairing CpG dinucleotides is methylated (99).

Interpretation of the DNA methylation mark is mediated by methyl-CpG-binding proteins (MBP) that specifically recognize methylated CpG dinucleotides through their methyl-CpG binding domain (MBD). All MBPs associate with corepressor complexes which modify chromatin and in this way suppress the gene expression (94, 100).

Both DNMTs (101) and MBPs (94) can act in a DNA sequence-specific way, potentially creating epigenetic variation between gene alleles.

1.2.4.3. Allele-specific DNA methylation and its effect on gene expression

A variant DNA methylation of two alleles of a single (non-imprinted) gene correlating with a sequence of a given allele has been described, as well as association of this variant methylation with an allele expression (102–104). This phenomenon can affect up to 10% of human genes (103).

1.2.4.4. HLA class II DNA methylation

1.2.4.4.1. Direct blocking of transcription factors binding in the HLA region

CREB and CTCF are the only factors involved in regulation of class II expression that contain a CpG dinucleotide in their recognition sequence and whose binding is sensitive to cytosine methylation.

CpG in the CREB binding site (X2 box) is present only in the SXY module in the promoter of the *HLA-DRA* gene. The CpG cytosine is mutated in all other class II isotypes, and thus *DRA* is the only class II gene in which CREB binding can be directly regulated by DNA methylation.

There are several CTCF sites interspersed through class II region. It was shown that DNA methylation of CTCF sites surrounding DQ locus is responsible for the lack of DQ expression in DQ negative B cell line. DNA methylation was associated with formation of heterochromatin and loss of CTCF binding to these sites; and with loss of RFX and CIITA binding to the *DQA1* and *DQB1* promoters (though methylation of promoter regions of these genes could also contribute to the effect) (105).

1.2.4.4.2. Effect on chromatin structure in the HLA region

Except the examples above, all CpGs in the class ll genes are located outside of known binding sites of regulatory proteins, therefore, if DNA methylation affects the gene expression, it has to be through modifying the chromatin accessibility.

Till now, most studies exploring HLA class II gene DNA methylation were performed on practically non-polymorphic *HLA-DRA* gene, on few sites recognized by methylation sensitive restriction endonucleases.

The CpG sites situated close to the start codon were usually methylation free (106–108) and their hypomethylation seemed to be important factor allowing for the gene expression (108). The association of non-promoter *DRA* gene methylation with expression was not so conclusive, some studies found that demethylation of a particular site/s is associated with genes expression (109, 110), some have found the contrary (111, 112).

Not much is known about class II allele-specific DNA methylation. We are aware of only two studies that studied methylation of individual alleles of HLA genes. In 1985, Uitterlinden *et al.* found variant DNA methylation between class I MHC haplotypes in rats (113). In 1992, using methylation sensitive restriction endonuclease, Toyoda *et al.* found differences in methylation of HLA-*DQB1* alleles (114). He observed interallelic differences in DNA methylation but constant pattern of methylation for a particular allele—an individual CpG site was either methylated in all examined sequences, or unmethylated in all sequences. The only allele with variant methylation was DR4-associated DQ7 (i.e., *DQB1*03:01* or **03:04*).

1.2.4.5. Histone modifications and their connection with DNA methylation

Histones undergo posttranslational modifications on specific amino acid residues in the N-terminal part of the protein. The message carried by modifications depends on the type of modification (acetylation, methylation, phosphorylation, ubiquitination, and more), its position within a nucleosome (which residue on which histone is modified) and on the degree of modification (e.g., mono-, di-, and trimethylation). The marks deposed on histones are specifically recognized by chromatin-remodeling proteins that translate them into the change of chromatin state. The resulting effect on chromatin and on gene expression depends on combination of modifications; genomic context (gene promoter, enhancer, body of the gene, CpG island) and on the pattern of the modification distribution over the region (115). Histone modifications are interconnected with each other as well as with DNA methylation (95).

Histone acetylation is a reversible lysine modification associated with open chromatin transcription. It is controlled by histone-acetyltransferases (HAT) and histone deacetylases (HDACs), which typically function as transcription co-activators (HAT) and corepressors (HDAC), respectively (95, 116). The role of link between histone modifications and HLA class II expression is played by a CIITA transcriptional coactivator. This protein not only associates with HATs, but also acts as one (53).

Histone methylation. Lysines and arginines can be modified by the addition of one to three methyl groups. DNA methylation is strongly associated with the absence of methylation

on lysine 4 and presence on lysine 9 on the histone H3 (H3K4 and H3K9). Histone methylation can be associated with gene activation or repression, depending on the exact modification and its genomic context (95, 117).

1.2.4.5.1. Allele-specific histone modifications and chromatin state

DNA sequence variation affects histone modifications in *cis* as well, as was repeatedly observed on association of polymorphic SNPs with histone modifications in the surrounding (115) as well as in distal (1 Mbp) region (through chromatin looping) (115, 118–120). Around 10% of active chromatin sites is allele-specific (120).

1.2.4.5.2. Allele-specific chromatin modifications of HLA class II genes

As discussed above, allele-specific histone modifications are quite a common feature in the human genome (119), however, there is no information on allele specific histone modifications of class II genes.

To conclude, HLA allele behavior is determined by its peptide sequence as well as by its production rate. High polymorphism on the genetic level may translate into variation in the epigenetic marks (DNA methylation, histone modifications) and consequently into uneven expression of the individual les, with an impact on its function in immune system. The aim of our work was to add missing points to this picture and analyze HLA allele-specific DNA methylation and mRNA expression.

2. Aims and Hypotheses

The overall aim of this thesis was to study allele-specific mRNA expression and DNA methylation of the HLA class II genes. The work was divided into two studies, each addressing a different aspects of the HLA class II allele expression regulation:

Study A: HLA class II DNA methylation and its effect on the mRNA expression

Hypothesis: DNA methylation can regulate expression of HLA class II genes.

Aim: Determine DNA methylation of individual alleles of the HLA-*DQA1* gene. Analyze inter-allelic differences in DNA methylation. Analyze the relationship between allele DNA methylation and mRNA expression.

Study B: HLA class II mRNA expression (allele- and cell-type-specific expression)

Hypothesis 1: There are differences in HLA class II expression, but are not well described.

Aim 1: Determine the mRNA expression hierarchy of HLA class II *DQA1* and *DQB1* alleles.

Hypothesis 2: Expression hierarchy of HLA class II alleles may differ in the different immune cell types.

Aim 2: Determine expression hierarchy of the *DQA1* and *DQB1* alleles in the whole blood cells, B cells and monocytes.

3. Material and Methods

Subject of the studies

The experiments were undertaken with the understanding and written consent of each subject. The study methodologies conformed to the standards set by the Declaration of Helsinki. This study was approved by The Ethical Committee of the Third Faculty of Medicine of Charles University in Prague.

All subjects were healthy, i.e., without an autoimmune disease (self-reported), volunteers of European descent.

Study A: DNA and RNA isolated from the whole blood cells of the healthy individuals (aged 19–39, n= 60).

Study B: DNA and RNA isolated from the whole blood cells, B lymphocytes and monocytes of the healthy subjects (aged 19–69, mean 31.2±12.1; n= 42) were recruited (for the study).

Methods

HLA genotyping
DNA isolation from whole peripheral blood
SSP-PCR genotyping of *HLA-DRB1, HLA-DQA1* and *HLA-DQB1* genes
Analysis of DNA methylation
DNA isolation from whole blood
Bisulfite conversion
Cloning into *E. coli*
DNA sequencing (provided by an external company)
DNA sequence analysis
Analysis of mRNA expression
Isolation of monocytes and B lymphocytes from peripheral blood
RNA isolation from the whole blood, monocytes and B lymphocytes
qPCR assay design
Gene and allele expression analysis using RT-qPCR

Statistical analysis

Data were analyzed using Prism 3.0 software with appropriate statistical tests. The level of significance was set at *$p < 0.05$, **$p < 0.005$ and ***$p < 0.0005$.

4. Results

The parts of this chapter are based on published articles (121, 122). A subset of data on *DQA1* promoter methylation was previously reported in my diploma thesis (123). DNA methylation and mRNA expression measurements from the study A were used to provide data for the group of healthy controls in the study of DNA methylation and mRNA expression of *HLA-DQA1* alleles in type 1 diabetes mellitus (121).

4.1. HLA genotyping

Results of genotyping of 158 individuals included into this study are summarized in **Table 1**. HLA genotyping results of the Czech subjects from this study combined with genotyping results of subjects from other studies conducted at the Department of Medical Genetics were published (124) and included into the public repository of HLA allele and haplotype frequencies www.allelefrequencies.net (125).

4.2. DNA methylation of *HLA-DQA1* promoter region (Study A)

Methylation status of 213 *HLA-DQA1* promoter (*QAP*) sequences from 35 individuals was analyzed. Out of these 35 subjects, mRNA expression data were available for 18 individuals. Out of them, 14 had both expression and DNA methylation analyses done on the RNA and DNA isolated from the same blood *DRA*w.

4.2.1. Methylation of individual CpG sites does not differ between alleles

The only difference in DNA methylation of individual CpG sites found to be significant after correction was between the *DQA1*01:01* and *DQA1*01:02* (DR15-linked) promoter at site −508 (pun = 0.0043/pc = 0.0473, RR CI = 0.9436 to 10.57). However, RR CI contained the value 1 and therefore we could not reject the possibility that the association observed is only due to a chance. No other differences in DNA methylation of any CpG site between any

2 alleles were observed (data not shown). The average mehtylation of individual CpG sites is shown in **Figure 1A.**

4.2.2. The most methylated alleles are *DQA1*02:01* and **04:01*

Because of the low number (3) of *DQA1*01:04* promoter allele sequences, we did not include it into methylation analysis. Promoters of *DQA1*02:01* and *DQA1*04:01* were significantly more methylated than most of the other alleles (*DQA1*01:01*, **01:02*— DR16-linked, **01:03*, **05:01*, **05:05*); promoter of *DQA1*04:01* was also methylated more than promoter of *DQA1*03*, and methylation of *DQA1*01:02* (DR13-linked) promoter was higher than that of *DQA1*01:01*. Overall methylation ranged from 7.6 methylcytosines on average per promoter for *DQA1*01:03* allele to 10.5 methylcytosines per promoter in *DQA1*02:01* and 10.8 methylcytosines in *DQA1*04:01*. Overall methylation of individual alleles is depicted in **Figure 1A.**

Table 1. ***DRB1-QAP-DQA1-DQB1* haplotypes and their frequencies in the study population.** The table lists only haplotypes with frequency > 1%. The *DQA1* promoter alleles (*QAP*) are listed as a part of haplotype. [a]The available nomenclature of does not distinguish all alleles present in our samples, so we split some of the existing alleles into 2 groups and use our own marking to denote them. Based on the sequence differences, allele *QAP 1.2* was split into groups *QAP 1.2K* and *QAP 1.2L*, *QAP 1.3* into groups *QAP 1.3a* and *1.3b*, and *QAP 4.1* into *QAP 4.1A* and *4.1B*. These "new" alleles differing in their sequence were also parts of different HLA-haplotypes; [b]Allele *DQA1*01:02* is associated with 3 different promoters (in 3 different haplotypes).

*DRB1**	QAP	*DQA1**	*DQB1**	n	f (%)
11	4.1A	05:05	03:01	50	15.9
15	1.2L	01:02 [a]	06:02	39	12.4
01	1.1	01:01	05:01	38	12.1
07	2.1	02:01	02:02	31	9.8
03	4.1B	05:01	02:01	29	9.2
13	1.3a	01:03	06:03	23	7.3
04	3.1	03:01	03:02	20	6.3
16	1.2K	01:02 [a]	05:02	13	4.1
12	4.1A	05:05	03:01	9	2.9
13	4.1A	05:05	03:01	9	2.9
08	4.2	04:01	04:02	8	2.5
13	1.4	01:02 [a]	06:04	8	2.5
14	1.3b	01:04	05:03	7	2.2
07	2.1	02:01	03:03	6	1.9
04	3.1	03:03	03:01	5	1.6
04	3.1	03:03	03:02	4	1.3

4.3. mRNA expression of HLA class II genes *DQA1* and *DQB1*

4.3.1. mRNA expression of *DQA1* alleles in whole blood cells (Study A)

We analyzed mRNA expression data from 43 individuals. Data included 10 *HLA-DQA1* alleles linked with 12 different promoter alleles. Each *DQA1* allele was linked to one promoter allele, except *DQA1*01:02* that was linked to 3 different *DQA1* promoters, *QAP 1.2K, QAP 1.2L* and *QAP 1.4*, that were part 3 different *DRB1* haplotypes, *DRB1*16*, *DRB1*15* and *DRB1*13*, respectively. *DQA1* alleles and their linked promoter alleles (*QAP*) are listed in **Table 1** as a part of the respective HLA class II haplsotype (126).

Because of a low number of sequences obtained (1, 2 and 1, respectively), expression data of *DQA1*01:02* (DR16-linked), *DQA1*01:04*, and *DQA1*01:05* alleles were not included into the analysis. Average mRNA expression levels of individual alleles normalized to *DRA* varied over 3-fold (0.07 for the least expressed *DQA1*05:05* and 0.23 for the most expressed *DQA1*03*). Statistical analysis showed that *DQA1*03* alleles were significantly overexpressed compared to most other alleles (*DQA1*01:01*, **01:02* (DR13-linked), **01:03*, **02:01*, **05:01*, **05:05*), and *DQA1*05:05* allele was expressed less than *DQA1*01:03*, **02:01*, **04:01* alleles; the *DQA1*01:03* allele was expressed more than *DQA1*01:01* allele (**Figure 2**).

To analyze the differences in more detail, we calculated the ratio between mean expression levels of the alleles. Mean expression ratio of alleles that proved to differ significantly ranged from 1.4 (*DQA1*02:01*/*DQA1*05:05*) to 3.6 (*DQA1*03*/*DQA1*05:05*) (**Table 2**). The allele

Table 2. Mean expression ratio of *DQA1* alleles. For each allele, the mean expression level was calculated (mean allele expression). To obtain the expression ratio of two alleles, the mean expression of one allele (in this table in the title row) was divided by the mean expression of the second allele (title column). Data for allele pairs whose expression proved to differ significantly (**Figure 2**) are highlighted in bold.

***DQA1* allel**	****01:01***	****01:02***[a]	****01:02***[b]	****01:03***	****02:01***	****03***	****04:01***	****05:01***	****05:05***
mean *DQA1* allele expression	0.081	0.088	0.127	0.120	0.091	0.235	0.109	0.085	0.066
(SEM)	(0.004)	(0.003)	(0.006)	(0.008)	(0.005)	(0.008)	(0.008)	(0.006)	(0.003)
****01:01***	1	1.08	1.57	**1.48**	1.12	**2.88**	1.34	1.04	0.81
****01:02***[a]	0.92	1	1.44	1.36	1.03	2.67	1.24	0.97	0.75
****01:02***[b]	0.64	0.69	1	0.94	0.72	**1.85**	0.86	0.67	**0.52**
****01:03***	**0.68**	0.73	1.06	1	0.76	**1.96**	0.91	0.71	**0.55**
****02:01***	0.89	0.97	1.40	1.32	1	**2.58**	1.20	0.93	**0.73**
****03***	**0.34**	0.37	**0.54**	**0.51**	**0.39**	1	**0.46**	**0.36**	**0.28**
****04:01***	0.74	0.81	1.17	1.10	0.83	**2.16**	1	0.78	**0.61**
****05:01***	0.95	1.04	1.49	1.41	1.07	**2.76**	1.28	1	0.78
****05:05***	1.23	1.33	**1.92**	**1.82**	**1.38**	**3.56**	**1.66**	1.29	1

[a] Linked with DR15 (*QAP 1.2L*), [b] DR13 (*QAP 1.4*).

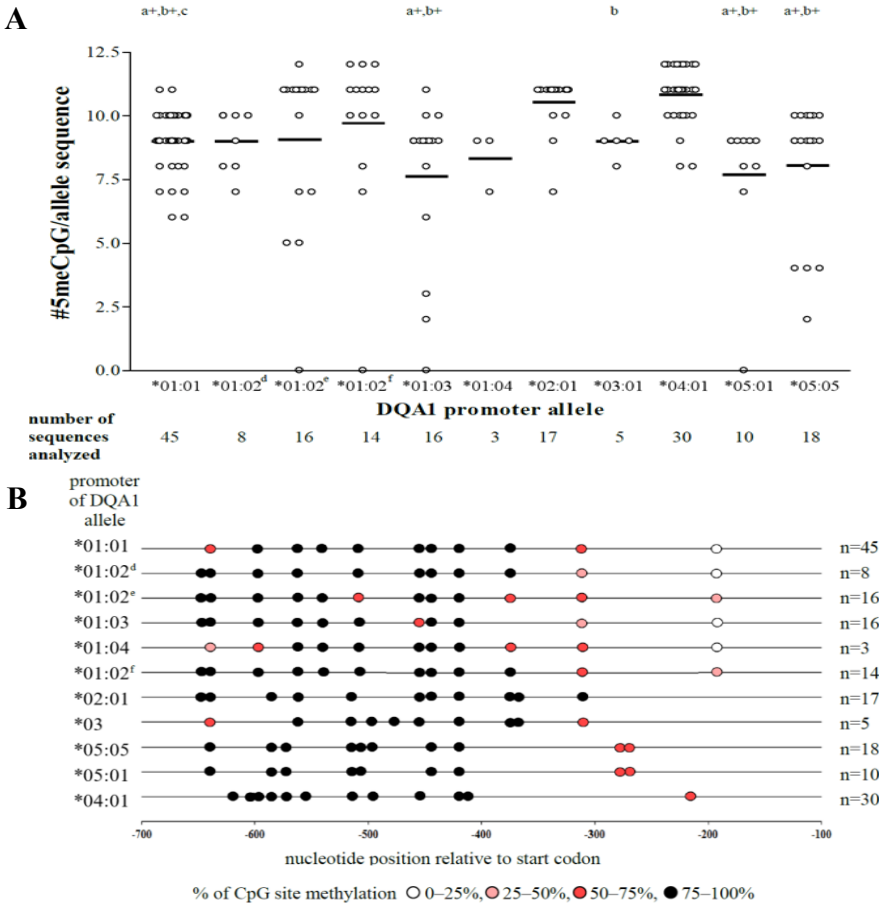
expression hierarchy was preserved in the heterozygous individuals. In heterozygotes, *DQA1*03* allele was always expressed more than the other allele, on the other side *DQA1*05:05* allele was always the less expressed one (with a single exception where *DQA1*01:05/*05:05* expression ratio was 1.09). In single heterozygous individuals we observed distinct *DQA1*01:03* > *DQA1*02:01*,**05:01*, **05:01*, and *DQA1*01:02* (DR13-linked) > *DQA1*05:01* relationship (allele ratio 1.3–1.7); and less pronounced (allele ratio close to one) *DQA1*02:01* > *DQA1*01:02* (DR15-linked) and *DQA1*02:01* > *DQA1*05:01* relationship. As these observations were done in single individuals only, no further conclusions regarding the hierarchy of allele expression could be made.

We also wondered whether and how the presence of specific *DQA1* alleles in an individual's genotype affects total *DQA1* mRNA expression. An inter-individual variation in the allele expression of a particular allele was low and it seemed that an allele mRNA expression level tends to stay the same irrespective of the identity of the other allele present in a heterozygous combination. If this is the case, one should be able to calculate the total *DQA1* mRNA level in the subject only by adding up the known values of an expression of the alleles carried by the subject. Indeed, we observed that amount of total *DQA1* gene mRNA seen in individual samples followed the "theoretical" *DQA1* total relative mRNA level, which was calculated as a sum of mean relative expression of alleles present in a sample.

4.3.2. mRNA expression of HLA class II genes *DQA1* and *DQB1* in B lymphocytes and monocytes (study B)

After the initial study A, we decided to explore the expression differences between the alleles more thoroughly: we included a *DQB1* gene to the study and compared the expression between the different cell types.

In total, the transcription rates in 42 individuals were analyzed. Data included 10 *HLA-DQA1* alleles and 13 *DQB1* alleles. Since allele *DQA1*01:02* is linked to 3 different *DQA1* promoters in 3 different haplotypes, we treated it as 3 different alleles (DR13-, DR15-, and DR16-linked *DQA1*01:02* allele).



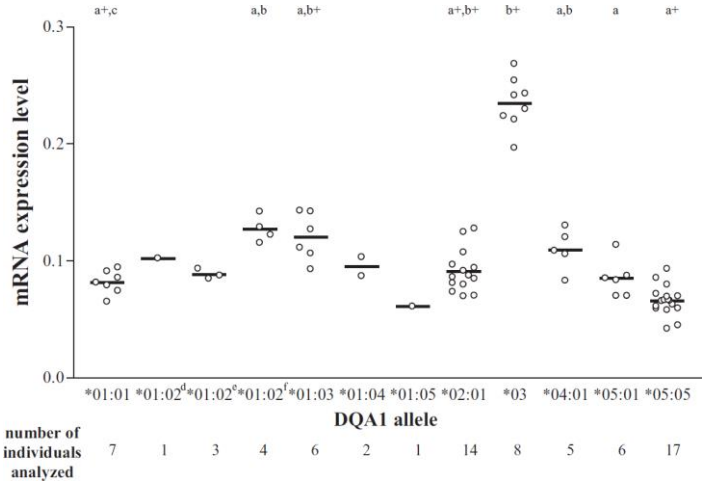


Figure 2. mRNA expression levels of *DQA1* alleles. Expression of *DQA1* alleles was determined by RT-qPCR with allele-specific primers and the mean of 3 independent measurements was used for the analysis. Each point in the graph represents normalized value for 1 individual and 1 allele. The horizontal line indicates mean relative expression of the allele.
[a(a+)]Allele expression significantly, $p < 0.05$ (highly significantly, $p < 0.005$) lower than that of *DQA1*03* allele; [b(b+)]Allele expression significantly, $p < 0.05$ (highly significantly, $p < 0.005$) higher than that of *DQA1*05:05* allele; [c]Allele expression significantly lower than that of *DQA1*01:03* allele;[d] Linked with DR16 (*QAP 1.2K*), [e] DR15 (*QAP 1.2L*), [f] DR13 (*QAP 1.4*).

Figure 1. *DQA1* promoter DNA methylation. According to the polymorphisms present on the sequencing read, sequences were classified into appropriate allelic groups. **(A) Overall methylation of *DQA1* promoter alleles.** For each sequence, the total number of methylated CpG sites in the region examined (i.e. overall methylation) was counted. Each point in the graph represents the number of methylated CpGs per 1 sequence. The horizontal line indicates mean methylation level of each allele. **(B) CpG methylation status of individual CpG sites of *DQA1* promoter alleles**.The matrix represents an amplicon that contains 9–12 CpG sites (number depends on promoter allele identity) obtained for 500 bp region overlapping promoter of the *DQA1* gene. Each colour-coded circle within the matrix represents one CpG site and its colour indicates the percentage of sequences that have the cytosine at the given site methylated. Unmethylated–low methylated sites are indicated by open circles. No interallelic difference in methylation of individual CpG sites was found to be significant after Bonferroni correction.
n – number of sequences analyzed; [a] [(a+)]Allele promoter methylation significantly, $p < 0.05$ (highly significantly, $p < 0.005$) lower than that of *DQA1*02:01* allele; [b] [(b+)]Allele promoter methylation significantly, $p < 0.05$ (highly significantly, $p < 0.005$) lower than that of *DQA1*04:01* allele; [c]Allele methylation significantly lower than that of *DQA1*01:02* (DR13-linked) allele; [d] Linked with DR16 (*OAP 1.2K*), [e] DR15 (*OAP 1.2L*),[f] DR13 (*OAP 1.4)*

4.3.2.1. Relative expression of HLA class II alleles in different cell types

***DQA1* and *DQB1* gene expression in cell types**

First, we analyzed differences in *DQA1* and *DQB1* gene alleles' expression between cell types irrespective of an allele identity. The highest ***DQA1*** level was observed in B lymphocytes. Whole blood cells (WBC) expressed around 1/2 and monocytes around 1/3 of the B cell levels, respectively (**Figure 3A**). The differences were significant for all cell-type pairs tested (B lymphocytes—monocytes; B lymphocytes—WBC; monocytes —WBC, each $p_c < 0.0003$).

For ***DQB1*** gene alleles, the only significant difference was observed between B lymphocytes and monocytes ($p_c = 0.0213$), where *DQB1* allele expression was on average 1.25 times higher in B cells (**Figure 3B**).

***DQA1* and *DQB1* allele expression in cell types**

Then we examined differences in expression between cell types on the level of individual alleles (**Figure 4**). For the ***DQA1*** gene alleles, mean allele mRNA level was always highest in B lymphocytes and lowest in monocytes. Difference in allele mRNA levels between these two cell populations varied from 1.3-fold for the *DQA1*03* group of alleles to more than 13-fold for the *DQA1*05* alleles. Expression levels in WBC lied between the values for B cells and monocytes. Due to a low number of samples in some allelic groups, we were able to statistically compare only expression of *DQA1*01:01*, **02:01*, **03:01* and **05:05* alleles. For *DQA1*02:01*,**03:01* and **05:05* alleles, differences in expression between B cells and monocytes proved to be significant, as well as differences between B cells and WBC for alleles *DQA1*02:01* and **05:05*.

The situation was not so unambiguous for the ***DQB1*** gene. The mean allele mRNA level was higher in B lymphocytes compared to monocytes. This was the same pattern as in the *DQA1* gene, but the differences were less pronounced (1.8 to 4-fold for all alleles in this group). The situation was reversed for the *DQB1*06* allele group where B cells expressed 0.6 to 1.0-fold of the monocyte mRNA levels. The expression of *DQB1*03:03* and **05* seemed to be similar in both cell types. We were able to statistically test differences between cell types for alleles *DQB1*02:02*, **03:01* and **05:01*. The expression of *DQB1*02:02* and **03:01* alleles was significantly higher in B lymphocytes compared to monocytes and compared to WBC.

Theoretical DQ dimer expression

The DQ molecule is a dimer and its amount on the cell surface depends on expression of both subunits. Thus we calculated theoretical expression of DQ dimers based on mRNA expression of the *DQA1* and *DQB1* alleles in B cells and monocytes (**Table 3**). The most and least „theoretically expressed" DQ molecules in both cell types would be DQ6 and molecules with *DQB1*02* and **03*-coded β chains, respectively. The differences in DQ expression between cell types would be highest for the DQ7.5 (4-9 fold) and lowest for the DQ9.2 (same expression on B cells and monocytes).

4.3.2.2. Inter-allelic differences in expression of *DQA1* and *DQB1* genes

Next, we inspected the differences in relative expression between individual alleles. In ***DQB1*** gene, alleles of **02*, **03*, **05* groups were less expressed than **06* group alleles, and this relationship was observed in all cell types examined. The inter-allelic differences were more pronounced in monocytes, where mean mRNA level of the most (*DQB1*06:03*) and least (**02:02*) expressed alleles differed up to 15-fold. In B cells, the expression of individual alleles was more balanced; the most (*DQB1*06:04*) and the least (**02:01*) expressed alleles differed 3.5-fold (**Figure 4**).

Expression hierarchy of the ***DQA1*** gene followed the same pattern in monocytes and WBC, with lowest mRNA levels in *DQA1*05:05* and **05:01* alleles, gradually increasing through **02:01* and **01:01* alleles, **01:02* and **01:03* alleles, and peaking in *DQA1*03* alleles (3-fold higher than in the **05* allele group). In B lymphocytes, similarly to monocytes, *DQA1*01:01*, **01:02* (DR15) and **02:01* were on the low end and **03* group alleles on the high end of the expression spectrum. However, it was surprising to see that *DQA1*05* alleles, whose expression level in monocytes was the lowest out of all alleles, were among the most

Table 3. Theoretical expression of DQ dimers. Expression of HLA-DQ dimers on the cell surface based on mRNA expression of *DQA1* and *DQB1* genes. DQ expression is determined as equal to the mRNA expression of the less expressed gene in the haplotype. Fold DQ expression for a particular DQ molecule is determined by dividing its DQ expression by 0.015 (DQ expression level of the least expressed DQ molecule). This is a simplified model that does not take into account post-translational mechanisms affecting the DQ expression on the cell surface.

Haplotype		DQ dimer	DQ expression		Fold DQ expression		
DRB1	*DQA1*-DQB1**		B cells	mono	B cells	mono	B/mono
01	01:01-05:01	DQ5.1	0.081	0.054	5.4	3.6	1.5
03	05:01-02:01	DQ2.5	0.044	0.015	2.9	1.0	2.9
04	03:01-03:02	DQ8.1	0.098	0.054	6.5	3.6	1.8
07	02:01-02:02	DQ2.2	0.067	0.016	4.5	1.1	4.2
07	02:01-03:03	DQ9.2	0.057	0.056	3.8	3.7	1.0
11	05:05-03:01	DQ7.5	0.073	0.015	4.9	1.0	4.9
13	01:02-06:04	DQ6.4	0.155	0.080	10.3	5.3	1.9
13	01:03-06:03	DQ6.3	0.154	0.066	10.3	4.4	2.3
15	01:02-06:02	DQ6.2	0.134	0.061	8.9	4.1	2.2

expressed ones in B lymphocytes, with levels similar to that of **03* allele group. As for *DQB1* gene, expression of individual alleles was more balanced in B cells; the most (*DQA1*05:01*) and the least (*DQA1*01:01*) expressed alleles differed 1.6-fold, compared to 10-fold difference between *DQA1*03:03* and **05:05* in monocytes.

4.3.2.3. Correlation of *PPIA* and *DRA* gene expression

To analyze the strength of HLA class II promoting stimuli in each cell type, we analyzed an expression of *DRA* gene normalized against a classic endogenous control, *PPIA*. *DRA* expression was significantly higher in B lymphocytes (mean relative *DRA* expression 4.2) and monocytes (4.9) compared to whole blood (1.9; both $p_c < 0.0003$) (**Figure 6A**). There was significant monotonic and linear correlation between Ct of *DRA* and Ct of *PPIA* in each cell type (**Figure 6B**). A value of Spearman's correlation coefficient was 0.88 in whole blood, 0.94 in B lymphocytes and 0.89 in monocytes, and Pearson's correlation coefficient was 0.99 in whole blood, 0.93 in B lymphocytes and 0.90 in monocytes (all six p values < 0.0001). Linear correlation between Cts means that ratio of *PPIA* and *DRA* in individual cell types tends to be constant among individuals. However, the slopes and constants of the equations describing this linear relationship were different for each cell type, indicating a cell type-specific relationship between the amount of *DRA* and *PPIA* transcripts.

4.3.3. mRNA expression of HLA class II genes *DQA1* and *DQB1* in B lymphocytes and monocytes in patients with T1D (pilot study)

mRNA expression of *HLA-DQA1* alleles in WBC of patients with T1D has been analyzed in our laboratory previously (127). In the study, higher expression of *DQA1*02:01* in healthy controls compared to patients was was found. As a next step, a study in was conducted to compare the expression of *DQA1* and *DQB1* alleles in cell populations in patients with T1D (Kotrbova-Kozak A, Cepek P, Zajacova M, Cerna M, 2016, unpublished). The results are shown in **Figure 5**.

Interallelic differences in expression of DQA1 and DQB1 genes

The expression pattern seemed to be the same in controls and patients for both genes in all cell types analyzed. Mean mRNA expression for each allele-cell type combination was calculated and compared between controls and patients with two-tailed Mann-Whitney test. There were enough samples for comparison for the following combinations: *DQA1*01:01, *03:01* and *DQB1*02:01, *02:02, *03:02, *05:01* in B cells, *DQA1*01:01, *02:01, *03:01* and *DQB1*02:01, *02:02, *03:02* in monocytes and *DQA1*03:01* and *DQB1*02:01, *03:02*

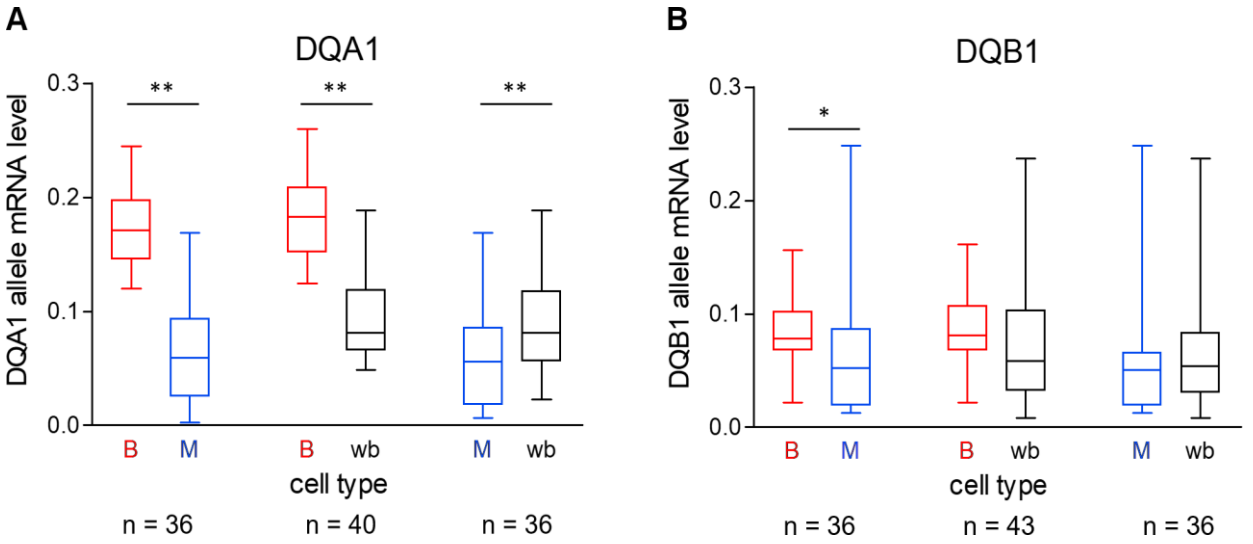


Figure 3. Relative mRNA expression of *DQA1* (A) and *DQB1* (B) alleles in different cell types. Normalized expression data of all *DQA1* (A) and *DQB1* (B) alleles in the indicated cell type were grouped together irrespective of allele identity, and differences in allele expression between cell types were analyzed by two-tailed Wilcoxon matched pairs test. Only individuals for whom measurements in both compared cell types were available (paired measurements) were included into analysis. Number of samples analyzed (n) and cell type (wb – whole blood cells, B – B lymphocytes, M – monocytes) is indicated. *$p_c < 0.05$, ** $p_c < 0.005$.

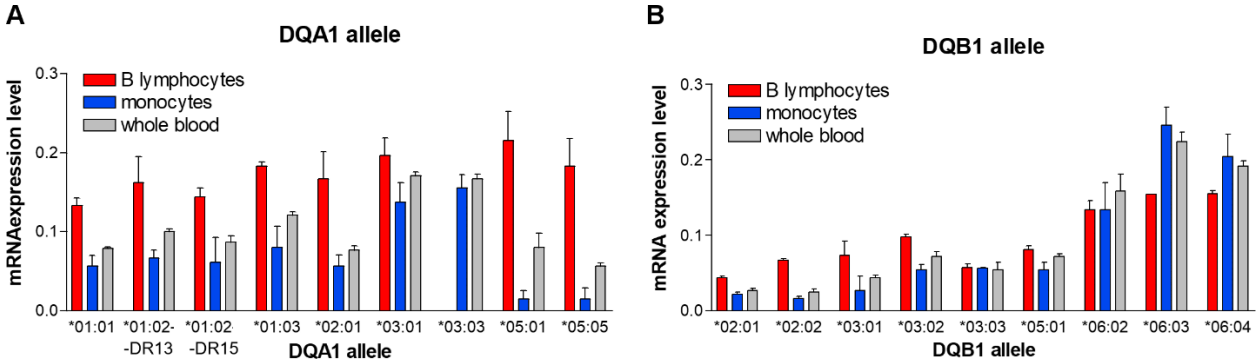


Figure 4. mRNA expression levels of *DQA1* (A) and *DQB1* (B) alleles in different cell types. Expression of *DQA1* (A) and *DQB1* (B) alleles was determined by RT-qPCR with allele-specific primers and the mean of 3 independent measurements was used for the analysis. Each column represents mean of data of at least 2 individuals, SD is indicated.

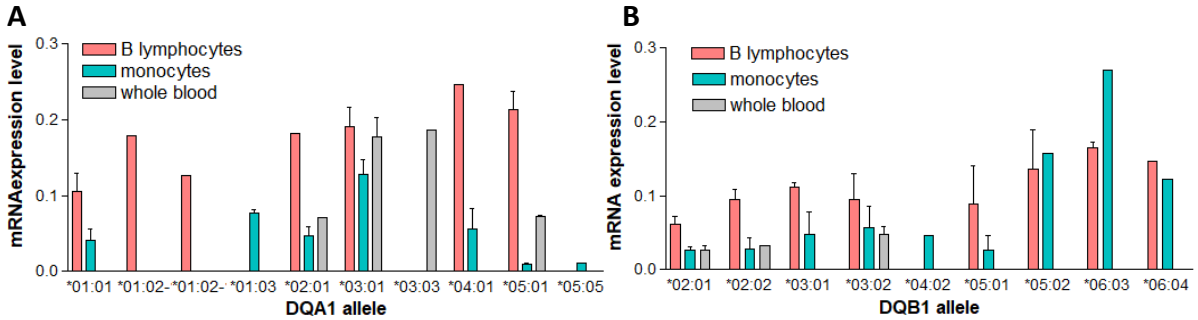


Figure 5. mRNA expression levels of *DQA1* (A) and *DQB1* (B) alleles in different cell types from T1D patients. Expression of *DQA1* (A) and *DQB1* (B) alleles was determined by RT-qPCR with allele-specific primers and the mean of 3 independent measurements was used for the analysis. SD is indicated.

in WBC. After Bonferroni correction ($p_c=p_{un}*4$), a single difference between expression of *DQB1*02:02* alleles in B cells was significant ($p_c=0.038$), with the expression higher in T1D patients.

Similarly to our previous studies, an allele expression ratio in heterozygotes was calculated. Again, the allele expression hierarchy was generally preserved in the heterozygous individuals.

Correlation of PPIA and DRA gene expression

Comparison of *DRA* expression normalized to *PPIA* between controls and patients (in WBC, control group from study A, group of patients from study (127)) did not reveal any differences between the two groups. The only notable difference were few outlying values in the patient group (**Figure 7**). We aimed to compare *DRA/PPIA* ratio between controls and patients as well, however, due to damaged batch of the *PPIA* assay which was recognized too late (and there was no opportunity to repeat the measurements), it was not possible.

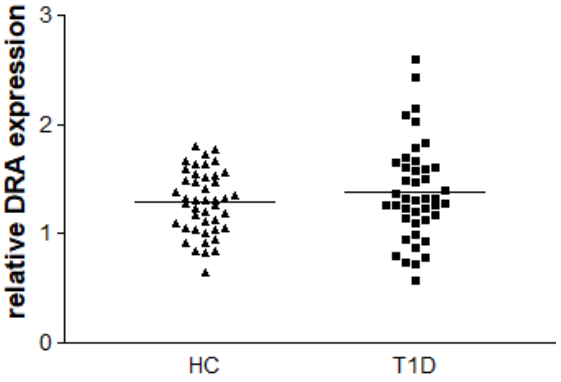


Figure 7. Relative *DRA* mRNA exoression in healthy controls (HC) and patients with T1D (T1D) in whole blood cells. Expression of *DRA* and *PPIA* gene was measured by RT-qPCR and the mean of 3 independent measurements was used for the analysis.

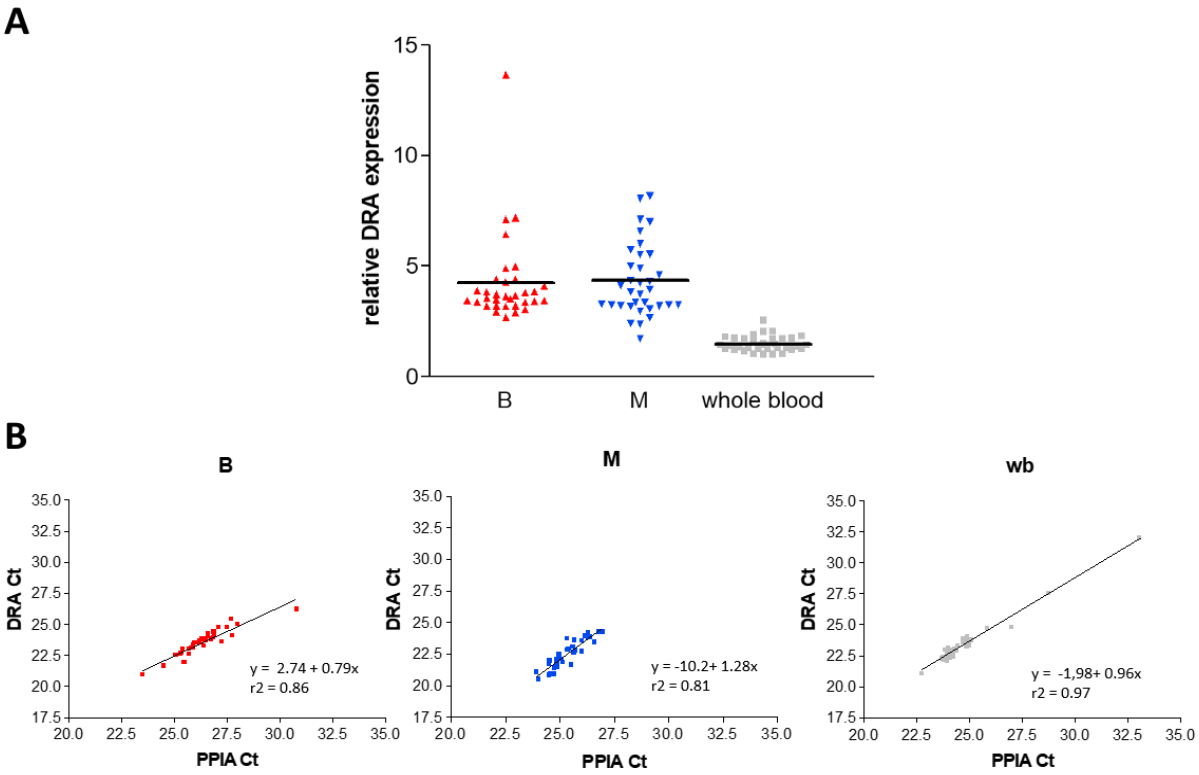


Figure 6. Expression of *DRA* mRNA in different cell types. Expression of *DRA* and *PPIA* gene was measured by RT-qPCR and the mean of 3 independent measurements was used for the analysis. (A) *DRA* mRNA levels normalized against *PPIA* in B lymphocytes (B), monocytes (M) and whole blood (wb). (B) Correlation between *PPIA* and *DRA* mRNA expression in different cell types. Ct values of *DRA* gene assay are plotted against Ct values of *PPIA* assay. Linear regression curves and r2 values are shown for each curve.

5. Discussion

5.1. DNA methylation of the polymorphic *HLA-DQA1* promoters and its effect on the mRNA expression

We did not observe the expected negative correlation between a number of methylated cytosines in promoter and allele expression.

The expression variability could possibly be caused by variable ability of the *DQA1* alleles to bind chromatin modifing methyl-CpG binding proteins. MBPs have specific requirements for a sequence surrounding the methylated cytosine (128–130). There are no MBD1 binding sites in any of the *DQA1* promoter alleles, but 1 Kaiso binding site in promoter of *DQA1*02:01* and *DQA1*01* group of alleles. It is interesting though that the least expressed alleles, *DQA1*05:01* and *DQA1*05:05*, have most theoretical MeCP2 binding sites (5 compared to 4 in *DQA1*03* and 3 in most *DQA1*01* alleles). As most of the CpG sites in all alleles are methylated (**Figure 1B**), these context differences have a chance to manifest. This could be one of the factors that add up and lead to low expression of the *DQA1*05* alleles. A pilot projects from our laboratory on histone modifications of the class II alleles hint low H3 acetylation of the *DQA1*05* group promoters (significantly (131) and insignificantly (132) lower than of the **01* group), which is in line with the low *DQA1*05* mRNA expression level observed here. Based on this, we could speculate that the combination of MBP binding sites in the **05* allele promoter make it especially prone to the silencing by DNA methylation, however, further experiments would be needed to confirm this (e.g., the ChIP studying the presence of the MBPs in the allele promoter). These pilot projects did not reveal correlation betweeen allele histone modification and mRNA expression – however, repeating these pilot experiments (to get measurements with lower SD), possibly in an extended genomic region (in proximal region, the specific signal presents very low percentage of input, which likely signifies an absence of histones in this region) may provide different results.

In all alleles we could observe a common pattern where most of the promoter CpG sites were methylated and as we approached the transcription start site, the methylation decreased to almost zero. DNA hypermethylation in this region is able to suppress the transcription of the *DQA1* gene even in the presence of CIITA and all other factors necessary for class II genes expression (38, 105). Thus it seems that in spite of a relatively low content of CpG dinucleotides in the *DQA1* gene promoter, high level of their methylation in an area very proximal to the transcription start site is able to decrease gene's expression and maybe this could be the site of the major regulation.

To conclude, the overall level of *DQA1* promoter methylation is not able to explain expression differences between alleles. We suggest that the genetic polymorphism in the region (especially region upstream of position -400, which is almost completely methylated in all alleles) may lead to different interpretation of the same 5meCpG mark in different allelic contex.

5.2. mRNA expression

5.2.1. mRNA expression of the *HLA-DQA1* alleles in the whole blood cells

Observed hierarchy of the *DQA1* alleles expression in monocytes and PBMC is in agreement with previously published studies (133–135) and the discrepancies can be explained by different methodology used (136).

Our results suggest that relative mRNA level of particular *DQA1* allele does not show much interindividual variability, tends to remain constant for a particular allele, and is not dependent on the second allele present in heterozygous individual.

5.2.2. mRNA expression of the *HLA-DQA1* alleles in B lymphocytes and monocytes

In the next phase, we analyzed the expression of *DQA1* and *DQB1* gene alleles in purified cell populations of CD14+ monocytes and CD19+ B lymphocytes. To provide a reference frame and compare data with the previous study, we again included whole blood cells into the analysis.

The ***DQA1*** expression hierarchy in the **whole blood** was **01:03*, **03:01* > **01:01* > **05:05* (statistically significant results) and *DQA1*03:01*, **03:03* > **01:03* > **01:02* (DR13, DR15-linked) > **01:01* >**02:01* > **05:01*, **05:05* (descriptive), which is in accordance with our pilot study on *DQA1* gene allele expression in whole blood and with other studies mentioned above.

Essentially the same expression hierarchy as for whole blood was observed in **monocytes** (**Figure 4A**). However, the situation was different in **B lymphocytes**: although *DQA1*03:01* was still highly expressed, it was surprising to see **05* group alleles among the most expressed alleles as well, as they were the least expressed ones in monocytes in our study, and were also consistently reported to be the lowest expressed allele group in various whole blood cell subsets by others (133, 135, 137). This discrepancy is likely to be caused by the actual difference in the *DQA1* allele expression regulation in B cells compared to mixed cell populations used in these studies, as the only other work that observed higher expression of *DQA1*05* compared to 01 allele was also done on B lymphocytes (92). The authors reported *DQA1*05:01* allele mRNA levels 1.9–4.6 higher compared to the **01:01* allele in heterozygous **01:01*/**05:01* B lymphocyte-derived lymphoblastoid cell lines from both celiac disease patients and the healthy controls. The same team observed identical *DQA1*01:01* > 05:01 relationship in monocyte-derived DCs (92), which is inconsistent with low expression of *DQA1*05* alleles observed in this study; however, we cannot exclude the possibility that during the maturation of monocyte derived DC, different alleles are induced to different levels (93) and expression ratio changes.

When different cell populations were compared, the *DQA1* allele expression was higher in B lymphocytes than in monocytes for each allele. Also, the expression of alleles in B lymphocytes was more balanced (maximum ratio between mean allele expression of 1.6), while in monocytes, the mean levels of the transcripts of most expressed allele *DQA1*03:03* were 10 times higher than those of the least expressed alleles of **05* group.

5.2.3. mRNA expression of the *HLA-DQB1* alleles in the whole blood cells, B lymphocytes and monocytes

We analyzed expression of 10 *DQB1* alleles belonging to *DQB1*02*, **03*, **05* and **06* allele groups. In all three cell populations examined, *DQB1*02* alleles were the least expressed ones, and *DQB1*06* alleles the most expressed (**Figure 4B**). This is in agreement with previous works (93, 135). The *DQB1*03:02* > **03:01* ratio and low expression of the *DQB1*02:01* allele is in contrast with findings of other groups (57, 92)

Relationship between the expression of *DQB1* alleles in different cell types was not so clear as was the case with *DQA1* gene. For the *DQB1*02* group, **03:01*, **03:02* and **05:01*, the allele expression was higher in B lymphocytes compared to monocytes, for *DQB1*03:03* and **06:02* expression was at the same level, and finally for *DQB1*06:03* and **06:04* alleles it was lower in B cells.

Similar to *DQA1* allele, the inter-allelic expression was more uniform in B cells (most/least expressed allele means ratio of 3.5) compared to monocytes (ratio of 15.4).

5.2.4. Cell-type specific regulation of HLA class II genes

The major observations on the cell-type specific expression are:

- For most *DQA1* and *DQB1* gene alleles, an expression in B lymphocytes is higher
- Inter-allelic expression differences are much less pronounced in B lymphocytes than in monocytes.

The first observation could be explained by a **regulatory protein present in B cells, but not in monocytes.** This function could be exerted by CIITA itself. **CIITA isoform III** which regulates constitutive class II expression in B lymphocytes has the highest transactivation

potential out of all isoforms, due to more efficient interaction with transcription machinery and promoter binding protein RFX (80), and could be the factor responsible for a higher expression observed in B lymphocytes.

The second observation, the one of distinct allele hierarchies in the B cells and monocytes is very intriguing**.** It implies that there must be **a transcription factor or coactivator that is both cell-type- and allele-specific** (and also isotype-specific, as it does not affect *DRA*). As discussed in the introduction, there are genetic differences between the promoter alleles that can affect the class II enhanceosome formation – possibly the **CIITA isoform III** could not only bind better to RFX, but also be more tolerant to variation in the enhanceosome structure, mitigating the effect of inter-allelic promoter variation on the transcription. However, this is just a hypothesis not explored in any study, and we did not follow up on this idea either. An alternative explanation involves a **transcription factor that would be more abundant in B cells than in monocytes**: saturation of all class II promoters in the presence of high protein concentration would lead to more allele-balanced expression in B cells; while a lower amount of the factor would lead to competition between promoters in monocytes. If this is the case, then **RFX5** protein is a potential candidate for such regulation, as its (mRNA) expression is approximately 3 to 4-fold higher in B cells compared to the monocytes (http://biogps.org/#goto=genereport&id=5993). All other basic subunits forming the enhanceosome complex (RFXAP, RFXANK, NF-YA, NF-YB, NF-YC) are expressed to similar extent in these cell types (http://biogps.org).

Therefore, it is possible that promoter polymorphisms, cell type-specific CIITA isoforms and maybe some other, yet unidentified, cell type-specific factors interact and generate observed complexity of HLA class II alleles expression. In conclusion, we described differences in expression of *DQA1* and *DQB1* gene alleles in whole blood cells, monocytes and B lymphocytes. Our findings support the idea that expression level of HLA class II alleles is the result of both, promoter-, and cell type-specific factors.

5.2.5. DRA/PPIA

The formula describing the linear relationship was somewhat different in each cell type, suggesting differential regulation of HLA class II expression in these cell types.

5.2.6. What are the advantages of using the *HLA-DRA* gene as an endogenous reference?

The proportion of HLA class II expressing cells in the whole blood can vary between different individuals and depends on the proportion of APCs among leukocytes and their activation state. Referring to classical endogenous control could lead to the false-positive finding of different *DQA1* (*DQB1*) allele expression between two samples only due to differences in the amount of HLA class II expressing cells or strength of inflammatory stimuli in the sample. As all HLA class II molecules are regulated coordinately, using another HLA class II molecule as an endogenous control eliminates this problem. *HLA-DRA* gene lacks typical class II polymorphism and interindividual variation in its expression should not be affected by genetic variation. As a result, the *DRA*-normalized expression of *DQA1* (*DQB1*) allele is proportionate purely to the strength of allele promoter (caused by *cis*-acting genetic and epigenetic differences). Using *DRA* as an endogenous control enables performing further expression studies on other HLA class II genes, normalize values to the *DRA* expression, and compare them with each other.

5.2.7. HLA class II autoimmunity relevant properties - example of T1D

Approximately half of the T1D risk can be attributed to the genetics factors, and HLA genes are responsible for a half of the genetic risk (138). The T1D predisposition can be carried independently by individual alleles as well as by whole haplotypes. Multiple risk

alleles can together form high-risk haplotypes as is *DRB1*04-DQA1*03:01-DQB1*03:02* (DR4-DQ8.1) and **03-*05:01-*02:01* (DR3-DQ2.5), and the highest risk is carried by the DR3/DR4 heterozygotes. The *DRB1*15-DQA1*01:02-DQB1*06:02* (DR15-DQ6.2) haplotype confers dominant protection (139).

When combining our results on DQ dimer expression and published results on DQ dimer stability (32) it seems that protective DQ dimers are intrinsically stable molecules with high expression level in B cells and monocytes and risk DQ dimers are unstable molecules with low expression. This is in line with theory that increased antigen presentation in thymus supports generation of protective Tregs (140). However, comparing OR (taken from (141)) of DQ dimers with their 'theoretical' expression rate did not reveal significant correlation of these parameters in B cells nor in monocytes (Spearman's correlation coeficient, data not shown). Evidence supports strong relationship between HLA class II mRNA and protein (46, 92), yet it would be interesting to actually measure surface expression of individual DQ dimers and then relate it to the risk for T1D.

It has been suggested that HLA-DQ6.2 can protect from diabetes by competing for the same antigen with the risk DQ8.1 and DQ2.5 molecules (142, 143). High expression of DQ6.2 observed in this study is in agreement with this theory: expression of DQ2.5 and DQ8.1 is around 0.3 and 0.7 of DQ6.2 expression in B cells, and around 0.4 and 0.9 of DQ6.2 expression in monocytes. These differences may become more pronounced when we consider that DQ6.2 is approximately 10-fold more stable than these molecules (32).

Comparison of the *DQA1* and *DQB1* allele mRNA expression in healthy controls and patients with T1D enables us to assess non-genetic factors influencing T1D development. Previously (127) we observed an increased expression of the *DQA1*02:01* allele in WBC from healthy *DQA1*02:01/*03* heterozygotes. *DQA1*02:01* is a part of neutral *DR7-DQA1*02:01-DQB1*02:02* haplotype, however, it is possible, that its increased expression in the carriers of risk haplotypes protected them from developing the disease (for example, by forming stable *trans*-dimers (32) with the DQα chain trascribed from high-risk *DQA1*03:01* and thus basically preventing function of the DQ8.1 dimer). However, in this study we observed an increased expression of the *DQB1*02:02* allele from the same haplotype in B cells from T1D patients. This is in contrast with the previous theory and we have no explanation for it.

Differences in histone modifications (including the *HLA-DQB1* gene) in T1D were shown between healthy people and T1D patients, however, some of them may be a result of a hyperglycaemia, not a cause of the disease (144–146).

The differences in DNA methylation of the *HLA-DOB*, *-DQA2* and *-DQA1* genes were shown as well (147–149), which could arise early in the proces, even before the disease manifestation (148, 149). However, comparison of DNA methylation of the *HLA-DQA1* gene between HLA-matched T1D patients and healthy controls in our laboratory revealed no difference in DNA methylation of the proximal promoter of this gene (127).

5.2.8. Study limitations

Whole blood cell subpopulations targeted in DNA methylation and mRNA expression analyses did not correspond
Did not analyze 5-hydroxymethylcytosine
Used low-throughput DNA sequencing method
Correlation of class II mRNA and protein expression
Did not consider non-HLA genes polymorphism

6. Conclusions

This thesis focused on allele-specific mRNA expression and promoter DNA methylation of HLA class II genes. The work was divided into two studies, each addressing different aspects of the HLA class II allele expression regulation:

Study A: HLA class II DNA methylation and its effect on the mRNA expression

Aim: Determine DNA methylation of individual alleles of the *HLA-DQA1* gene. Analyze inter-allelic differences in DNA methylation. Analyze the relationship between allele DNA methylation and mRNA expression.

Results: In the **study A**, we determined expression hierarchy and DNA methylation profile of proximal promoter of the *HLA-DQA1* alleles. We observed inter-allelic differences in the **overall methylation**, which was highest for the *DQA1*02:01* and **04:01* alleles and did not match their average expression level, and the low expression of the *DQA1*05:01* allele was not in accordance with its low methylation density. There was no correlation between number of methylated cytosines per allele and its expression. Frequency of methylation of the **individual CpG sites** did not differ between alleles, however, due to the high polymorphism of class II genes, the CpG sites were often located in different positions and sequence contexts which could translate into epigenetic differences between alleles and subsequently into differences in mRNA expression levels.

To conclude, based on our results, DNA methylation density of the *DQA1* promoter alleles does not regulate their expression. It is possible that allele expression is regulated by methylation of individual CpG sites in different allelic contexts, but this hypothesis would require furtner testing.

Study B: HLA class II mRNA expression (allele- and cell-type-specific expression)

Aim 1: Determine the mRNA expression hierarchy of HLA class II *DQA1* and *DQB1* alleles.

Aim 2: Determine expression hierarchy of the *DQA1* and *DQB1* alleles in the whole blood cells, B cells and monocytes.

In the **study B,** we determined the mRNA expression hierarchy of the individual alleles and described its variation between the three cell populations. Class II transcription level is higher (with the exception of *DQB1*06* alleles) and shows lower interallelic variation in B cells compared to monocytes. The *DQB1* expression hierarchy can be generalized into *DQB1*06* > **03*,**05* > **02* pattern in monocytes and *DQB1*06* > **02*, **03*, **05* in B cells. The *DQA1* expression hierarchy is *DQA1*03* > *01 ≥ **02* > **05* in monocytes and *DQA1*03*, **05* > *01, **02* in B cells. Because of the low number of samples for certain allele-cell type combinations, only some of these results were confirmed statistically. CIITA isoforms are discussed as the factors that could drive the cell-type specific expression of class II alleles.

Finally, a relevance of these findings to the type 1 diabetes as an example of autoimmune disease is discussed. If we simplify 'theoretical' amount of DQ dimers on the cell surface as an mRNA amount of the less expressed of its two chains, then the expression level by itself is not able to explain dimer association with T1D.

7. List of Publications

Publications which are part of the thesis:

Cepek P, **Zajacova M**, Kotrbova-Kozak A, Silhova E, Cerna M (2016): DNA methylation and mRNA expression of HLA-*DQA1* alleles in type 1 diabetes mellitus. Immunology. Jun;148(2):150-9. doi: 10.1111/imm.12593 PMID: 26854762. ***IF: 4.078***

Zajacova M, Kotrbova-Kozak A, Cerna M (2016): HLA-*DRB1*, -*DQA1* and -*DQB1* genotyping of 180 Czech individuals from the Czech Republic pop 3. Hum Immunol. Apr;77(4):365-6. doi: 10.1016/j.humimm.2016.02.003 PMID: 26867812. ***IF: 2.127***

Zajacova M, Kotrbova-Kozak A, Cepek P, Cerna M (2015): Differences in promoter DNA methylation and mRNA expression of individual alleles of the HLA class II *DQA1* gene. Immunol Lett 167(2):147-54. doi: 10.1016/j.imlet.2015.08.006 PMID: 26297837. ***IF: 2.512***

Zajacova M, Kotrbova-Kozak A, Cerna M (2018): Expression of HLA-*DQA1* and HLA-*DQB1* genes in B lymphocytes, monocytes and whole blood. Int J Immunogenet. In Press. ***IF: 1.093***

Publications which are not part of the thesis:

Cibulova A, **Zajacova M**, Fojtikova M, Stolfa J, Sedova L, Cejkova P, Lippert J, Arenberger P, Cerna M (2013): The HLA-Cw**06* allele and -1149 G/T polymorphism of extrapituitary promoter of PRL gene as a possible common genetic predisposing factors to psoriasis vulgaris and psoriatic arthritis in Czech population. Rheumatol Int. 33: 913-9. doi: 10.1007/s00296-012-2472-7 Rheumatology International. ***IF: 1.885***

8. References

1. Al-Daccak R, Mooney N, Charron D. 2004. *Curr. Opin. Immunol.* 16(1):108–13
2. Cerna M. 2008. *Wiener Medizinische Wochenschrift*. 158(1–2):2–12
3. Alfonso C, Karlsson L. 2000. *Annu. Rev. Immunol*. 18:113–42
4. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SGE. 2015. *Nucleic Acids Res.* 43(D1):D423–31
5. Schmid D, Pypaert M, Münz C. 2007. *Immunity*. 26(1):79–92
6. Perry JSA, Hsieh C-S. 2016. *Immunol. Rev.* 271(1):141–55
7. Rogers PR, Croft M. 1999. *J. Immunol.* 163:1205–13
8. Gottschalk RA, Hathorn MM, Beuneu H, Corse E, Dustin ML, et al. 2012. *Proc. Natl. Acad. Sci.* 109(3):881–86
9. Corse E, Gottschalk RA, Krogsgaard M, Allison JP. 2010. *PLoS Biol.* 8(9):e1000481
10. Sewell AK. 2012. *Nat. Rev. Immunol.* 12(9):669–77
11. Lever M, Lim H-S, Kruger P, Nguyen J, Trendel N, et al. 2016. *Proc. Natl. Acad. Sci. U. S. A.* 113(43):E6630–38
12. Gottschalk RA, Corse E, Allison JP. 2010. *J. Exp. Med.* 207(8):1701–11
13. Gallegos AM, Bevan MJ. 2004. *J. Exp. Med.* 200(8):1039–49
14. Bautista JL, Lio C-WJ, Lathrop SK, Forbush K, Liang Y, et al. 2009. *Nat. Immunol.* 10(6):610–17
15. Hosken NA, Shibuya K, Heath AW, Murphy KM, Garra AO. 1995. *J. Exp. Med.* 182(November):20–22
16. Carballido JM, Carballido-Perrig M, Terres G, Heusser CH, Blaser K. 1992. *Eur J Immunol*. 22:1357
17. Constant BS, Pfeiffer C, Woodard A, Pasqualini T, Bottomly K. 1995. *J. Exp. Med.* 182(November):0–5
18. DiMolfetto L, Neal HA, Wu A, Reilly C, Lo D. 1998. *Cell. Immunol.* 183(1):70–79
19. Iezzi G, Sonderegger I, Ampenberger F, Schmitz N, Marsland BJ, Kopf M. 2009. *Proc. Natl. Acad. Sci.* 106(3):876–81
20. Shiner EK, Holbrook BC, Alexander-Miller MA. 2014. *PLoS One*. 9(6):e100175
21. Swanberg M, Lidman O, Padyukov L, Eriksson P, Åkesson E, et al. 2005. *Nat. Genet.* 37(5):486–94
22. Alvaro-Benito M, Morrison E, Wieczorek M, Sticht J, Freund C. 2016. *Open Biol.* 6(8):160165
23. Harndahl M, Rasmussen M, Roder G, Dalgaard Pedersen I, Sørensen M, et al. 2012. *Eur. J. Immunol.* 42(6):1405–16
24. Lazarski CA, Chaves FA, Jenks SA, Wu S, Richards KA, et al. 2005. *Immunity*. 23(1):29–40
25. Hall FC, Rabinowitz JD, Busch R, Visconti KC, Belmares M, et al. 2002. *Eur. J. Immunol.* 32(3):662–70
26. Nelson CA, Petzold SJ, Unanue ER. 1994. *Nature*. 371(6494):250–52
27. Hwang I, Kim K, Choi S, Lomunova M. 2017. *Mol. Cells*. 40(1):24–36
28. Edwards LJ, Evavold BD. 2013. *J. Autoimmun.* 44:82–90
29. Busch DH, Pamer EG. 1998. *J. Immunol.* 160(9):4441–48
30. Sant AJ, Chaves FA, Jenks SA, Richards KA, Menges P, et al. 2005. *Immunol. Rev.* 207:261–78
31. Bondinas GP, Moustakas AK, Papadopoulos GK. 2007. *Immunogenetics*. 59(7):539–53
32. Miyadera H, Ohashi J, Lernmark Å, Kitamura T, Tokunaga K, et al. 2015. *J. Clin. Invest.* 125(1):275–91
33. Ettinger RA, Liu AW, Nepom GT, Kwok WW. 1998. *J. Immunol.* 161:6439–45
34. Mangalam AK, Taneja V, David CS. 2013. *J. Immunol.* 190(2):513–18
35. Kwok WW, Kovats S, Thurtle P, Nepom GT. 1993. *J. Immunol.* 150(6):2263–72
36. van Lummel M, van Veelen PA, Zaldumbide A, de Ru A, Janssen GMC, et al. 2012. *J. Biol. Chem.* 287(12):9514–24
37. Reinius B, Sandberg R. 2015. *Nat. Rev. Genet.* 16(11):653–64
38. Majumder P, Gomez JA, Chadwick BP, Boss JM. 2008. *J. Exp. Med.* 205(4):785–98
39. Neefjes J, Jongsma ML, Paul P, Bakke O. 2011. *Nat Rev Immunol*. 11(12):823–36
40. Busch R, Rinderknecht CH, Roh S, Lee AW, Harding JJ, et al. 2005. *Immunol. Rev.* 207(1):242–60
41. Denzin LK. 2013. *Front. Immunol.* 4:465
42. Hornell TMC, Burster T, Jahnsen FL, Pashine A, Ochoa MT, et al. 2006. *J. Immunol.* 176(6):3536–47
43. Rinderknecht CH, Roh S, Pashine A, Belmares MP, Patil NS, et al. 2010. *Immunology*. 131(1):18–32
44. Zhou Z, Reyes-Vargas E, Escobar H, Rudd B, Rockwood AL, et al. 2016. *Eur. J. Immunol.* 46(4):834–45
45. Bronson PG, Caillier S, Ramsay PP, McCauley JL, Zuvich RL, et al. 2010. *Hum. Mol. Genet.* 19(11):2331–40
46. Raj P, Rai E, Song R, Khan S, Wakeland BE, et al. 2016. *Elife*. 5:e12089
47. Vander Lugt B, Khan A a, Hackney J a, Agrawal S, Lesch J, et al. 2014. *Nat. Immunol.* 15(2):161–67
48. De Riva A, Varley MC, Bluck LJ, Cooke A, Deery MJ, Busch R. 2013. *J. Immunol.* 190(12):5961–71
49. Prevosto C, Usmani MF, McDonald S, Gumienny AM, Key T, et al. 2016. *PLoS One*. 11(8):1–25
50. The MHC sequencing consortium. 1999. *Nature*. 401(6756):921–23
51. Andersson G. 1998. *Front. Biosci.* 3:d739-45
52. Lenormand C, Bausinger H, Gross F, Signorino-Gelo F, Koch S, et al. 2012. *J. Immunol.* 188(8):3903–11

53. Reith W, LeibundGut-Landmann S, Waldburger J-M. 2005. *Nat. Rev. Immunol.* 5(10):793–806
54. Mischiati C, Feriotto G, Bianchi N, Rutigliano C, Giacomini P, Gambari R. 1999. *Biochimie*. 81(3):219–28
55. van den Elsen PJ. 2011. *Front. Immunol.* 2(OCT):1–9
56. Muhlethaler-Mottet A, Krawczyk M, Masternak K, Spilianakis C, Kretsovali A, et al. 2004. *J. Biol. Chem.* 279(39):40529–35
57. Beaty JS, West KA, Nepom GT. 1995. *Mol. Cell. Biol.* 15(9):4771–82
58. Benoist C, Mathis D. 1990. *Annu. Rev. Immunol.* 8:681–715
59. Reith W, Mach B. 2001. *Annu. Rev. Immunol*. 19(Figure 1):331–73
60. Reith W, Barras E, Satola S, Kobr M. 1989. . 86(June):4200–4204
61. Hasegawa SL, Sloan JH, Reith W, Mach B, Boss JM. 1991. *Nucleic Acids Res.* 19(6):1243–49
62. Morzycka-Wroblewska E, Munshi A, Ostermayer M, Harwood JI, Kagnoff MF. 1997. *Immunogenetics*. 45(3):163–70
63. Moreno CS, Beresford GW, Louis-plence P, Morris AC, Boss JM. 1999. . 10:143–51
64. Wang J, Zhuang J, Iyer S, Lin X-Y, Greven MC, et al. 2013. *Nucleic Acids Res.* 41(D1):D171–76
65. Angel P, Imagawa M, Chiu R, Stein B, Imbra RJ, et al. 1987. *Cell*. 49(6):729–39
66. Ono SJ, Bazil V, Levi BZ, Ozato K, Strominger JL. 1991. *Proc. Natl. Acad. Sci. U. S. A.* 88(10):4304–8
67. Vilen BJ, Cogswell JP, Ting JP. 1991. *Mol. Cell. Biol.* 11(5):2406–15
68. Zhu XS, Linhoff MW, Li G, Chin KC, Maity SN, Ting JP. 2000. *Mol. Cell. Biol.* 20(16):6051–61
69. Robinson J, Malik A, Parham P, Bodmer J, Marsh S. 2000. *Tissue Antigens*. 55:280–87
70. Indovina P, Megiorni F, Ferrante P, Apollonio I, Petronzelli F, Mazzilli MC. 1998. *Hum. Immunol.* 59(12):758–67
71. Adams BS, Cha HC, Cleary J, Haiying T, Wang H, et al. 2003. *Arthritis Res. Ther.* 5(4):226–33
72. Morzycka-Wroblewska E, Harwood JI, Smith JR, Kagnoff MF. 1993. *Immunogenetics*. 37(5):364–72
73. Westerheide SD, Boss JM. 1999. . 27(7):1635–41
74. Otten LA, Steimle V. 1998. *Eur. J. Immunol.* 28:473–78
75. Douhan J, Lieberson R, Knoll JH, Zhou H, Glimcher LH. 1997. *J. Exp. Med.* 185(11):1885–95
76. Cazalis M, Friggeri A, Cavé L, Demaret J, Barbalat V, et al. 2013. *Crit. Care*. 17(6):R287
77. LeibundGut-Landmann S, Waldburger J-M, Krawczyk M, Otten LA, Suter T, et al. 2004. *Eur. J. Immunol.* 34(6):1513–25
78. Muhlethaler-Mottet A, Otten LA, Steimle V, Mach B. 1997. *EMBO J.* 16(10):2851–60
79. Wong AW, Ghosh N, McKinnon KP, Reed W, Piskurich JF, et al. 2002. *J. Immunol.* 169(6):3112–19
80. Beaulieu YB, Leon Machado JA, Ethier S, Gaudreau L, Steimle V. 2016. *PLoS One*. 11(2):e0148753
81. Masternak K, Peyraud N, Krawczyk M, Barras E, Reith W. 2003. *Nat. Immunol.* 4(2):132–37
82. Krawczyk M, Peyraud N, Rybtsova N, Masternak K. 2004. *J Immunol.* 173:6200–6210
83. Gomez JA, Majumder P, Nagarajan UM, Boss JM. 2005. *J. Immunol.* 175(2):1030–40
84. Ghirlando R, Felsenfeld G. 2016. *Genes Dev.* 30(8):881–91
85. de Wit E, Vos ESM, Holwerda SJB, Valdes-Quezada C, Verstegen MJAM, et al. 2015. *Mol. Cell*. 60(4):676–84
86. Majumder P, Boss JM. 2010. *Mol. Cell. Biol.* 30(17):4211–23
87. Majumder P, Gomez JA, Boss JM. 2006. . 281(27):18435–43
88. Hoarau J-J, Cesari M, Caillens H, Cadet F, Pabion M. 2004. *Tissue Antigens*. 63(1):58–71
89. Senju S, Kimura A, Yasunami M, Kamikawaji N, Yoshizumi H, Sasazuki T. 1992. *Immunogenetics*. 36(5):319–25
90. Harton J, Jin L, Hahn A, Drake J. 2016. *F1000Research*. 5(0):368
91. Briata P, Radka SF, Sartoris S, Lee JS. 1989. *Proc. Natl. Acad. Sci. U. S. A.* 86(3):1003–7
92. Pisapia L, Camarca A, Picascia S, Bassi V, Barba P, et al. 2016. *J. Autoimmun.* 70:63–72
93. Cesari M, Caillens H, Cadet F, Pabion M, Cesari M, et al. 1999. *Immunogenetics*. 50:309–18
94. Klose RJ, Bird AP. 2006. *Trends Biochem. Sci.* 31(2):89–97
95. Vaissiere T, Sawan C, Herceg Z. 2008. *Mutat. Res.* 659:40–48
96. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, et al. 2007. *Nat. Genet.* 39(4):457–66
97. Hural JA, Kwan M, Henkel G, Benjamin M, Brown MA, et al. 2000. *J. Immunol.* 165(6):3239–49
98. Cheng X, Blumenthal RM. 2008. *Structure*. 16(3):341–50
99. Cheng X, Blumenthal RM. 2010. *Biochemistry*. 49(14):2999–3008
100. Hendrich B, Tweedie S. 2003. *Trends Genet.* 19(5):269–77
101. Hervouet E, Peixoto P, Delage-Mourroux R, Boyer-Guittaut M, Cartron P-F. 2018. *Clin. Epigenetics*. 10:17
102. Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, et al. 2008. *Nat. Genet.* 40(7):904–8
103. Zhang Y, Rohde C, Reinhardt R, Voelcker-Rehage C, Jeltsch A. 2009. *Genome Biol.* 10(12):R138
104. Ghotbi R, Gomez A, Milani L, Tybring G, Syvänen A-C, et al. 2009. *Pharmacogenomics J.* 9(3):208–17
105. Majumder P, Boss JM. 2011. *Genes Immun.* 12(4):291–99
106. Feriotto G, Pozzi L, Ciucci A, Camarda G, Mischiati C, et al. 1993. *Eur. J. Biochem.* 218(2):485–92
107. Barbieri R, Nastruzzi C, Volinia S, Villa M, Piva R, et al. 1990. *J. Immunogenet.* 17(1–2):51–66
108. Gambari R, Barbieri R, Piva R, Tecce R, Fisher PB, et al. 1987. *Ann. N. Y. Acad. Sci.* 511:292–307
109. Carrington MN, Salter RD, Cresswell P, Ting JP. 1985. *Immunogenetics*. 22(3):219–29
110. Wang Y, Peterlin BM. 1986. *Immunogenetics*. 24(5):298–303
111. Lee JS, O'Neill L. 1987. *Immunogenetics*. 26(1–2):92–98
112. Levine F, Pious D. 1985. *Immunogenetics*. 22(5):427–40
113. Uitterlinden AG, Vijg J, Giphart MJ, Knook DL. 1985. *Exp. Clin. Immunogenet.* 2(4):215–22
114. Toyoda H, Redford A, Magalong D. 1992. *Dis. Markers*. 10(1):7–18
115. Taudt A, Colomé-tatché M, Johannes F. 2016. *Nat. Rev. Genet.* 17(6):319–32
116. Zentner GE, Henikoff S. 2013. *Nat. Struct. Mol. Biol.* 20(3):259–66
117. Rose NR, Klose RJ. 2014. *Biochim. Biophys. Acta*. 1839(12):1362–72
118. Kasowski M, Kyriazopoulou-panagiotopoulou S, Grubert F, Judith B, Kundaje A, et al. 2013. *Science*. 342(6159):750–52
119. McVicker G, van de Geijn B, Degner JF, Cain CE, Banovich NE, et al. 2013. *Science*. 342(6159):747–49
120. McDaniell R, Lee B, Song L, Liu Z, Boyle AP, et al. 2010. *Science*. 328(5975):235–39
121. Zajacova M, Kotrbova-Kozak A, Cepek P, Cerna M. 2015. *Immunol. Lett.* 167(2):147–54
122. Zajacova M, Kotrbova-Kozak A, Cerna M. 2018. *Int. J. Immunogenet.* In Press:
123. Zajacová M. 2010. *Assessment of expression markers of HLA class II genes*. Charles University in Prague, Faculty of Science
124. Zajacova M, Kotrbova-Kozak A, Cerna M. 2016. *Hum. Immunol.* 77(4):365–66
125. Gonzalez-Galarza FF, Takeshita LYC, Santos EJM, Kempson F, Maia MHT, et al. 2015. *Nucleic Acids Res.* 43(D1):D784–88
126. Brünnler G, Haas J, Fan L, Petzl-Erler M, Volgger A, et al. 1997. In *Proceedings of the Twelfth International Histocompatibility Workshop and Conference, Vol I: Workshop*. 1:171–75

127. Cepek P, Zajacova M, Kotrbova-Kozak A, Silhova E, Cerna M. 2016. *Immunology*. 148(2):150–59
128. Klose RJ, Sarraf SA, Schmiedeberg L, McDermott SM, Stancheva I, Bird AP. 2005. *Mol. Cell*. 19(5):667–78
129. Clouaire T, de Las Heras JI, Merusi C, Stancheva I. 2010. *Nucleic Acids Res.* 38(14):4620–34
130. Prokhortchouk A, Hendrich B, Jørgensen H, Ruzov A, Wilm M, et al. 2001. *Genes Dev.* 15(13):1613–18
131. Chmel M. 2015. *The study of epigenetic regulation of gene HLA II. Class within family relationships*. Charles University in Prague, Faculty of Science
132. Cernoch M. 2014. *The study of epigenetic regulation of HLA class II genes at the level of histone modifications*. Charles University in Prague, Faculty of Science
133. Donner H, Seidl C, Rau H, Herwig J, Seifried E, et al. 2002. *Eur. J. Immunogenet.* 29(4):321–30
134. Maffei A, Pozzo GD, Prisco A, Ciullo M, Harris PE, et al. 1997. *Hum. Immunol.* 53(2):167–73
135. Britten AC, Mijovic CH, Barnett AH, Kelly MA. 2009. *Int. J. Immunogenet.* 36(1):47–57
136. Fernandez S, Wassmuth R, Knerr I, Frank C, Haas JP. 2003. *Eur. J. Immunogenet.* 30(2):141–48
137. Maffei A, Harris PE, Reed EF, Del Pozzo G, Ciullo M, et al. 1997. *Eur. J. Immunol.* 27:1549–56
138. Polychronakos C, Li Q. 2011. *Nat. Rev. Genet.* 12(11):781–92
139. Jones EY, Fugger L, Strominger JL, Siebold C. 2006. *Nat. Rev. Immunol.* 6(4):271–82
140. Unanue ER. 2014. *Antigen Presentation in the Autoimmune Diabetes of the NOD Mouse*, Vol. 32. 579-608 pp.
141. Erlich H, Valdes AM, Noble J, Carlson JA, Varney M, et al. 2008. . 57(4):1084–92
142. Nepom GT. 1990. *Diabetes*. 39(10):1153–57
143. Parry CS, Brooks BR. 2008. *Biol. Direct*. 3:42
144. Miao F, Chen Z, Zhang L, Liu Z, Wu X, et al. 2012. *J. Biol. Chem.* 287(20):16335–45
145. Miao F, Gonzalo IG, Lanting L, Natarajan R. 2004. *J. Biol. Chem.* 279(17):18091–97
146. Miao F, Wu X, Zhang L, Yuan YC, Riggs AD, Natarajan R. 2007. *J. Biol. Chem.* 282(18):13854–63
147. Stefan M, Zhang W, Concepcion E, Yi Z, Tomer Y. 2014. *J. Autoimmun.* 50:33–37
148. Rakyan VK, Hildmann T, Novik KL, Lewin J, Tost J, et al. 2004. *PLoS Biol.* 2(12):e405
149. Disanto G, Vcelakova J, Pakpoor J, Elangovan RI, Sumnik Z, et al. 2013. *Diabetologia*. 56(9):2093–95