



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ**  
**ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ**  
ΤΜΗΜΑ ΙΑΤΡΙΚΗΣ

---

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ  
ΣΠΟΥΔΩΝ « ΠΛΗΡΟΦΟΡΙΚΗ ΕΠΙΣΤΗΜΩΝ ΖΩΗΣ »

**ΜΕΤΑΠΤΥΧΙΑΚΗ**  
**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**« Εκτίμηση συχνότητας των HLA-A, -B, -DRB1  
αλληλόμορφων και απλοτύπων σε σύνολο δεδομένων  
5.131 Ελλήνων εθελοντών δοτών μυελού των οστών και  
Εφαρμογές για στρατηγικό σχεδιασμό μητρώου δοτών »**

---

**ΠΑΝΑΓΙΩΤΗΣ Κ. ΚΟΡΜΠΑΚΗΣ**  
*A.M. : 1951*

**ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ :**

**ΑΛΕΞΑΝΔΡΟΣ ΣΠΥΡΙΔΩΝΙΔΗΣ** - Αναπληρωτής Καθηγητής

*Πάτρα , Ιούνιος 2015*



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ**  
**ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ**  
**ΤΜΗΜΑ ΙΑΤΡΙΚΗΣ**

---

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ  
ΣΠΟΥΔΩΝ « ΠΛΗΡΟΦΟΡΙΚΗ ΕΠΙΣΤΗΜΩΝ ΖΩΗΣ »

**ΜΕΤΑΠΤΥΧΙΑΚΗ**  
**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**« Εκτίμηση συχνότητας των HLA-A, -B, -DRB1  
αλληλόμορφων και απλοτύπων σε σύνολο δεδομένων  
5.131 Ελλήνων εθελοντών δοτών μυελού των οστών και  
Εφαρμογές για στρατηγικό σχεδιασμό μητρώου δοτών »**

---

**ΠΑΝΑΓΙΩΤΗΣ Κ. ΚΟΡΜΠΑΚΗΣ**

**A.M. : 1951**

**ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ :**

**ΑΛΕΞΑΝΔΡΟΣ ΣΠΥΡΙΔΩΝΙΔΗΣ - Αναπληρωτής Καθηγητής**

*Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την .....*

.....  
**ΑΛΕΞΑΝΔΡΟΣ ΣΠΥΡΙΔΩΝΙΔΗΣ**  
*Αναπληρωτής Καθηγητής Ιατρικής*

.....  
**ΑΘΑΝΑΣΙΑ ΜΟΥΖΑΚΗ**  
*Καθηγήτρια Ιατρικής*

.....  
**ΓΕΩΡΓΙΟΣ ΑΝΔΡΟΥΛΑΚΗΣ**  
*Αναπληρωτής Καθηγητής  
Διοίκησης Επιχειρήσεων*

*Πάτρα , Ιούνιος 2015*

# ΜΕΤΑΠΤΥΧΙΑΚΗ

## ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**« Εκτίμηση συχνότητας των HLA-A, -B, -DRB1  
αλληλόμορφων και απλοτύπων σε σύνολο δεδομένων  
5.131 Ελλήνων εθελοντών δοτών μυελού των οστών και  
Εφαρμογές για στρατηγικό σχεδιασμό μητρώου δοτών »**

---

**ΠΑΝΑΓΙΩΤΗΣ Κ. ΚΟΡΜΠΑΚΗΣ**

**A.M. : 1951**

.....  
ΠΑΝΑΓΙΩΤΗΣ Κ. ΚΟΡΜΠΑΚΗΣ

*«Η ολοκλήρωση της εργασίας αυτής έγινε στο πλαίσιο της υλοποίησης του διατμηματικού μεταπτυχιακού προγράμματος ΠΛΗΡΟΦΟΡΙΚΗ ΕΠΙΣΤΗΜΩΝ ΖΩΗΣ του τμήματος Ιατρικής του Πανεπιστημίου Πατρών, το οποίο συγχρηματοδοτήθηκε μέσω της Πράξης «Πρόγραμμα χορήγησης υποτροφιών ΙΚΥ με διαδικασία εξατομικευμένης αξιολόγησης ακαδημαϊκού έτους 2012-2013» από πόρους του Ε.Π. «Εκπαίδευση και Δια Βίου Μάθηση» του Ευρωπαϊκού Κοινωνικού Ταμείου (ΕΚΤ) και του ΕΣΠΑ (2007-2013)»*



---

## Π Ρ Ο Λ Ο Γ Ο Σ

---

Η παρούσα μεταπτυχιακή διπλωματική εργασία εκπονήθηκε στο Κέντρο Ενημέρωσης & Προσέλευσης Εθελοντών Δοτών Μυελού των Οστών (ΚΕΔΜΟΠ) στα πλαίσια του διατμηματικού προγράμματος μεταπτυχιακών σπουδών «ΠΛΗΡΟΦΟΡΙΚΗ ΕΠΙΣΤΗΜΩΝ ΖΩΗΣ» του Τμήματος Ιατρικής του Πανεπιστημίου Πατρών.

Σκοπός της εργασίας είναι η διαχείριση, η επεξεργασία και η εφαρμογή μεθοδολογιών βιοστατιστικής ανάλυσης σε βιολογικά δεδομένα του HLA γονιδιακού συστήματος, το οποίο είναι υπεύθυνο για την ανάλυση της ιστοσυμβατότητας κατά την μεταμόσχευση αιμοποιητικών κυττάρων μυελού των οστών. Έτσι, αρχικά έγινε αναφορά σε ζητήματα που σχετίζονται με τη μεταμόσχευση αιμοποιητικών κυττάρων και την ιστοσυμβατότητα μεταξύ δοτών και ληπτών μοσχευμάτων και παρουσιάστηκαν τα βασικά τους χαρακτηριστικά. Επιπλέον, παρουσιάστηκαν βασικές παράμετροι και μετρικές βιοστατιστικής ανάλυσης HLA δεδομένων, τα οποία περιλαμβάνονται σε εξειδικευμένα μητρώα εθελοντών δοτών μυελού των οστών σε διεθνές και εθνικό επίπεδο, καθώς επίσης μελετήθηκαν τεχνικές διαχείρισης και ανάλυσης των δεδομένων των συγκεκριμένων μητρώων με ειδικά εργαλεία λογισμικού. Τέλος, υλοποιήθηκε επεξεργασία και ανάλυση των HLA δεδομένων 5.131 Ελλήνων εθελοντών δοτών μυελού των οστών που είχαν εγγραφεί στο μητρώο του Κέντρου Ενημέρωσης & Προσέλευσης Εθελοντών Δοτών Μυελού των Οστών (ΚΕΔΜΟΠ) του Πανεπιστημίου Πατρών και παρουσιάστηκαν τα αποτελέσματα της ανάλυσης ως προς την εκτίμηση των συχνοτήτων των αλληλόμορφων γονιδίων HLA-A, HLA-B και HLA-DRB1 και των απλοτύπων που προκύπτουν από τον συνδυασμό τους, με σκοπό την σύγκριση των HLA χαρακτηριστικών γνωρισμάτων του Ελληνικού πληθυσμού με άλλους πληθυσμούς και εφαρμογές για στρατηγικό σχεδιασμό μητρώου εθελοντών δοτών μυελού των οστών.

---

## ΕΥΧΑΡΙΣΤΙΕΣ

---

Στο σημείο αυτό, θα ήθελα να εκφράσω τις ειλικρινείς μου ευχαριστίες σε όλους όσους συνέβαλαν στην ολοκλήρωση της εκπόνησης της παρούσας μεταπτυχιακής διπλωματικής εργασίας.

Ιδιαίτερα θα ήθελα να ευχαριστήσω τον επιβλέποντα της εργασίας αυτής, αναπληρωτή καθηγητή Αιματολογίας / Μεταμόσχευσης Αιμοποιητικών Κυττάρων του Τμήματος Ιατρικής του Πανεπιστημίου Πατρών και Υπεύθυνο της Μονάδος Μεταμόσχευσης Μυελού των Οστών του ΠΓΝΠ. κ. Αλέξανδρο Σπυριδωνίδη για την πολύτιμη βοήθειά του και τη διαρκή υποστήριξή του, τόσο κατά τη διεξαγωγή της έρευνας όσο και κατά τη συγγραφή της παρούσας εργασίας, καθώς επίσης την υπεύθυνη του Τμήματος Αρχείου & Διακίνησης Δειγμάτων κ. Γεωργία Οικονομοπούλου και την υπεύθυνη του Τμήματος Προσέλκυσης, Δημοσιότητας & Εκδηλώσεων του ΚΕΔΜΟΠ κ. Αλεξάνδρα Σαρμά για την συνεργασία κατά την συλλογή και επεξεργασία των δεδομένων του μητρώου των εθελοντών δοτών μυελού των οστών του ΚΕΔΜΟΠ.

Τέλος, θα ήθελα να εκφράσω τις ευχαριστίες μου στην οικογένειά μου για την αμέριστη συμπαράσταση, βοήθεια και προ πάντων κατανόηση και ανοχή καθ' όλο το χρονικό διάστημα των σπουδών μου.

*Πάτρα, Ιούνιος 2015*

*ΠΑΝΑΓΙΩΤΗΣ ΚΟΡΜΠΑΚΗΣ*

---

## Π Ε Ρ Ι Λ Η Ψ Η

---

Η ανάλυση της συμβατότητας των HLA (Human Leukocyte Antigen - Ανθρώπινων Λευκοκυτταρικών Αντιγόνων) γονιδίων μεταξύ δότη και λήπτη βελτιώνει την επιβίωση των ασθενών μετά από μεταμόσχευση αιμοποιητικών βλαστικών κυττάρων. Κατά συνέπεια, η αναζήτηση του κατάλληλου δότη έχει στόχο την εξεύρεση δοτών οι οποίοι είναι συμβατοί με τους αντίστοιχους ασθενείς τουλάχιστον για αυτά τα γονίδια, τα οποία έχουν τυποποιηθεί σε υψηλή ανάλυση. Σε αυτό το πλαίσιο, η διαθεσιμότητα των κατανομών των συχνοτήτων των HLA αλληλόμορφων γονιδίων και των αντίστοιχων απλοτύπων που προκύπτουν από τον συνδυασμό τους είναι χρήσιμη για δύο λόγους: Πρώτον, αυξάνεται η αποδοτικότητα των επιμέρους αναζητήσεων συμβατού δότη, εφόσον πολλοί εγγεγραμμένοι δότες στο μητρώο εθελοντών δοτών μυελού των οστών παρατίθενται με ελλειπίες HLA πληροφορίες. Δεύτερον, οι κατανομές των συχνοτήτων των HLA απλοτύπων και η βιοστατιστική ανάλυση των HLA δεδομένων του πληθυσμού των δοτών, καθορίζουν πόσοι ασθενείς βρίσκουν συμβατούς δότες σε ένα μητρώο με  $N$  εθελοντές δότες, από την ανάλυση του πληθυσμού. Στα πλαίσια της εκπόνησης της παρούσας μεταπτυχιακής διπλωματικής εργασίας θα προσδιοριστεί η κατανομή των συχνοτήτων των HLA αλληλόμορφων γονιδίων και των HLA απλοτύπων σε έναν πληθυσμό Ελλήνων εθελοντών δοτών μυελού των οστών για τα γονίδια HLA-A, HLA-B και HLA-DRB1, η οποία εν συνεχεία θα χρησιμοποιηθεί για τη βελτιστοποίηση της αναζήτησης μεμονωμένων βλαστικών κυττάρων των δοτών και την αποδοτική διαχείριση του αντίστοιχου μητρώου των εθελοντών δοτών. Επιπλέον, οι συχνότητες των HLA απλοτύπων που θα υπολογιστούν, θα χρησιμοποιηθούν για την εκτίμηση της πιθανότητας εύρεσης συμβατών δοτών εντός του συνόλου των εθελοντών δοτών. Οι αναλύσεις και η επεξεργασία του συνόλου των δεδομένων των εθελοντών δοτών θα υλοποιηθούν με ειδικά πακέτα λογισμικού βιοστατιστικής ανάλυσης, υπολογιστικούς αλγόριθμους και μεθοδολογίες. Τέλος, θα πραγματοποιηθεί μια προσπάθεια σύγκρισης των στατιστικών αποτελεσμάτων που θα προκύψουν από την ανάλυση των δεδομένων του πληθυσμού των Ελλήνων δοτών, με τα αντίστοιχα αποτελέσματα άλλων πληθυσμών (Γερμανών δοτών) για την εξαγωγή συμπερασμάτων ως προς το ποσοστό της HLA ιστοσυμβατότητας μεταξύ των πληθυσμών αυτών και την ποιότητα των μητρώων.

**ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ :** Μεταμόσχευση Αιμοποιητικών Βλαστικών Κυττάρων, Μυελός των Οστών, Ιστοσυμβατότητα, MHC, HLA, NMDP, Ανισορροπία Σύνδεσης, Γενετική Ισορροπία, Hardy-Weinberg, Arlequin, Αλγόριθμος EM, Αλληλόμορφο, Γονίδιο, Αντιγόνο, Απλότυπος, Γονότυπος, Φαινότυπος, Μητρώο Εθελοντών Δοτών, Συχνότητα, Βιοστατιστική.

---

# A B S T R A C T

---

*The analysis of compatibility of HLA (Human Leukocyte Antigens) genes between donor and recipient improves survival of patients after unrelated hematopoietic stem cell transplantation. Consequently, the search for a suitable donor is aimed at finding donors who match with the corresponding patients at least for these genes, which are formulated at high resolution. In this context the availability of frequency distributions of HLA alleles and the corresponding haplotypes as a result from their combination is useful for two reasons: First, increasing the efficiency of individual compatible donor searches, as many registered donors are listed only with incomplete HLA information. Second, frequency distributions of HLA haplotypes and biostatistical analysis of HLA data of donors, determine how many patients find compatible donors in a registry with N volunteer donors from the analyzed population. As part of the preparation of this dissertation, we will determine the frequency distribution of HLA alleles and HLA haplotypes in a population of Greeks bone marrow volunteer donors for HLA-A, HLA-B and HLA-DRB1 genes, in order to be used to optimize the search of individual stem cell donors and for the efficient management of the corresponding registry of volunteer donors. Moreover, the frequencies of HLA haplotypes that will be calculated, will be used to estimate the probability of finding a compatible donor within all of volunteer donors. This analysis and processing of all the data of volunteer donors will be implemented with special software packages of biostatistic analysis, computational algorithms and methodologies. Finally, there will be an attempt of comparison of statistical results obtained by analyzing the data of the population of Greek donors, with the corresponding results of other populations (German donors) to draw conclusions about the percentage of HLA histocompatibility between these populations and the quality of donor registries.*

**Key Words :** Hematopoietic Stem Cell Transplantation, Bone Marrow, Histocompatibility, MHC, HLA, NMDP, Linkage Disequilibrium, Genetic Equilibrium, Hardy-Weinberg, Arlequin, EM Algorithm, Allele, Gene, Antigen, Haplotype, Genotype, Phenotype, Volunteer Donor Registry, Frequency, Biostatistics.

---

# Π Ε Ρ Ι Ε Χ Ο Μ Ε Ν Α

---

|                                                                          |           |
|--------------------------------------------------------------------------|-----------|
| ΠΡΟΛΟΓΟΣ .....                                                           | i         |
| ΕΥΧΑΡΙΣΤΙΕΣ.....                                                         | ii        |
| ΠΕΡΙΛΗΨΗ .....                                                           | iii       |
| ΠΕΡΙΕΧΟΜΕΝΑ .....                                                        | v         |
| ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ .....                                                  | vii       |
| ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ.....                                                   | ix        |
| <b>1 ΕΙΣΑΓΩΓΗ .....</b>                                                  | <b>1</b>  |
| 1.1 Αντικείμενο της διπλωματικής εργασίας.....                           | 1         |
| 1.2 Διάρθρωση της διπλωματικής εργασίας.....                             | 3         |
| <b>2 ΜΕΤΑΜΟΣΧΕΥΣΗ ΑΙΜΟΠΟΙΗΤΙΚΩΝ ΚΥΤΤΑΡΩΝ &amp; ΙΣΤΟΣΥΜΒΑΤΟΤΗΤΑ .....</b> | <b>6</b>  |
| 2.1 Μυελός των Οστών .....                                               | 6         |
| 2.1.1 Είδη Μυελού των Οστών.....                                         | 7         |
| 2.1.2 Λειτουργία Μυελού των Οστών .....                                  | 7         |
| 2.2 Μεταμόσχευση Αιμοποιητικών Κυττάρων .....                            | 8         |
| 2.2.1 Αιμοποιητικά Βλαστικά Κύτταρα (HSC).....                           | 8         |
| 2.2.2 Πηγές Αιμοποιητικών Βλαστικών Κυττάρων.....                        | 9         |
| 2.2.3 Μεταμόσχευση Αιμοποιητικών Βλαστικών Κυττάρων (HSCT).....          | 10        |
| 2.2.4 Κατηγορίες Μεταμόσχευσης Αιμοποιητικών Βλαστικών Κυττάρων.....     | 12        |
| 2.3 Συμπλέγματα Ιστοσυμβατότητας .....                                   | 14        |
| 2.3.1 Μείζον Σύμπλεγμα Ιστοσυμβατότητας (MHC).....                       | 14        |
| 2.3.2 Ελάσσονα Συστήματα Ιστοσυμβατότητας .....                          | 15        |
| 2.3.3 Οργάνωση HLA γονιδίων.....                                         | 16        |
| 2.3.4 Δομή HLA γονιδίων .....                                            | 20        |
| 2.3.5 Λειτουργία HLA μορίων .....                                        | 22        |
| 2.3.6 Κληρονομική Μεταβίβαση .....                                       | 24        |
| 2.3.7 Κλινικές Εφαρμογές HLA Συστήματος .....                            | 25        |
| 2.3.7.1 Μεταμοσχεύσεις Ιστών & Οργάνων.....                              | 25        |
| 2.3.7.2 Έλεγχος Αμφισβητούμενης Πατρότητας .....                         | 27        |
| 2.3.7.3 Ανθρωπολογικές Μελέτες & Φυλογενετική Ανάλυση Πληθυσμών .....    | 27        |
| 2.3.7.4 Συσχετισμοί με Ασθένειες .....                                   | 28        |
| 2.3.8 Πολυμορφισμός HLA Συστήματος.....                                  | 29        |
| 2.3.8.1 Χαρακτηριστικά Πολυμορφισμού.....                                | 29        |
| 2.3.8.2 Βιολογική Σημασία Πολυμορφισμού.....                             | 30        |
| 2.3.9 Τυποποίηση HLA Συστήματος.....                                     | 30        |
| 2.3.10 Ονοματολογία HLA Συστήματος .....                                 | 32        |
| <b>3 ΣΤΟΙΧΕΙΑ ΒΙΟΣΤΑΤΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ HLA ΔΕΔΟΜΕΝΩΝ .....</b>            | <b>36</b> |
| 3.1 Έλεγχος Γενετικής Ισορροπίας Hardy-Weinberg.....                     | 36        |
| 3.2 Ανισορροπία Σύνδεσης (LD - Linkage Disequilibrium) .....             | 39        |

|          |                                                                                                                                         |            |
|----------|-----------------------------------------------------------------------------------------------------------------------------------------|------------|
| 3.2.1    | Ποσοτικοποίηση Ανισορροπίας Σύνδεσης .....                                                                                              | 39         |
| 3.3      | Γενετική Απόσταση (Genetic Distance) .....                                                                                              | 41         |
| 3.4      | Εκτίμηση Συχνοτήτων HLA Απλοτύπων .....                                                                                                 | 41         |
| 3.5      | Ο Αλγόριθμος Μεγιστοποίησης Προσδοκίας (EM – Expectation-Maximization) .....                                                            | 43         |
| 3.5.1    | Περιγραφή Αλγορίθμου EM .....                                                                                                           | 44         |
| 3.5.2    | Ιδιότητες Αλγορίθμου EM .....                                                                                                           | 46         |
| <b>4</b> | <b>ΔΙΑΧΕΙΡΙΣΗ ΜΗΤΡΩΩΝ ΕΘΕΛΟΝΤΩΝ ΔΟΤΩΝ ΜΥΕΛΟΥ ΤΩΝ ΟΣΤΩΝ<br/>&amp; ΕΡΓΑΛΕΙΑ ΑΝΑΛΥΣΗΣ HLA ΔΕΔΟΜΕΝΩΝ .....</b>                              | <b>48</b>  |
| 4.1      | Μητρώα Εθελοντών Δοτών Μυελού των Οστών .....                                                                                           | 48         |
| 4.1.1    | Διεθνή Μητρώα Εθελοντών Δοτών Μυελού των Οστών .....                                                                                    | 49         |
| 4.1.1.1  | BMDW .....                                                                                                                              | 50         |
| 4.1.1.2  | WMDA .....                                                                                                                              | 51         |
| 4.1.1.3  | NMDP .....                                                                                                                              | 51         |
| 4.1.1.4  | CIBMTR .....                                                                                                                            | 51         |
| 4.1.2    | Ελληνικά Μητρώα Εθελοντών Δοτών Μυελού των Οστών .....                                                                                  | 53         |
| 4.1.2.1  | ΚΕΔΜΟΠ .....                                                                                                                            | 55         |
| 4.2      | Εργαλεία Διαχείρισης HLA Δεδομένων .....                                                                                                | 56         |
| 4.2.1    | Ανάλυση HLA με χρήση του λογισμικού ARLEQUIN .....                                                                                      | 56         |
| <b>5</b> | <b>ΕΠΕΞΕΡΓΑΣΙΑ &amp; ΑΝΑΛΥΣΗ HLA ΔΕΔΟΜΕΝΩΝ ΕΛΛΗΝΙΚΟΥ ΜΗΤΡΩΟΥ<br/>5.131 ΕΘΕΛΟΝΤΩΝ ΔΟΤΩΝ ΜΥΕΛΟΥ ΤΩΝ ΟΣΤΩΝ ΜΕΣΩ ΤΟΥ ARLEQUIN 3.5 .....</b> | <b>60</b>  |
| 5.1      | Ανάπτυξη Μητρώου Ελλήνων Εθελοντών Δοτών Μυελού των Οστών ΚΕΔΜΟΠ .....                                                                  | 60         |
| 5.2      | Αντικείμενα & Μέθοδοι HLA Ανάλυσης .....                                                                                                | 62         |
| 5.2.1    | Στοιχεία Μητρώου Εθελοντών Δοτών ΚΕΔΜΟΠ .....                                                                                           | 62         |
| 5.2.2    | Στοιχεία Δείγματος Πληθυσμού .....                                                                                                      | 63         |
| 5.2.3    | HLA Τυποποίηση & Επεξεργασία HLA Δεδομένων .....                                                                                        | 63         |
| 5.2.3.1  | DNA Lookup Tool .....                                                                                                                   | 64         |
| 5.2.4    | Δεδομένα Εισόδου .....                                                                                                                  | 67         |
| 5.2.4.1  | Δημιουργία Αρχείου ARLEQUIN .....                                                                                                       | 67         |
| 5.2.4.2  | Τιμές Υπολογιστικών Παραμέτρων .....                                                                                                    | 70         |
| 5.2.5    | Αποτελέσματα Βιοστατιστικής Ανάλυσης Δεδομένων .....                                                                                    | 77         |
| 5.2.5.1  | Πλήθος & Συχνότητες Αλληλόμορφων .....                                                                                                  | 77         |
| 5.2.5.2  | Πλήθος & Συχνότητες Απλοτύπων .....                                                                                                     | 88         |
| 5.2.5.3  | Πλήθος & Συχνότητες Αντιγόνων .....                                                                                                     | 96         |
| 5.2.5.4  | Έλεγχος Ισορροπίας Hardy-Weinberg .....                                                                                                 | 100        |
| 5.2.5.5  | Ανισορροπία Σύνδεσης .....                                                                                                              | 103        |
| 5.2.5.6  | Ανάλυση Πιθανοτήτων Ισοσυμβατότητας & Βελτιστοποίηση Μητρώου .....                                                                      | 107        |
| <b>6</b> | <b>ΣΥΜΠΕΡΑΣΜΑΤΑ &amp; ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ .....</b>                                                                                      | <b>115</b> |
| 6.1      | Σύνοψη & Συμπεράσματα .....                                                                                                             | 115        |
| 6.2      | Μελλοντική Εργασία .....                                                                                                                | 119        |
|          | <b>ΑΝΑΦΟΡΕΣ &amp; ΔΙΑΔΙΚΤΥΑΚΕΣ ΠΗΓΕΣ .....</b>                                                                                          | <b>120</b> |

---

## ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

---

|                    |                                                                                                                                   |    |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------|----|
| <b>ΕΙΚΟΝΑ 1.1</b>  | Το λογότυπο του ΚΕΔΜΟΠ.....                                                                                                       | 3  |
| <b>ΕΙΚΟΝΑ 2.1</b>  | Τα κυτταρικά στοιχεία του αίματος.....                                                                                            | 6  |
| <b>ΕΙΚΟΝΑ 2.2</b>  | Η τοπολογία των HLA γονιδίων στο χρωμόσωμα 6 .....                                                                                | 15 |
| <b>ΕΙΚΟΝΑ 2.3</b>  | Η οργάνωση των HLA γονιδίων σε αθροίσματα τάξης I, II και III.....                                                                | 17 |
| <b>ΕΙΚΟΝΑ 2.4</b>  | Οι περιοχές των HLA γονιδιακών θέσεων και τα γονίδια που περιλαμβάνουν .....                                                      | 18 |
| <b>ΕΙΚΟΝΑ 2.5</b>  | Η δομή του HLA μορίου τάξης I.....                                                                                                | 21 |
| <b>ΕΙΚΟΝΑ 2.6</b>  | Η δομή του HLA μορίου τάξης II .....                                                                                              | 22 |
| <b>ΕΙΚΟΝΑ 2.7</b>  | Η σχισμοειδής θήκη ως θέση σύνδεσης αντιγόνου και HLA μορίου .....                                                                | 22 |
| <b>ΕΙΚΟΝΑ 2.8</b>  | Σύνδεση αντιγόνου, MHC μορίου & υποδοχέα TCR.....                                                                                 | 23 |
| <b>ΕΙΚΟΝΑ 2.9</b>  | Κληρονομική μεταβίβαση HLA γονιδίων .....                                                                                         | 24 |
| <b>ΕΙΚΟΝΑ 2.10</b> | Ο ρόλος των HLA μορίων στις κύριες οδούς παρουσίασης των αντιγόνων του<br>μοσχεύματος του δότη στα T-λεμφοκύτταρα του λήπτη ..... | 26 |
| <b>ΕΙΚΟΝΑ 2.11</b> | Ονοματολογία αλληλόμορφων HLA γονιδίων.....                                                                                       | 33 |
| <b>ΕΙΚΟΝΑ 4.1</b>  | Η εξέλιξη του πλήθους εθελοντών δοτών μυελού των οστών στο χρόνο σε παγκόσμιο<br>επίπεδο, σύμφωνα με το αρχείο BWMD .....         | 50 |
| <b>ΕΙΚΟΝΑ 4.2</b>  | Ετήσια εκτίμηση πλήθους αυτόλογων και αλλογενών μεταμοσχεύσεων στις ΗΠΑ,<br>σύμφωνα με το κέντρο CIBMTR .....                     | 52 |
| <b>ΕΙΚΟΝΑ 4.3</b>  | Πηγές αλλογενών μοσχευμάτων ανά χρονικές περιόδους στις ΗΠΑ, σύμφωνα με το<br>κέντρο CIBMTR.....                                  | 52 |
| <b>ΕΙΚΟΝΑ 4.4</b>  | Η γενική καρτέλα πληροφοριών του πακέτου ARLEQUIN 3.5.....                                                                        | 57 |
| <b>ΕΙΚΟΝΑ 5.1</b>  | Η ειδική αίτηση εγγραφής εθελοντή δότη μυελού των οστών του ΚΕΔΜΟΠ.....                                                           | 60 |
| <b>ΕΙΚΟΝΑ 5.2</b>  | Τμήμα δεδομένων ταυτότητας εγγεγραμμένων δοτών μητρώου ΚΕΔΜΟΠ .....                                                               | 61 |
| <b>ΕΙΚΟΝΑ 5.3</b>  | Τμήμα δεδομένων HLA τυποποίησης εγγεγραμμένων δοτών μητρώου ΚΕΔΜΟΠ ....                                                           | 62 |
| <b>ΕΙΚΟΝΑ 5.4</b>  | Μετατροπή κωδικού αλληλόμορφου σε 4-digit μορφή μέσω του DNA Lookup Tool ...                                                      | 65 |
| <b>ΕΙΚΟΝΑ 5.5</b>  | Η καρτέλα Project του ARLEQUIN 3.5.....                                                                                           | 70 |
| <b>ΕΙΚΟΝΑ 5.6</b>  | Η επιλογή General Settings της καρτέλας Settings του ARLEQUIN 3.5 .....                                                           | 71 |
| <b>ΕΙΚΟΝΑ 5.7</b>  | Η επιλογή Molecular diversity indices της καρτέλας Settings του ARLEQUIN 3.5 ....                                                 | 72 |
| <b>ΕΙΚΟΝΑ 5.8</b>  | Η επιλογή Haplotype inference – EM algorithm της καρτέλας Settings του<br>ARLEQUIN 3.5.....                                       | 74 |
| <b>ΕΙΚΟΝΑ 5.9</b>  | Η επιλογή Linkage disequilibrium – Pairwise linkage της καρτέλας Settings του<br>ARLEQUIN 3.5.....                                | 75 |
| <b>ΕΙΚΟΝΑ 5.10</b> | Η επιλογή Linkage disequilibrium – Hardy-Weinberg της καρτέλας Settings του<br>ARLEQUIN 3.5.....                                  | 76 |
| <b>ΕΙΚΟΝΑ 5.11</b> | Γραφική παράσταση αθροιστικών συχνοτήτων αλληλόμορφων γονιδίων.....                                                               | 79 |
| <b>ΕΙΚΟΝΑ 5.12</b> | Γραφική παράσταση συχνοτήτων των 20 συχνότερων απλοτύπων του<br>Ελληνικού μητρώου .....                                           | 92 |
| <b>ΕΙΚΟΝΑ 5.13</b> | Γραφική παράσταση συχνοτήτων των 20 συχνότερων απλοτύπων του<br>Γερμανικού μητρώου.....                                           | 93 |

|                    |                                                                                                                                                                                                            |     |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>EIKONA 5.14</b> | Γραφική παράσταση συχνοτήτων των 15 κοινών συχνότερων απλοτύπων του Ελληνικού και του Γερμανικού μητρώου .....                                                                                             | 94  |
| <b>EIKONA 5.15</b> | Γραφική παράσταση συχνοτήτων των 20 κοινών συχνότερων απλοτύπων του Ελληνικού και του Γερμανικού μητρώου ως προς το Γερμανικό μητρώο .....                                                                 | 95  |
| <b>EIKONA 5.16</b> | Γραφική παράσταση της πιθανότητας ιστοσυμβατότητας $p(n)$ ως προς το μέγεθος του μητρώου των εθελοντών δοτών (DRS) .....                                                                                   | 110 |
| <b>EIKONA 5.17</b> | Γραφική παράσταση της πιθανότητας ιστοσυμβατότητας $p(n)$ ως προς το μέγεθος του μητρώου των εθελοντών δοτών (DRS) σε λογαριθμική κλίμακα του άξονα $x$ , σύμφωνα με το διακριτό μη γραμμικό μοντέλο ..... | 113 |
| <b>EIKONA 5.18</b> | Γραφική παράσταση της πιθανότητας ιστοσυμβατότητας $p(n)$ ως προς το μέγεθος του μητρώου των εθελοντών δοτών (DRS) σε λογαριθμική κλίμακα του άξονα $x$ , σύμφωνα με το συνεχές μη γραμμικό μοντέλο.....   | 114 |

---

## ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

---

|                  |                                                                                                                                                       |     |
|------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>ΠΙΝΑΚΑΣ Α</b> | Συσχετίσεις HLA αντιγόνων τάξης I και II με νοσήματα.....                                                                                             | 29  |
| <b>ΠΙΝΑΚΑΣ Β</b> | Γενική αρχή ονοματολογίας HLA συστήματος.....                                                                                                         | 34  |
| <b>ΠΙΝΑΚΑΣ Γ</b> | Παραδείγματα ονοματολογίας και τυποποίησης HLA παραγόντων.....                                                                                        | 34  |
| <b>ΠΙΝΑΚΑΣ Δ</b> | Επίπεδα ανάλυσης HLA τυποποίησης.....                                                                                                                 | 35  |
| <b>ΠΙΝΑΚΑΣ Ε</b> | Στατιστικά στοιχεία μεταμοσχεύσεων στην Ελλάδα.....                                                                                                   | 54  |
| <b>ΠΙΝΑΚΑΣ Ζ</b> | Αποτελέσματα πλήθους παρατηρούμενων αλληλόμορφων.....                                                                                                 | 77  |
| <b>ΠΙΝΑΚΑΣ Η</b> | Αποτελέσματα συχνοτήτων αλληλόμορφων.....                                                                                                             | 78  |
| <b>ΠΙΝΑΚΑΣ Θ</b> | Αθροιστικές συχνότητες αλληλόμορφων.....                                                                                                              | 79  |
| <b>ΠΙΝΑΚΑΣ Ι</b> | Αποτελέσματα τιμών εκτιμώμενων συχνοτήτων των 100 συχνότερων απλοτύπων.....                                                                           | 89  |
| <b>ΠΙΝΑΚΑΣ Κ</b> | Συχνότητες 20 συχνότερων απλοτύπων του Ελληνικού μητρώου.....                                                                                         | 92  |
| <b>ΠΙΝΑΚΑΣ Λ</b> | Συχνότητες 20 συχνότερων απλοτύπων του Γερμανικού μητρώου.....                                                                                        | 93  |
| <b>ΠΙΝΑΚΑΣ Μ</b> | Συχνότητες 15 κοινών συχνότερων απλοτύπων Ελληνικού & Γερμανικού<br>μητρώου.....                                                                      | 94  |
| <b>ΠΙΝΑΚΑΣ Ν</b> | Συχνότητες 20 κοινών συχνότερων απλοτύπων Ελληνικού & Γερμανικού<br>μητρώου ως προς φθίνουσα ταξινόμηση του Γερμανικού μητρώου.....                   | 95  |
| <b>ΠΙΝΑΚΑΣ Ξ</b> | Αποτελέσματα πλήθους παρατηρούμενων αντιγόνων.....                                                                                                    | 96  |
| <b>ΠΙΝΑΚΑΣ Ο</b> | Αποτελέσματα συχνοτήτων αντιγόνων.....                                                                                                                | 97  |
| <b>ΠΙΝΑΚΑΣ Π</b> | Αποτελέσματα ελέγχου ισορροπίας Hardy-Weinberg (HWE).....                                                                                             | 101 |
| <b>ΠΙΝΑΚΑΣ Ρ</b> | Αποτελέσματα 20 μεγαλύτερων τιμών ανισορροπίας σύνδεση για τα 20 συχνότερα<br>παρατηρούμενα ζεύγη αλληλόμορφων των εξεταζόμενων γονιδιακών τόπων..... | 103 |
| <b>ΠΙΝΑΚΑΣ Σ</b> | Αποτελέσματα 10 υψηλότερων θετικών και αρνητικών τιμών ανισορροπίας<br>σύνδεσης για τα παρατηρούμενα ζεύγη των εξεταζόμενων γονιδιακών τόπων.....     | 104 |
| <b>ΠΙΝΑΚΑΣ Τ</b> | Εκτιμήσεις τιμών των μεγεθών $p(n)$ και $d(n)$ για τα τρία σενάρια της HLA<br>ανάλυσης για το αρχικό μητρώο των $n=5.131$ εθελοντών δοτών.....        | 110 |

# Κ Ε Φ Α Λ Α Ι Ο 1<sup>ο</sup>

## ΕΙΣΑΓΩΓΗ

### 1.1 Αντικείμενο της διπλωματικής εργασίας

Τα αντιγόνα ιστοσυμβατότητας είναι βιομόρια που βρίσκονται στην επιφάνεια των κυττάρων, τα οποία καθορίζουν την ταυτότητα του οργανισμού και δεν αναγνωρίζονται από τα Τ-λεμφοκύτταρα ως ξένα και κατά συνέπεια δεν προκαλούν την ανοσοβιολογική απόκριση του οργανισμού. Η ομάδα των εξαιρετικά πολυμορφικών γονιδίων που εκφράζει τα αντιγόνα ιστοσυμβατότητας και τα προϊόντα της οποίας σχετίζονται με την κυτταρική αναγνώριση και με την διάκριση μεταξύ αυτο-αντιγόνων και αλλο-αντιγόνων αποτελεί το **Μείζον Σύμπλεγμα Ιστοσυμβατότητας (MHC - Major Histocompatibility Complex)**. Τα γονίδια MHC στον άνθρωπο εκφράζονται με τη μορφή του συστήματος των **Ανθρώπινων Λευκοκυτταρικών Αντιγόνων (HLA - Human Leukocyte)**. Τα αντιγόνα αυτά ονομάστηκαν έτσι, επειδή, αναφέρονται στις πρωτεΐνες που κωδικοποιούνται από τα γονίδια του συμπλέγματος MHC και εκφράζονται στην επιφάνεια της κυτταρικής μεμβράνης των λευκοκυττάρων. Τα HLA γονίδια αποτελούν ένα σύνολο άνισα κατανομημένων γονιδίων, τα οποία βρίσκονται πάνω στο βραχύ σκέλος του χρωμοσώματος 6, αντιπροσωπεύουν το 0,1% της όλης γενετικής πληροφορίας και διακρίνονται με τεχνητό τρόπο σε τρία μεγάλα αθροίσματα (τάξης I, II και III). Κάθε άθροισμα περιέχει πολλές γονιδιακές θέσεις (loci), κάθε μια από τις οποίες εκφράζεται με πολλές γονιδιακές μορφές (αλληλόμορφα γονίδια ή αλλήλια) στο χρωμόσωμα. Ο συνδυασμός των HLA γονιδίων που συνδέονται στενά πάνω στο ίδιο χρωμόσωμα και κληρονομούνται από τον ένα γονέα ονομάζεται απλότυπος. Η **γενετική σύνδεση** είναι η τάση που έχουν συγκεκριμένες γονιδιακές θέσεις στα χρωμοσώματα ή συγκεκριμένα αλληλόμορφα γονίδια να κληρονομούνται μαζί. Οι γενετικές θέσεις που βρίσκονται κοντά ή μια στην άλλη πάνω στο ίδιο χρωμόσωμα έχουν την τάση να μην διαχωρίζονται κατά την μείωση και να παραμένουν γενετικά συνδεδεμένες. Σχετικό φαινόμενο αποτελεί η **ανισορροπία γενετικής σύνδεσης (LD - Linkage Disequilibrium)**, κατά το οποίο η συχνότητα εμφάνισης

κάποιων μεταλλάξεων ή πολυμορφισμών δεν ακολουθεί την αναμενόμενη συχνότητα έτσι ώστε, κατά συνέπεια να προκύπτει διαφορά στην αναμενόμενη συχνότητα (expected frequency) που υπολογίζεται θεωρητικά και στην αντίστοιχη παρατηρούμενη συχνότητα (observed frequency) που υπολογίζεται στην πραγματικότητα .

Η αλλογενής **μεταμόσχευση αιμοποιητικών βλαστικών κυττάρων (HSCT - Hematopoietic Stem Cell Transplantation)** είναι μια ολοένα και περισσότερο διαδεδομένη θεραπεία για τη λευχαιμία και άλλες σοβαρές ασθένειες του αίματος. Οι αναζητήσεις μη συγγενών συμβατών δοτών διενεργούνται βάσει HLA δεδομένων, τα οποία καταχωρούνται σε ειδικά μητρώα εθελοντών δοτών. Λόγω της εκτεταμένης ποικιλομορφίας του HLA συστήματος και της ανάγκης ύπαρξης, ικανής για διενέργεια μεταμόσχευσης, ιστοσυμβατότητας μεταξύ δότη και λήπτη, έχουν δημιουργηθεί διεθνή και εθνικά μητρώα και δεξαμενές εθελοντών δοτών μυελού των οστών. Από τον Ιούνιο του 2013, υπήρχαν περισσότεροι από 21.000.000 δυνητικοί εθελοντές δότες αιμοποιητικών βλαστοκυττάρων σε 69 μητρώα από 50 χώρες.

Η πιθανότητα εύρεσης ενός τουλάχιστον HLA-συμβατού δότη για έναν ασθενή από έναν συγκεκριμένο πληθυσμό σε ένα καθορισμένο αρχείο εθελοντών δοτών είναι μια βασική παράμετρος για τον στρατηγικό σχεδιασμό του μητρώου των δοτών. Ο ορισμός της HLA συμβατότητας έχει αλλάξει με την πάροδο του χρόνου. Επί του παρόντος, ένας δότης ο οποίος ταιριάζει με έναν ασθενή τουλάχιστον για τους HLA γονιδιακούς τύπους HLA-A, HLA-B και HLA-DRB1, τυποποιημένοι σε υψηλή ανάλυση («8/8 HLA απλότυποι υψηλής ανάλυσης»), θεωρείται ως HLA-συμβατός. Η τυποποίηση υψηλής ανάλυσης απαιτεί αλληλούχηση των εξονίων 2 και 3 για τα γονίδια τάξης I (HLA-A και HLA-B) και του εξονίου 2 για τα HLA-DRB1.

Μέθοδοι για τον υπολογισμό των **πιθανοτήτων ιστοσυμβατότητας** συναρτήσεως του μεγέθους του μητρώου δοτών έχουν αναπτυχθεί από το 1980. Οι μέθοδοι αυτές βασίζονται στην εκτίμηση των συχνοτήτων των HLA απλοτύπων, οι οποίες μπορούν να εκτιμηθούν από τα φαινοτυπικά δεδομένα των δοτών, χρησιμοποιώντας τα κατάλληλα εργαλεία, όπως τον αλγόριθμο μεγιστοποίησης προσδοκίας (EM - Expectation-Maximization). Λαμβάνοντας υπόψη το σημαντικό ποσό των διεθνών ανταλλαγών προϊόντων αιμοποιητικών βλαστικών κυττάρων, τα μοντέλα βιοστατιστικών αναλύσεων των απλοτυπικών συχνοτήτων και των πιθανοτήτων ιστοσυμβατότητας φαίνεται να είναι υπεραπλουστευμένα.

Στην παρούσα εργασία, έχει περιγραφεί ένα γενικό πλαίσιο για την διενέργεια εκτιμήσεων των τιμών των πιθανοτήτων ιστοσυμβατότητας και εύρεσης συμβατών δοτών για μια ενδεχόμενη μεταμόσχευση μυελού των οστών, το οποίο βασίζεται στην βιοστατιστική ανάλυση και εφαρμογή ειδικών μεθοδολογιών σε HLA δεδομένα που περιλαμβάνονται στο μητρώο ενός πληθυσμού 5.131 Ελλήνων εθελοντών δοτών μυελού των οστών του Κέντρου Ενημέρωσης & Προσέλκυσης Εθελοντών Δοτών Μυελού των Οστών του Πανεπιστημίου Πατρών (ΚΕΔΜΟΠ). Τα αποτελέσματα των αναλύσεων, τα οποία προκύπτουν από την επεξεργασία των HLA δεδομένων με εξειδικευμένα βιοστατιστικά εργαλεία λογισμικού (Arlequin) και σχετίζονται με τιμές συχνοτήτων HLA αλληλόμορφων και HLA απλοτύπων, καθώς επίσης και ειδικών παραμέτρων όπως η ανισορροπία σύνδεσης, η γενετική απόσταση, ο έλεγχος υποθέσεων γενετικής ισορροπίας Hardy-Weinberg κλπ., χρησιμοποιούνται στη συνέχεια για βελτιστοποίηση και στρατηγικό σχεδιασμό του μητρώου των δοτών.



**ΕΙΚΟΝΑ 1.1** Το λογότυπο του ΚΕΔΜΟΠ

## **1.2 Διάρθρωση της διπλωματικής εργασίας**

Στα πλαίσια της εκπόνησης της διπλωματικής εργασίας θα καθοριστεί η κατανομή των συχνοτήτων των HLA αλληλόμορφων γονιδίων και των αντίστοιχων HLA απλοτύπων, καθώς επίσης θα υπολογιστεί ο δείκτης ανισορροπίας γενετικής σύνδεσης μεταξύ των αλληλόμορφων γονιδίων HLA, σε έναν πληθυσμό Ελλήνων εθελοντών δοτών μυελού των οστών για τα γονίδια HLA-A, HLA-B και HLA-DRB1, η οποία εν συνεχεία θα χρησιμοποιηθεί για τη βελτιστοποίηση της αναζήτησης μεμονωμένων βλαστικών κυττάρων των δοτών και την αποδοτική διαχείριση του αντίστοιχου μητρώου των δοτών.

Επιπλέον, οι συχνότητες των απλοτύπων HLA που θα υπολογιστούν, θα χρησιμοποιηθούν για την εκτίμηση της πιθανότητας εύρεσης συμβατών δοτών εντός του συνόλου των δοτών. Οι αναλύσεις και η επεξεργασία του συνόλου των δεδομένων των δοτών θα υλοποιηθούν με συγκεκριμένα ειδικά πακέτα λογισμικού και υπολογιστικούς αλγόριθμους και μεθοδολογίες.

Τέλος, θα πραγματοποιηθεί μια προσπάθεια σύγκρισης των στατιστικών αποτελεσμάτων που θα προκύψουν από την ανάλυση των δεδομένων του πληθυσμού των Ελλήνων δοτών, με τα αντίστοιχα αποτελέσματα άλλων πληθυσμών (Γερμανών, Κυπρίων δοτών) για την εξαγωγή συμπερασμάτων ως προς το ποσοστό της ιστοσυμβατότητας HLA μεταξύ των πληθυσμών αυτών.

Η διάρθρωση της εργασίας ακολουθεί μια προοδευτική προσέγγιση, ξεκινώντας από την ανάλυση των βασικότερων χαρακτηριστικών της μεταμόσχευσης αιμοποιητικών κυττάρων, της ιστοσυμβατότητας και των στοιχείων και παραμέτρων βιοστατιστικής ανάλυσης των HLA δεδομένων, στην συνέχεια παρουσιάζονται ζητήματα διαχείρισης διεθνών και εθνικών μητρώων και δεξαμενών εθελοντών δοτών μυελού των οστών και εξειδικευμένα εργαλεία ανάλυσης των αντίστοιχων HLA δεδομένων που περιλαμβάνονται στα μητρώα, ενώ έπειτα παρουσιάζονται τα στάδια επεξεργασίας και διαχείρισης των HLA δεδομένων του μητρώου των Ελλήνων εθελοντών δοτών μυελού των οστών του ΚΕΔΜΟΠ, καθώς και τα αποτελέσματα της ανάλυσης μέσω του βιοστατιστικού πακέτου Arlequin.

Πιο συγκεκριμένα :

Στο *δεύτερο κεφάλαιο* παρουσιάζονται τα βασικά είδη και οι λειτουργίες του μυελού των οστών , τα χαρακτηριστικά της μεταμόσχευσης αιμοποιητικών κυττάρων, καθώς επίσης έννοιες και εφαρμογές που σχετίζονται με τα συμπλέγματα ιστοσυμβατότητας και το HLA σύστημα.

Στο *τρίτο κεφάλαιο* γίνεται αναφορά σε ποσοτικά στοιχεία και ειδικές παραμέτρους και μετρικές βιοστατιστικής ανάλυσης HLA δεδομένων, με στόχο την διενέργεια εκτιμήσεων των τιμών των συχνοτήτων HLA αλληλόμορφων και HLA απλοτύπων του πληθυσμού του μητρώου του ΚΕΔΜΟΠ.

Το *τέταρτο κεφάλαιο* αποτελεί μια διερεύνηση των μητρώων και των δεξαμενών εθελοντών δοτών μυελού των οστών σε διεθνές και εθνικό επίπεδο, καθώς και συγκεκριμένα στα χαρακτηριστικά της ανάπτυξης και της διαχείρισης του μητρώου των εθελοντών δοτών του ΚΕΔΜΟΠ. Επιπλέον, γίνεται αναφορά και σε συγκεκριμένα εργαλεία και πακέτα λογισμικού βιοστατιστικής ανάλυσης HLA δεδομένων όπως το Arlequin.

Στο *πέμπτο κεφάλαιο*, αφού παρουσιάστηκαν στοιχεία του δείγματος του πληθυσμού των δοτών του μητρώου του ΚΕΔΜΟΠ και της τυποποίησης των αντίστοιχων HLA δεδομένων, αναπτύχθηκε η επεξεργάσιμη από το πακέτο Arlequin μορφή των δεδομένων εισόδου και καθορίστηκαν οι τιμές των σχετικών υπολογιστικών παραμέτρων και μετρικών. Τέλος, παρουσιάζονται τα αποτελέσματα της βιοστατιστικής ανάλυσης των HLA δεδομένων των Ελλήνων εθελοντών δοτών του μητρώου του ΚΕΔΜΟΠ ως προς τον υπολογισμό των συχνοτήτων και του πλήθους των αλληλόμορφων, των αντιγόνων και την εκτίμηση των συχνοτήτων των απλοτύπων, του ελέγχου ισορροπίας Hardy-Weinberg και του υπολογισμού του ποσού της ανισορροπίας σύνδεσης μεταξύ των ζευγών των διαφορετικών γονιδιακών τόπων. Βάσει των παραπάνω αποτελεσμάτων, πραγματοποιείται ανάλυση των πιθανοτήτων ιστοσυμβατότητας με σκοπό τη βελτιστοποίηση και τον στρατηγικό σχεδιασμό του μητρώου των δοτών.

Τέλος, στο *έκτο κεφάλαιο*, παρουσιάζονται τα αποτελέσματα και τα συμπεράσματα της εργασίας, ενώ παράλληλα αναφέρονται πιθανές μελλοντικές επεκτάσεις πάνω στο αντικείμενο που πραγματεύτηκε η εργασία.

# Κ Ε Φ Α Λ Α Ι Ο 2<sup>ο</sup>

## ΜΕΤΑΜΟΣΧΕΥΣΗ ΑΙΜΟΠΟΙΗΤΙΚΩΝ ΚΥΤΤΑΡΩΝ & ΙΣΤΟΣΥΜΒΑΤΟΤΗΤΑ

### 2.1 Μυελός των Οστών

Ο **μυελός των οστών** είναι ρευστός ιστός που βρίσκεται στις μυελικές κυψέλες των οστών του σώματος και κυρίως στα οστά της λεκάνης και του στέρνου. Περιέχει αρχέγονα κύτταρα τα οποία όταν ωριμάσουν μετατρέπονται σε λευκά αιμοσφαίρια, ερυθρά αιμοσφαίρια και αιμοπετάλια. Αρχέγονα κύτταρα υπάρχουν και στο αίμα, σε μικρή όμως ποσότητα. Αυτά τα κύτταρα όταν μεταμοσχεύονται σε ασθενείς δημιουργούν νέο υγιή μυελό.



**ΕΙΚΟΝΑ 2.1** Τα κυτταρικά στοιχεία του αίματος

Αξιοσημείωτο είναι το γεγονός, ότι οι μυελικές κυψέλες των οστών κατά τη γέννηση αντιπροσωπεύουν το 1,5% του σωματικού βάρους, ενώ στον ενήλικα το 4.5%, παρόλο που ο ολικός όγκος αίματος ελλατώνεται από το 8% του σωματικού βάρους κατά τη γέννηση, στο 7% στον ενήλικα. (Εθνικός Οργανισμός Μεταμοσχεύσεων 2006, Βοργιάς Ν.Ι., Λαουτάρης Ν.Π. 1991).

### **2.1.1 Είδη Μυελού των Οστών**

Ο μυελός των οστών διακρίνεται σε δυο είδη, στον **ερυθρό** και στον **λιπώδη** μυελό των οστών.

Ο **ερυθρός μυελός** των οστών, που βρίσκεται στα σπογγώδη οστά (σπόνδυλοι, πλευρές, άνω άκρο μηριαίου οστού και τα πλατιά οστά της πυέλου) και στις επιφύσεις των επιμηκών οστών, αποτελεί ένα μεγάλο και διάχυτο αιμοποιητικό όργανο και παράγει τα κύτταρα του αίματος και την αιμοσφαιρίνη. Σε ορισμένες περιπτώσεις, όπως σε μεγάλη αιμόλυση, μπορεί να εμφανισθεί και σε άλλα σημεία, όπως στις διαφύσεις των επιμηκών οστών και οδηγεί, στα παιδιά, σε διαταραχές της ανάπτυξης του οστού και σε οστικές παραμορφώσεις. Στις περιπτώσεις αυτές, λόγω της πλήρους κατάληψης των οστών από ερυθρό μυελό, μπορεί να εμφανισθεί αιμοποίηση στο ήπαρ και στον σπλήνα.

Ο **λιπώδης ή ωχρός μυελός των οστών** βρίσκεται στον αυλό των επιμηκών οστών, συνίσταται από λιπώδη κύτταρα και συνδετικό ιστό και είναι αδρανής, δηλαδή δεν συμμετέχει στην αιμοποίηση (Πλέσσας Σ., Κανέλλος Ε. 1997).

### **2.1.2 Λειτουργία Μυελού των Οστών**

Η λειτουργία του μυελού των οστών είναι πολλαπλή. Πρόκειται για το κύριο όργανο του συστήματος μονοκυττάρων – μακροφάγων, που συμμετέχει στις διαδικασίες αναγνώρισης των αντιγόνων, στην κυτταρική και χημική ανοσία και στην αναγνώριση και απομάκρυνση με φαγοκυττάρωση των γερασμένων κυττάρων. Η κύρια όμως λειτουργία του μυελού είναι η αιμοποίηση. Η διαφοροποίηση και ωρίμανση των διαφόρων κυττάρων του αίματος ξεκινάει από πηγές αρχέγονων κυττάρων, πολυδύναμων αδέσμευτων ή μερικώς δεσμευμένων, από τις οποίες προέρχονται πηγές δεσμευμένων προγονικών κυττάρων, τα οποία μορφολογικά δεν διαφέρουν μεταξύ τους, μοιάζουν όμως απόλυτα με τα μικρά ή μέσου μεγέθους λεμφοκύτταρα. Τα κύτταρα αυτά, έχουν την ιδιότητα της αυτόματης ανανέωσής τους με μιτώσεις αλλά και της διαφοροποίησής τους, επίσης ύστερα από μιτώσεις (Βοργιάς Ν.Ι., Λαουτάρης Ν.Π. 1991).

## **2.2 Μεταμόσχευση Αιμοποιητικών Κυττάρων**

Η μεταμόσχευση αιμοποιητικών κυττάρων αποτελεί μια θεραπευτική μέθοδο που αποσκοπεί στην ίαση αιματολογικών, νεοπλασματικών, ανοσολογικών, γενετικών, μεταβολικών και άλλων παθήσεων. Χρησιμοποιείται στη θεραπευτική διαχείριση νοσημάτων που δεν είναι δυνατόν να αντιμετωπισθούν με άλλες γνωστές θεραπευτικές μεθόδους, όπως π.χ. χημειοθεραπεία, ακτινοβολία κλπ., ως αποκλειστική θεραπευτική διέξοδος ή ως εναλλακτική θεραπεία στις περιπτώσεις που η χρήση της έχει να επιδείξει καλύτερα αποτελέσματα έναντι των υπόλοιπων εναλλακτικών μεθόδων.

### **2.2.1 Αιμοποιητικά Βλαστικά Κύτταρα (HSC)**

Μόσχευμα αιμοποιητικών κυττάρων καλείται η συλλογή αιμοποιητικών κυττάρων από έναν δότη με σκοπό την χορήγησή τους σε λήπτη ασθενή. Ένα μόσχευμα περιέχει μεταξύ άλλων, πολυδύναμα αρχέγονα αιμοποιητικά κύτταρα. Τα αιμοποιητικά βλαστικά κύτταρα (HSC - Hematopoietic Stem Cells), ως στελεχιαία πολυδύναμα κύτταρα, έχουν τρεις βασικές ιδιότητες :

- ❖ Είναι αδιαφοροποίητα, δηλαδή δεν έχουν κάποια ειδική ιστολογική δομή που να τους επιτρέπει την πραγματοποίηση εξειδικευμένων λειτουργιών.
- ❖ Έχουν τη δυνατότητα της διαίρεσης-αναπαραγωγής και αυτοανανέωσης για μεγάλες χρονικές περιόδους, δηλαδή γεννούν όμοιους απογόνους, κάτι που έχει ως αποτέλεσμα τη διατήρηση αυτού του πληθυσμού εφόρου ζωής, σε αντίθεση με τα μυϊκά κύτταρα, τα κύτταρα του αίματος και τα νευρικά κύτταρα, τα οποία υπό κανονικές συνθήκες δεν αναπαράγονται. Έτσι, ένας αρχικός πληθυσμός αιμοποιητικών βλαστικών κυττάρων μπορεί να πολλαπλασιαστεί στο εργαστήριο δίνοντας εκατομμύρια κύτταρα. Τα κύτταρα-απόγονοι, που συνεχίζουν να παραμένουν αδιαφοροποίητα, όπως τα μητρικά κύτταρα, είναι ικανά για μακροχρόνια αυτοανανέωση.
- ❖ Έχουν τη δυνατότητα διαφοροποίησης προς ωριμότερες μορφές κυττάρων. Η διαδικασία από την οποία προκύπτουν διαφοροποιημένα κύτταρα από αδιαφοροποίητα ονομάζεται διαφοροποίηση.

Λόγω των παραπάνω ιδιοτήτων τους, στην περίπτωση επιτυχούς μεταμόσχευσής τους, τα αιμοποιητικά βλαστικά κύτταρα εξασφαλίζουν μακροχρόνια αιμοποιΐα από πλευράς του λήπτη του μοσχεύματος. Καθοριστικός παράγοντας για μια επιτυχή μεταμόσχευση είναι η ποσότητα των αιμοποιητικών βλαστικών κυττάρων που περιέχει το μόσχευμα ανά χιλιόγραμμα βάρους του ασθενή (κυτταρική δόση).

### **2.2.2 Πηγές Αιμοποιητικών Βλαστικών Κυττάρων**

Υπάρχουν τρεις πηγές αιμοποιητικών βλαστικών κυττάρων :

- 1) **Ο μυελός των οστών**, ο οποίος αποτελούσε αρχικά την κύρια πηγή λήψης των κυττάρων αυτών και εντός του οποίου τα κύτταρα αυτά πολλαπλασιάζονται και διαφοροποιούνται.
- 2) **Το περιφερικό αίμα**, όπου βρίσκονται σε μικρότερες συγκεντρώσεις και μόνο υπό συγκεκριμένες συνθήκες διέγερσης η ποσότητά τους αυξάνεται.
- 3) **Το αίμα του ομφάλιου λώρου και του πλακούντα**, όπου περιέχεται σημαντικός αριθμός κυττάρων.

Ο εντοπισμός και η καταμέτρηση του αριθμού των αιμοποιητικών βλαστικών κυττάρων σε ένα μόσχευμα γίνεται με τη βοήθεια του αντιγόνου CD34 που φέρουν μόνον αυτά τα κύτταρα.

Οι καθιερωμένες πηγές για μεταμόσχευση αιμοποιητικών κυττάρων είναι ο μυελός των οστών και το περιφερικό αίμα. Παρά το γεγονός ότι υπάρχουν εκατομμύρια δότες καταχωρημένοι στα παγκόσμια μητρώα μη συγγενών εθελοντών δοτών μυελού των οστών, τα αποτελέσματα των αναζητήσεων κατάλληλου δότη έχουν αποδείξει τη μερική μόνον επάρκεια των μητρώων αυτών όσον αφορά στην κάλυψη του συνόλου των αναγκών σε αιμοποιητικά βλαστικά κύτταρα. Ειδικότερα για την πλειοψηφία των εθνικών μειονοτήτων, η εύρεση και πρόσβαση σε κατάλληλο δότη είναι μια δύσκολη υπόθεση. Το γεγονός αυτό, κατά συνέπεια, καθιστά επιτακτική την ανάγκη επινόησης εναλλακτικών λύσεων για τη μεταμόσχευση αιμοποιητικών βλαστικών κυττάρων. Η καλύτερη λύση από αυτές είναι η χορήγηση αίματος που λαμβάνεται από τον ομφάλιο λώρο και τον πλακούντα. Τόσο ο ομφάλιος λώρος, όσο και ο πλακούντας είναι άχρηστος για το νεογνό και τη μητέρα μετά τη γέννηση και σε πολλές περιπτώσεις απορρίπτεται με τα λοιπά βιολογικά απόβλητα των κλινικών.

### **2.2.3 Μεταμόσχευση Αιμοποιητικών βλαστικών Κυττάρων (HSCT)**

Η μεταμόσχευση αιμοποιητικών βλαστικών κυττάρων (HSCT - Hematopoietic Stem Cell Transplantation), που συνήθως υλοποιείται ως μεταμόσχευση μυελού των οστών, είναι μια ιατρική διαδικασία στον τομέα της αιματολογίας και της ογκολογίας. Η μεταμόσχευση μυελού των οστών είναι η θεραπεία επιλογής για τα άτομα με αιμοποιητικές κακοήθειες (π.χ. λευχαιμία), ανεπάρκεια του μυελού των οστών και ορισμένες μορφές καρκίνου (π.χ. λέμφωμα), οι οποίες οδηγούν σε ένα εξασθενημένο ανοσοποιητικό σύστημα. Η βασική αρχή της διαδικασίας της μεταμόσχευσης μυελού των οστών βασίζεται στην ενδοφλέβια έγχυση βλαστικών κυττάρων που συλλέγονται από τον μυελό των οστών του δότη, το περιφερικό αίμα ή το αίμα ομφάλιου λώρου, ώστε να χρησιμοποιηθούν για να αντικαταστήσουν τις αιμοποιητικές λειτουργίες ενός ασθενούς. Η έκβαση της μεταμόσχευσης μυελού των οστών εξαρτάται από διάφορους παράγοντες όπως η ηλικία του ασθενούς, το υποκείμενο νόσημα, η φάση της νόσου κατά τη διάρκεια της μεταμόσχευσης κ.ά. Ωστόσο, ο πιο σημαντικός παράγοντας για την επιτυχή έκβαση της μεταμόσχευσης μυελού των οστών είναι ο βαθμός ιστοσυμβατότητας μεταξύ ασθενούς και δότη, ο οποίος καθορίζει το κατά πόσο ταιριάζουν ως προς τα **Ανθρώπινα Λευκοκυτταρικά Αντιγόνα (HLA - Human Leukocyte Antigen)**. Ενδεχόμενη αναντιστοιχία μεταξύ αλληλόμορφων HLA γονιδίων ανάμεσα σε έναν λήπτη και έναν δότη αυξάνει τη συχνότητα εμφάνισης αλλά και τη σοβαρότητα μιας αλλοδραστικής ανοσοαπόκρισης, κατά τη μεταμόσχευση αιμοποιητικών βλαστικών κυττάρων. Το επίπεδο της προσαρμογής που απαιτείται ποικίλλει ανάλογα με την πηγή των βλαστικών κυττάρων που χρησιμοποιούνται για μεταμόσχευση μυελού των οστών. Στις περισσότερες περιπτώσεις, οι ασθενείς δεν διαθέτουν κατάλληλο HLA συμβατό δότη εντός της οικογένειάς τους. Για ασθενείς που στερούνται HLA ταυτόσημου αδελφού ή φαινοτυπικά συγγενή δότη, οι ιατροί θα πρέπει να ενεργοποιήσουν μια διαδικασία αναζήτησης κατάλληλου ιστοσυμβατού δότη, μέσω της αλληλεπίδρασης ανάμεσα στα εθνικά και διεθνή μητρώα εθελοντών δοτών μυελού των οστών και της διερεύνησης των βάσεων δεδομένων τους, οι οποίες περιλαμβάνουν την αποτύπωση των HLA γονιδίων του λήπτη και των δυνητικών του μη συγγενών εθελοντών δοτών ή μονάδων ομφαλοπλακουντιακού αίματος των δοτών. Στην συνέχεια επιλέγεται ο πλέον κατάλληλος από τους διαθέσιμους ιστοσυμβατούς δότες και η διαδικασία ολοκληρώνεται με επιβεβαιωτικές δοκιμασίες συμβατότητας δότη-λήπτη και με ιολογικό έλεγχο του δότη. Η επιλογή του κατάλληλου δότη αποτελεί ιδιαίτερα δύσκολη απόφαση. Αρχικά και για μεγάλο χρονικό διάστημα, η συμβατότητα

περιοριζόταν στα ορολογικώς καθοριζόμενα HLA αντιγόνα. Σήμερα, η εξέλιξη των τεχνικών τα μοριακής βιολογίας βοήθησε πολύ στην καλύτερη επιλογή του πλέον κατάλληλου δότη και ως εκ τούτου στην επιτυχημένη μεταμόσχευση.

Οι τεχνολογίες της πληροφορικής και των επικοινωνιών, τα υπολογιστικά εργαλεία και οι μεθοδολογίες της πληροφορικής διαδραματίζουν καίριο ρόλο στη διαδικασία αναζήτησης συμβατού δότη στα διαθέσιμα μητρώα εθελοντών δοτών, τόσο σε εθνικό όσο και σε διεθνές επίπεδο. Μία από τις μεγαλύτερες προκλήσεις για τα ηλεκτρονικά υπολογιστικά συστήματα που διαχειρίζονται δεδομένα και επεξεργάζονται πληροφορίες μητρώων εθελοντών δοτών μυελού των οστών είναι η ανάπτυξη κατάλληλων στρατηγικών σχεδιασμού μητρώων, μέσω κατάλληλης βιοστατιστικής ανάλυσης και αξιόπιστου υπολογισμού συχνοτήτων αλληλόμορφων HLA γονιδίων και αντιγόνων, HLA απλοτύπων και εν συνεχεία πιθανοτήτων ιστοσυμβατότητας. Οι εφαρμογές αυτών των μεθοδολογιών χρησιμοποιούνται συνήθως κατά την αναζήτηση συμβατού δότη, όπως ορίζεται και από την **Παγκόσμια Δεξαμενή Δοτών Μυελού των Οστών (BMDW - Bone Marrow Donors Worldwide)**.

Σήμερα, υπάρχουν πάνω από 20 εκατομμύρια δότες αιμοποιητικών βλαστικών κυττάρων και μονάδων αίματος ομφάλιου λώρου, διαθέσιμα σε όλο τον κόσμο. Λόγω του χαρακτήρα του HLA συστήματος, των μεθόδων HLA τυποποίησης και του περιορισμού των πόρων, δεν έχουμε πλήρη ενημέρωση σχετικά με τους HLA τύπους αυτών των δοτών.

Η μεταμόσχευση οργάνων και ιστών που αναπτύχθηκε κυρίως τις τελευταίες δεκαετίες του 20ού αιώνα έδωσε μεν διέξοδο και λύση σε ένα μεγάλο φάσμα νοσημάτων, αλλά δεν έδωσε ακόμη λύση στις έντονες παρενέργειες που προκαλούνται από την παρατεταμένη ανοσοκαταστολή και τις ανοσολογικού τύπου διαταραχές, που καταλήγουν στην απόρριψη του μοσχεύματος. Η νέα τεχνολογία βελτίωσε τις παρενέργειες αυτές με την περεταίρω ανάπτυξη εξειδικευμένων τρόπων επιλογής μοσχευμάτων και ανοσοκατασταλτικής θεραπευτικής προσέγγισης. Η συνεχιζόμενη πρόοδος και οι νεότερες μεθοδολογικές προσεγγίσεις θα δώσουν, ενδεχομένως, οριστική λύση στο πρόβλημα.

#### **2.2.4 Κατηγορίες Μεταμόσχευσης Αιμοποιητικών Βλαστικών Κυττάρων**

Οι μεταμοσχεύσεις μυελού των οστών μπορούν να κατηγοριοποιηθούν βάσει δύο βασικών κριτηρίων διάκρισης :

- 1) Σχέση δότη-λήπτη
- 2) Πηγή προέλευσης του μοσχεύματος

Ανάλογα με τη **σχέση δότη-λήπτη** διακρίνουμε τα εξής είδη μεταμόσχευσης :

**Αυτόλογη μεταμόσχευση.** Ο δότης και ο λήπτης είναι το ίδιο πρόσωπο. Αυτό πρακτικά σημαίνει ότι το μόσχευμα προέρχεται από τον ίδιο το λήπτη και λαμβάνεται είτε από το μυελό των οστών ή το περιφερικό του αίμα, είτε από το ομφαλικό αίμα του αν αυτό έχει διατηρηθεί κατά την γέννησή του. Στην περίπτωση της αυτόλογης μεταμόσχευσης δεν τίθεται θέμα ιστοσυμβατότητας δότη και λήπτη, αφού αυτή είναι απόλυτη.

**Αλλογενής μεταμόσχευση.** Ο δότης και ο λήπτης είναι διαφορετικά πρόσωπα.

Υπάρχουν, μεταξύ άλλων, δύο βασικές προϋποθέσεις που πρέπει να εξασφαλίζονται:

- α) Θα πρέπει ο δότης να είναι υγιής.
- β) Να υπάρχει αποδεκτός βαθμός ιστοσυμβατότητας μεταξύ δότη και λήπτη. Το αλλογενές μόσχευμα συνήθως προέρχεται είτε από ταυτόσημα αδέρφια ή εναλλακτικά, από άλλα μέλη της οικογένειας, είτε από μη συγγενείς εθελοντές δότες. Γενικότερα όσον αφορά στα αλλογενή μοσχεύματα, το 25-30% προέρχεται από ταυτόσημα αδέρφια, ενώ μόλις το 5% από τα υπόλοιπα μέλη της οικογένειας. Στο 40% των περιπτώσεων το μόσχευμα που χρησιμοποιείται προέρχεται από μη συγγενή εθελοντή δότη, ενώ σε μεγάλο αριθμό ασθενών δεν βρίσκεται συμβατός δότης.

Αριθμητικά επιδιώκεται η περιεκτικότητα του μοσχεύματος με επαρκή αριθμό μητρικών αιμοποιητικών κυττάρων, ώστε να ξεπεραστεί ο κίνδυνος απόρριψης και να εξασφαλιστεί η ταχύτερη αποκατάσταση της αιμοποιίας. Η αλλογενής μεταμόσχευση αρχέγονων αιμοποιητικών κυττάρων αποτελεί μια μέθοδο ίασης που εφαρμόζεται, μεταξύ άλλων, σε αιμοποιητικά κακοήγη νοσήματα, σύνδρομο ανεπάρκειας μυελού των οστών και κληρονομικές δυσλειτουργίες ανοσοανεπάρκειας.

**Συγγενική μεταμόσχευση.** Ο δότης και ο λήπτης είναι μονοωογενή δίδυμα αδέρφια. Έτσι, ενώ εκ πρώτης όψεως η μεταμόσχευση θα μπορούσε να χαρακτηριστεί ως αλλογενής, στην πράξη το μόσχευμα συμπεριφέρεται ως αυτόλογο. Το μόσχευμα δεν διαθέτει αντινεοπλασματική δράση. Γενικότερα, η περίπτωση αυτή μεταμόσχευσης είναι η σπανιότερη.

Ανάλογα με την **πηγή προέλευσης του μοσχεύματος**, διακρίνουμε τα ακόλουθα είδη μεταμοσχεύσεων:

**Μυελική μεταμόσχευση.** Το μόσχευμα λαμβάνεται από το μυελό των οστών, υπό γενική αναισθησία του δότη. Η λήψη του γίνεται κυρίως από τα λαγόνια οστά.

**Αιματική μεταμόσχευση.** Το μόσχευμα αποτελείται από κύτταρα του περιφερικού αίματος, τα οποία συλλέγονται από το αίμα με ειδικό μηχάνημα. Ο ασθενής ή ο δότης δεν υποβάλλεται σε γενική αναισθησία. Ο αριθμός των αρχέγονων αιμοποιητικών κυττάρων που κυκλοφορούν στο περιφερικό αίμα, υπό κανονικές συνθήκες, δεν είναι αρκετά υψηλός ώστε να παρέχει αξιόπιστο μόσχευμα προς αιμοποιία. Η συγκέντρωση των κυττάρων αυξάνεται με τη χρήση αυξητικών παραγόντων (αλλογενής μεταμόσχευση) ή και χημειοθεραπείας (αυτόλογη μεταμόσχευση), οι οποίοι κινητοποιούν τα αρχέγονα κύτταρα και τα αναγκάζουν να εξέλθουν μαζικά από τον μυελό στο αίμα. Είναι χαρακτηριστικό ότι η ποσότητα των κυττάρων σε αυτά τα μοσχεύματα είναι συνήθως μεγαλύτερη από αυτά του μυελού των οστών. Γενικότερα η μέθοδος είναι ασφαλής και δεν συνοδεύεται από αυξημένο κίνδυνο λευχαιμογένεσης.

**Ομφαλοπλακουντιακή μεταμόσχευση.** Το μόσχευμα προέρχεται από το αίμα του πλακούντα και του ομφάλιου λώρου και λαμβάνεται αμέσως μετά τον τοκετό.

## 2.3 Συμπλέγματα Ιστοσυμβατότητας

### 2.3.1 Μείζον Σύμπλεγμα Ιστοσυμβατότητας (MHC)

Σε όλα σχεδόν τα σπονδυλωτά, τα οποία έχουν έως σήμερα μελετηθεί και ιδίως στα θηλαστικά, συμπεριλαμβανομένου και του ανθρώπου, έχει διαπιστωθεί η ύπαρξη μιας στενά συνδεδεμένης ομάδας γονιδίων, τα προϊόντα της οποίας σχετίζονται με την κυτταρική αναγνώριση και με την διάκριση των αυτο- και αλλο- αντιγόνων.

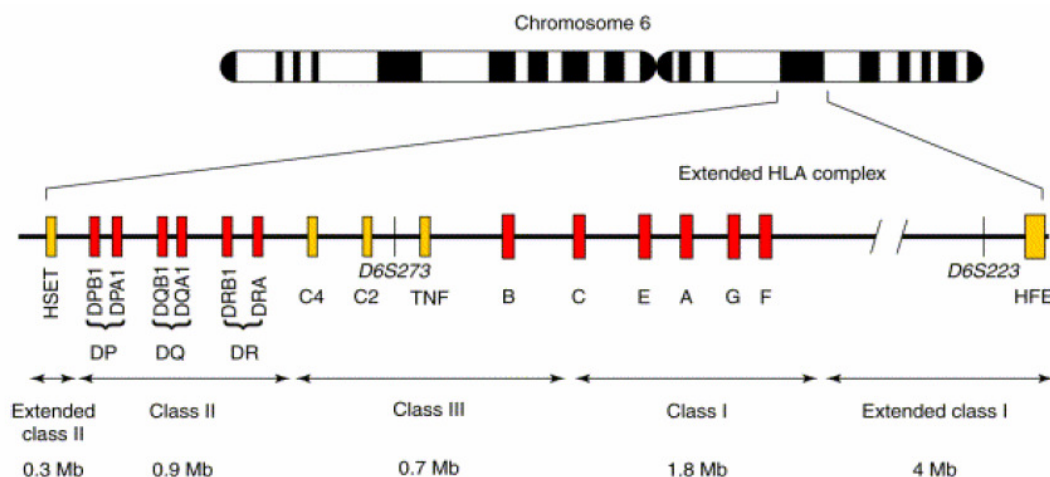
Η ομάδα αυτή των γονιδίων, που είναι σε μεγάλο βαθμό πολυμορφικά, καλείται **Μείζον Σύμπλεγμα Ιστοσυμβατότητας (MHC - Major Histocompatibility Complex)**. Το σύμπλεγμα MHC στον ποντικό αναφέρεται ως H-2 και στον άνθρωπο, ύστερα από πρόταση του *Jean Dausset* (βραβείο Nobel 1981), ως σύστημα **ανθρώπινων λευκοκυτταρικών αντιγόνων (HLA)**. Τα αντιγόνα αυτά ονομάστηκαν έτσι, επειδή, αναφέρονται στις πρωτεΐνες που κωδικοποιούνται από τα γονίδια του HLA συστήματος και εκφράζονται στην επιφάνεια της κυτταρικής μεμβράνης των λευκοκυττάρων. Ωστόσο, η ονομασία του συστήματος με τον όρο HLA, θεωρείται σήμερα ως η πλέον αντιπροσωπευτική.

Το νοηματικό περιεχόμενο του MHC προσδιορίζεται ως εξής:

- **Μείζον (Major)**, διότι, σε αντίθεση με τα ελάσσονα συστήματα ιστοσυμβατότητας (*minor Histocompatibility loci, mH*), είναι πολύ σημαντικό για την ταχεία ή μη απόρριψη μοσχευμάτων και την δημιουργία μεγάλου αριθμού κυτταροτοξικών αντισωμάτων.
- **Σύμπλεγμα (Complex)**, επειδή αποτελείται από πολλούς γενετικούς τόπους που είναι στενά συνδεδεμένοι μεταξύ τους και κληρονομούνται σαν σύνολο (*en block*).
- **Ιστοσυμβατότητας (Histocompatibility)**, γιατί η προσπάθεια αφορούσε, τουλάχιστον στα αρχικά στάδια των ερευνών, τη σημασία της ιστοσυμβατότητας στις μεταμοσχεύσεις οργάνων.

Τα γονίδια του HLA συμπλέγματος καθορίζουν την απαντητικότητα ή μη απαντητικότητα των T-λεμφοκυττάρων σε ποικίλα αντιγόνα. Για τον λόγο αυτό ονομάζονται και **γονίδια της ανοσιακής απάντησης (immune response genes, Ir)**. Το μείζον σύμπλεγμα ιστοσυμβατότητας του ανθρώπου είναι ένα πολύμορφο γενετικό σύστημα αποτελούμενο από περίπου 200 άνισα κατανεμημένα γονίδια. Κάθε άτομο έχει δύο σύνολα γονιδίων και κατά συνέπεια, οι συνδυασμοί των HLA

γονιδίων του κάθε ατόμου είναι σπάνιοι ή σχεδόν μοναδικοί σε διάφορους πληθυσμούς. Ο HLA πολυμορφισμός είναι ευεργετικός για την διενέργεια μελετών ανάμεσα σε διαφορετικούς πληθυσμούς, επειδή επιτρέπει τον προσδιορισμό της γενετικής συγγένειας μεταξύ διαφορετικών πληθυσμών. Η μελέτη των HLA αλληλοτύπων είναι επίσης σημαντική κατά την έρευνα πολύπλοκων γενετικών ασθενειών.



**ΕΙΚΟΝΑ 2.2** Η τοπολογία των HLA γονιδίων στο χρωμόσωμα 6

Το HLA σύστημα εντοπίζεται στο βραχύ σκέλος (p) του χρωμοσώματος 6 σε απόσταση 17cM από το κεντρομερίδιο. Το σύστημα αυτό αντιπροσωπεύει το 1/1000 της όλης γενετικής πληροφορίας του ανθρώπου και καταλαμβάνει έκταση η οποία κυμαίνεται από 2,5 Mbp (mega basis pair) έως 4 Mbp ή 4.000 Kb και αντιπροσωπεύει ένα από τα πλέον καλά μελετημένα βιολογικά συστήματα.

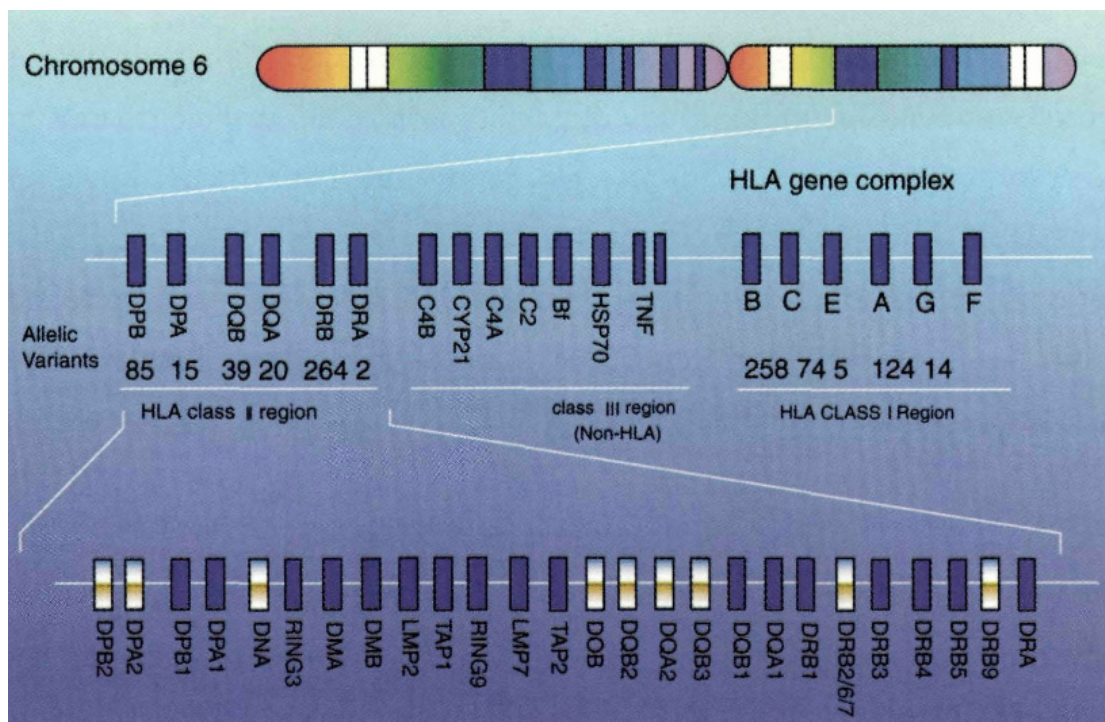
### 2.3.2 Ελάσσονα Συστήματα Ιστοσυμβατότητας

Βασική αιτία της απόρριψης ενός μοσχεύματος, παρά την πλήρη συμβατότητα μεταξύ δότη και λήπτη ως προς τα μείζονα συστήματα ιστοσυμβατότητας, είναι πιθανή η ύπαρξη **ελασσόνων συστημάτων ιστοσυμβατότητας (mH - minor Histocompatibility loci)**, η μελέτη των οποίων απασχολεί τους ερευνητές επί μακρόν. Η ονοματολογία τους και η τοπογραφία τους στο χρωμόσωμα χρονολογούνται από την εποχή που ο G. Snell περιέγραψε τους 13 γενετικούς τόπους (loci) στο σύστημα ιστοσυμβατότητας H-2 του ποντικού. Τελευταία, έχουν προσδιορισθεί περισσότεροι από 50 mH γενετικοί τόποι στο χρωμόσωμα του ποντικού. Ακόμη και σήμερα, η γνώση για τα ελάσσονα συστήματα ιστοσυμβατότητας στον άνθρωπο είναι πολύ περιορισμένη. Τελευταία δεδομένα συνηγορούν με την άποψη ότι αποτελούν ομάδα ενδογενών πεπτιδίων ή

υπεραντιγόνων, τα οποία σε συνεργασία με τα MHC μόρια, παίζουν καθοριστικό ρόλο στη ρύθμιση της ανοσιακής απάντησης.

### **2.3.3 Οργάνωση HLA γονιδίων**

Το HLA σύστημα αποτελείται από εκφραζόμενα γονίδια και ψευδογονίδια, τα οποία οργανώνονται σε **τρία μεγάλα αθροίσματα (clusters)** με κατεύθυνση από το κεντρομερίδιο προς το τελομερές ως εξής: τάξης I, II και III. Κάθε ένα από αυτά τα αθροίσματα αντιστοιχεί σε πολλές **γονιδιακές θέσεις ή τόπους (loci)** και κάθε ένας από τους τόπους αυτούς μπορεί να έχει στο χρωμόσωμα πολλές **γονιδιακές μορφές ή αλληλόμορφα γονίδια (alleles)**, ενώ στην επιφάνεια του κυττάρου εκφράζεται με μία μορφή, δηλαδή το αντίστοιχο HLA αντιγόνο. Οι σειρές των αλληλόμορφων των γονιδίων για τα τάξης I κωδικοποιούνται από τις γονιδιακές θέσεις HLA-A, -B, -C, για τα τάξης II από τις γονιδιακές θέσεις HLA-D, -DR, -DP, -DQ και για τα τάξης III από τις γονιδιακές θέσεις C2, BF, C4A, C4B. Άλλα γονίδια της HLA περιοχής είναι τα GLO (γλυοξαλάση των ερυθροκυττάρων), η 21-υδροξυλάση (κυτταρόχρωμα των επινεφριδίων υπεύθυνο για την 21-υδροξυλίωση των κορτικοειδών), καθώς και τα γονίδια  $\alpha$  και  $\beta$  για τον παράγοντα νέκρωσης των όγκων (TNF – Tumor Necrosis Factor). Δεν έχει ακόμα διευκρινιστεί εάν τα «αλλότρια» αυτά γονίδια περιλαμβάνονται τυχαία και αντιπροσωπεύουν αλληλουχίες από τα αρχικά στάδια της εξέλιξης ή έχουν κάποιον ιδιαίτερα σπουδαίο βιολογικό ρόλο. Αρκετά πιθανή είναι η ύπαρξη και άλλων παρόμοιων γονιδίων στην HLA περιοχή, η παρουσία των οποίων στο χρωμόσωμα πρέπει να προσδιοριστεί.



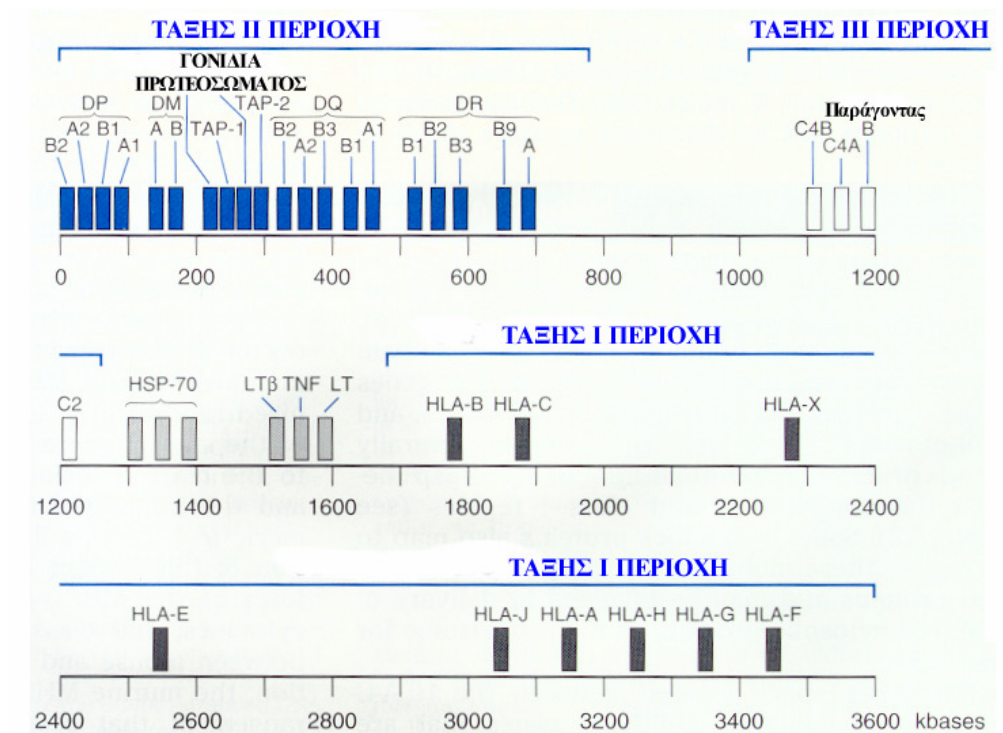
**ΕΙΚΟΝΑ 2.3** Η οργάνωση των HLA γονιδίων σε αθροίσματα τάξης I, II & III

Η **τάξης I γονιδιακή θέση (region)** περιλαμβάνει τρεις γενετικούς τόπους (loci) : τους HLA-A, -B και -C, που κωδικοποιούν τη βαριά αλυσίδα των τάξης I HLA μορίων και αποτελούν τα κλασσικά HLA γονίδια, ως θεμελιώδη συστατικά της ανοσιακής απάντησης. Με τις τεχνικές της μοριακής βιολογίας, έχουν πρόσφατα διαπιστωθεί νέα, μη κλασσικά HLA γονίδια τάξης I : τα HLA-E, -F, -G, και -H, των οποίων ο βιολογικός ρόλος δεν είναι ακόμα γνωστός. Ο HLA-E τύπος εντοπίζεται μεταξύ των HLA-B και HLA-C τύπων, ενώ οι HLA-F και HLA-G τόποι εδράζονται τελομερικά του HLA-A τύπου. Τα HLA-G μόρια εκφράζονται μόνο σε τροφοβλαστικά κύτταρα, σε πλακούντες πρώτου τριμήνου και σε κυτταρικές σειρές προερχόμενες από τέτοιους ιστούς. Διαδραματίζουν σημαντικό ρόλο στην ανοσιακή αλληλεπίδραση μητέρας και εμβρύου, εμποδίζοντας έτσι την πρόκληση ανοσιακής απάντησης της μητέρας έναντι του πλακούντα.

Η **τάξης II γονιδιακή θέση** περιλαμβάνει τις α και β αλυσίδες των DR και DQ γονιδίων, τα οποία εδράζονται κοντά στο κεντρομέρος του χρωμοσώματος 6 και καταλαμβάνουν έκταση περίπου 1.000 Kbp (1 Mbp). Στην ίδια γονιδιακή θέση περιλαμβάνονται και άλλα γονίδια, των οποίων ο βιολογικός ρόλος δεν έχει ακόμα διευκρινιστεί. Πλησίον της περιοχής DP εδράζεται άθροισμα γονιδίων όπως είναι τα RING1, RING2, TAP1, TAP2, DMB, DOB κ.ά., τα οποία παίζουν σημαντικό ρόλο στην επεξεργασία και παρουσίαση του HLA αντιγόνου. Κοινό χαρακτηριστικό των

πρωτεϊνών που κωδικοποιούνται από τα γονίδια της τάξης II γονιδιακής θέσης είναι οι δομικές ομοιότητες τους, καθώς και ο υψηλός βαθμός πολυμορφισμού τους.

Στην **τάξης III γονιδιακή θέση** έχουν διαπιστωθεί μέχρι σήμερα περισσότερα από 36 γονίδια, τα οποία διαδραματίζουν σημαντικό ρόλο στην ειδική ανοσιακή απάντηση. Το γονιδιακό αυτό τμήμα περιέχει γονιδιακές θέσεις που είναι υπεύθυνες για την παραγωγή πρωτεϊνών συμπληρώματος, ορμονών, ή για την επεξεργασία ενδοκυτταρικών πεπτιδίων και άλλων αναπτυξιακών χαρακτηριστικών. Έτσι η τάξης III γονιδιακή θέση δεν είναι στην πραγματικότητα ένα μέρος του HLA συστήματος, αλλά βρίσκεται εντός της HLA περιοχής, επειδή τα συστατικά της, είτε σχετίζονται με τις λειτουργίες των HLA αντιγόνων ή βρίσκονται υπό παρόμοιους μηχανισμούς ελέγχου σε σχέση με τα HLA γονίδια.



**ΕΙΚΟΝΑ 2.4** Οι περιοχές των HLA γονιδιακών θέσεων και τα γονίδια που περιλαμβάνουν

### *Άλλα γονίδια στην HLA γονιδιακή περιοχή*

Άλλοι γονιδιακοί τύποι που περιλαμβάνονται στην τάξης II HLA-περιοχή είναι:

1. Οι **HLA-DNA** και **-DOB**.
2. Οι **HLA-DMA** και **-DMB**. Βρίσκονται ανάμεσα στα DNA και DOB γονίδια και αποτελούν μέλη της οικογένειας των γονιδίων των ανοσοσφαιρινών (Ig gene family).

3. Οι **LMP2**, **LMP7**, **TAP1** και **TAP2**. Αυτοί οι γενετικοί τόποι κωδικοποιούν πρωτεΐνες που σχετίζονται με την επεξεργασία και μεταφορά του αντιγονικού πεπτιδίου από τα τάξης I HLA-αντιγόνα (ενδογενής ή τάξης I οδός παρουσίασης των αντιγόνων).
4. Ο **RING9**, του οποίου η λειτουργία δεν έχει προς το παρόν προσδιορισθεί.
5. Ο **RING3**, βρίσκεται ανάμεσα στα DNA και DMA γονίδια. Αυτό, εκφράζεται σε μια μεγάλη ποικιλία ιστών, η λειτουργία του όμως παραμένει προς το παρόν άγνωστη.
6. Οι **Y4** και **Y5** στη θέση του RING3. Πιθανόν το Y4 να ταυτίζεται με το RING3.
7. Ένα σύμπλεγμα γονιδίων που περιλαμβάνει το γονίδιο **11A2** του κολλαγόνου (**COL1A2**) και έξι γονίδια άγνωστης λειτουργίας (**RING1**, **RING2**, **RING5**, **KE3**, **KE4** και **KE5**). Τα γονίδια αυτά, όχι μόνο δεν φαίνεται να σχετίζονται λειτουργικά με το HLA ή το ανοσιακό σύστημα γενικότερα, αλλά ίσως και να οριοθετούν το τέλος του HLA.

Ανάμεσα στις περιοχές I και II, εντοπίζεται μιας υψηλής περιεκτικότητας σε γουανίνη και κυτοσίνη οικογένεια γονιδίων που καταλαμβάνει έκταση 1.100 Kb (1,1 Mbp). Αποτελεί την πιο πυκνή περιοχή του HLA με ένα γονίδιο κάθε 20Kb DNA. Τριάντα εννέα γονίδια έχουν ως τώρα ανακαλυφθεί.

Σε αυτή την HLA-περιοχή περιλαμβάνονται:

1. Τα γονίδια της κλασσικής οδού του συμπληρώματος **C4A**, **C4B** και **C2**.
2. Τα γονίδια του **παράγοντα B** (Bf, παλαιότερη ονομασία Factor B, BF), ο οποίος αποτελεί συστατικό της εναλλακτικής οδού του συμπληρώματος.
3. Τα γονίδια που κωδικοποιούν τους **παράγοντες νέκρωσης των όγκων** (**Tumor Necrosis Factor, TNF**) **TNF $\alpha$**  (156 αμινοξέα) και **TNF $\beta$**  (171 αμινοξέα), στα οποία αρχικά είχε αποδοθεί η ονομασία λεμφοτοξίνες (lymphotoxin).
4. Τα γονίδια που κωδικοποιούν τις **πρωτεΐνες του θερμικού σοκ (heat shock proteins, HSP70)**.
5. Κεντρομερικά του C4 γονιδίου εντοπίστηκαν τα γονίδια της **21-υδροξυλάσης**, **CYP21P** (ψευδογονίδιο, παλαιότερη ονομασία CYP21A), και **CYP21** (λειτουργικό γονίδιο, παλαιότερη ονομασία CYP21B).
6. Πρόσφατα εντοπίστηκαν, στην περιοχή των γονιδίων του συμπληρώματος, τρία νέα γονίδια: τα **RD**, **OSG** και **G11**.

### **2.3.4 Δομή HLA μορίων**

Η στερεοχημική δομή των τάξης I και II HLA αντιγόνων περιλαμβάνει τέσσερα τμήματα, παρεμφερή και στις δύο τάξεις. Το αμινοτελικό εξωκυττάριο τμήμα που συνδέεται με το αντιγόνο, το ανάλογο της ανοσοσφαιρίνης μη πολυμορφικό εξωκυττάριο τμήμα, το διαμεμβρανικό και το κυτταροπλασματικό τμήμα.

#### ***HLA μόρια τάξης I***

Τα HLA μόρια τάξης I είναι πολυμορφικές διαμεμβρανικές γλυκοπρωτεΐνες που βρίσκονται στην κυτταρική επιφάνεια όλων των εμπύρηνων κυττάρων του οργανισμού, με διαφορές ως προς την έκφραση και την ιστική κατανομή. Φυσιολογικός τους ρόλος είναι η παρουσίαση ξένων αντιγονικών πεπτιδίων, όπως οι ιοί, στα κυτταροτοξικά T-λεμφοκύτταρα. Στη μεταμόσχευση, τα μόρια τάξης I είναι τα κύρια αντιγόνα, που αναγνωρίζονται από τα κυτταροτοξικά T-λεμφοκύτταρα του λήπτη κατά την απόρριψη.

Κάθε HLA μόριο τάξης I αποτελείται από μια βαριά πολυπεπτιδική αλυσίδα των 340 αμινοξέων και από μια ελαφριά αλυσίδα, τη β2-μικροσφαιρίνη, των 99 αμινοξέων η οποία κωδικοποιείται από γονίδιο του χρωμοσώματος 15. Η βαριά αλυσίδα διασχίζει τη διπλοστοιβάδα της κυτταρικής μεμβράνης με το αμινοτελικό τμήμα της προς το εξωτερικό του κυττάρου, όπου ακριβώς συνδέεται με τη β2-μικροσφαιρίνη. Χωρίς αυτή τη σύνδεση το λειτουργικό τάξης I μόριο δεν είναι δυνατόν να εκφραστεί στην κυτταρική επιφάνεια.

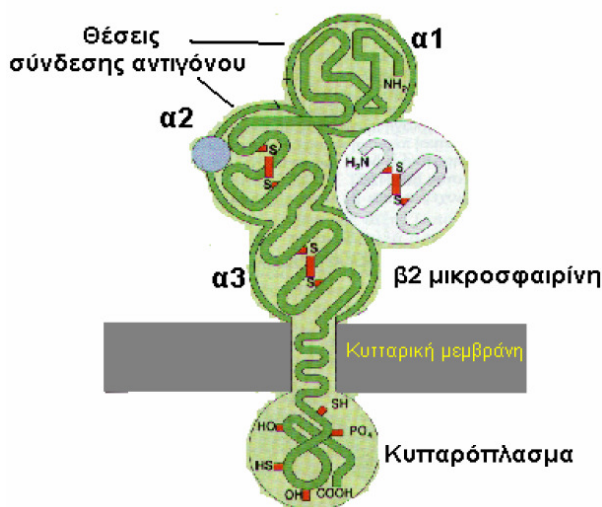
Στη βαριά αλυσίδα διακρίνονται τρεις διαφοροποιημένες δομικές περιοχές :

- α) Ένα εξωκυττάριο τμήμα από 274 αμινοξέα.
- β) Ένα υδρόφοβο διαμεμβρανικό τμήμα των 25 περίπου αμινοξέων υπεύθυνο για τη σταθεροποίηση του μορίου στη μεμβράνη.
- γ) Ένα υδρόφιλο ενδοκυττάριο ακραίο τμήμα αποτελούμενο από 30 αμινοξέα.

Με την κρυσταλλογράφηση των HLA μορίων τάξης I στις αρχές της δεκαετίας του 1990 μελετήθηκε η τρισδιάστατη δομή τους. Σε αυτήν τη διαμόρφωση δύο α-έλικες σχηματίζουν τα πλευρικά τοιχώματα μιας μακριάς και βαθιάς σχισμής, ο πυθμένας της οποίας αποτελείται από μία β-πτυχωτή επιφάνεια. Το αποτέλεσμα είναι η δημιουργία σχισμοειδούς θήκης, όπου συνδέεται το αντιγονικό πεπτίδιο, για να παρουσιαστεί στο CD8<sup>+</sup> T λεμφοκύτταρο. Τα εσωτερικά τοιχώματα της κοιλότητας της σχισμοειδούς θήκης δεν είναι ομαλά αλλά σχηματίζονται από μία σειρά κοιλωμάτων. Το αντιγονικό πεπτίδιο εισέρχεται μέσα στη θήκη, όπου μία ή

περισσότερες από τις πλευρικές του αλυσίδες προσαρμόζονται μέσα στα κοιλώματα της διαμορφώνοντας έτσι ένα μοριακό σύμπλεγμα ικανό να ανιχνευτεί από τον T κυτταρικό υποδοχέα.

Παρόλο που τα μόρια τάξης I βρίσκονται στην επιφάνεια όλων των εμπύρηνων κυττάρων το επίπεδο έκφρασής τους ποικίλει μεταξύ των διαφόρων τύπων των κυττάρων. Επιπλέον, η έκφραση αυξάνει κατά τη διάρκεια μιας ανοσολογικής απόκρισης. Γενικά, η έκφρασή τους επηρεάζεται θετικά ή αρνητικά από πολλούς παράγοντες, όπως οι κυτοκίνες, οι οποίες αυξάνουν ακόμα περισσότερο την έκφρασή τους. Η αυξημένη έκφραση συνοδεύει συνήθως αύξηση της ποσότητας του mRNA.

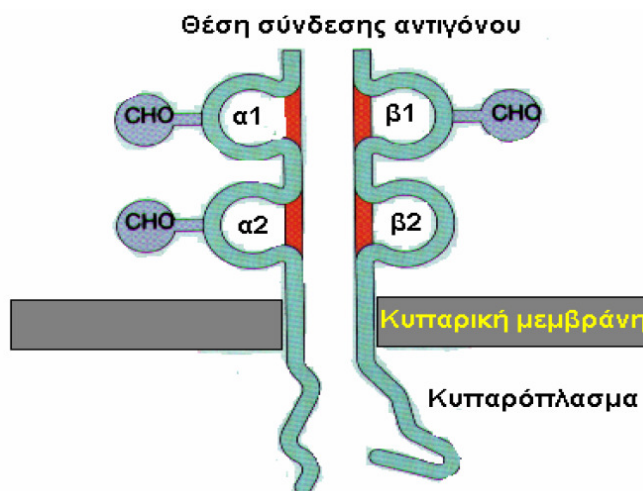


**ΕΙΚΟΝΑ 2.5** Η δομή του HLA μορίου τάξης I

### ***HLA μόρια τάξης II***

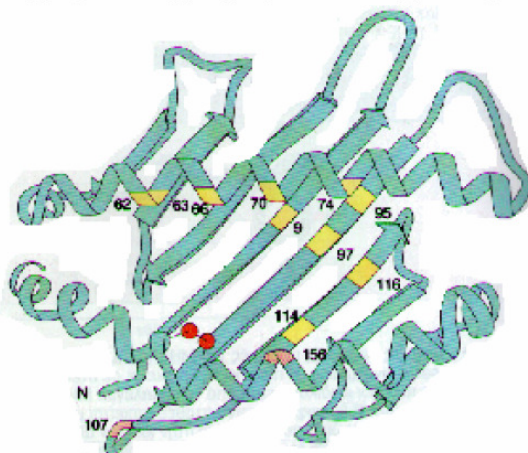
Τα HLA μόρια τάξης II είναι ετεροδιμερή που βρίσκονται κυρίως στα αντιγονοπαρουσιαστικά κύτταρα (APC - Antigen Presenting Cells) συμπεριλαμβανομένων των B-λεμφοκυττάρων των μακροφάγων, των δενδριτικών και των ενεργοποιημένων T-λεμφοκυττάρων. Αποτελούνται από μία βαριά γλυκοπρωτεϊνική αλυσίδα με αντίστοιχα α1, α2 εξωμεμβρανικά τμήματα, ένα διαμεμβρανικό τμήμα 30 αμινοξέων και ένα ενδομεμβρανικό 15 αμινοξέων και από μία ελαφριά αλυσίδα με β1 και β2 επικράτειες και όμοια τα δύο υπόλοιπα τμήματα. Υπάρχουν ενδείξεις, σύμφωνα με τις οποίες οι δύο αλυσίδες έχουν όμοια δομή. Έχει βρεθεί ότι οι α1, α2 και β1 είναι N-γλυκοσυλιωμένες πρωτεΐνες και ότι όλες οι επικράτειες εκτός της α1 σταθεροποιούνται με ενδοσουλφιδικούς δεσμούς και επίσης συναντούμε υδατανθρακικές ομάδες και στις δύο αλυσίδες. Κύριο χαρακτηριστικό των HLA μορίων τάξης II είναι ο έντονος πολυμορφισμός, που διαπιστώνεται σε γενετικό, ορολογικό και λειτουργικό επίπεδο και εντοπίζεται περισσότερο στις β από ότι στις α επικράτειες. Η κρυσταλλογράφηση του HLA-DR1 αντιγόνου επιβεβαίωσε

την υπόθεση ότι τα HLA μόρια τάξης II παρουσιάζουν μια τριτοταγή δομή ανάλογη με αυτή των HLA μορίων τάξης I. Βρέθηκε ότι οι  $\alpha 1$  και  $\beta 1$  περιοχές των τάξης II μορίων σχηματίζουν μια σχισμή όπου ενώνεται με το αντιγόνο.



**ΕΙΚΟΝΑ 2.6** Η δομή του HLA μορίου τάξης II

#### Σχισμοειδής θήκη σύνδεσης αντιγόνου



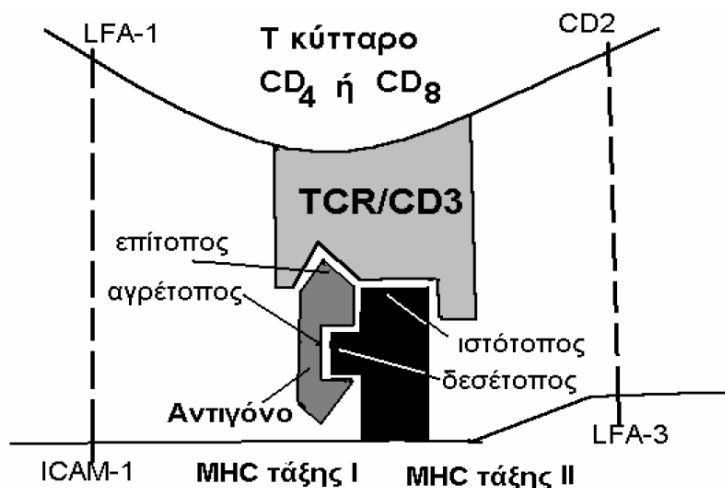
**ΕΙΚΟΝΑ 2.7** Η σχισμοειδής θήκη ως θέση σύνδεσης αντιγόνου & HLA μορίου

### 2.3.5 Λειτουργία HLA μορίων

Είναι γενικά αποδεκτό ότι το HLA σύστημα παίζει σημαντικό βιολογικό ρόλο στον οργανισμό. Μερικά στοιχεία, ενδεικτικά του MHC για τον οργανισμό, είναι η παρουσία του σε όλα τα σπονδυλωτά, ο τεράστιος πολυμορφισμός των HLA γονιδίων, η αποφυγή HLA ομοζυγωτισμού, που πιθανόν να είναι θανατηφόρος για το έμβρυο, η σημασία της HLA συμβατότητας στη μεταμόσχευση κ.ά. Κύρια λειτουργία του HLA συστήματος είναι η συμμετοχή του στην ανοσολογική απόκριση. Αυτή συνιστά ένα βιολογικό φαινόμενο, που ξεκινά όταν στον οργανισμό εισβάλλει μια ξένη ουσία. Σκοπός της αντίδρασης αυτής είναι η απομάκρυνση και η καταστροφή τέτοιων ουσιών. Η **ανοσολογική απόκριση** διακρίνεται σε **χημική** που συνίσταται

στην παραγωγή ανοσοσφαιρινών από τα πλασματοκύτταρα και σε **κυτταρική** με την οποία συντελείται η αναγνώριση του «ξένου» σώματος από τα T λεμφοκύτταρα. Το HLA σύστημα συμμετέχει στην αναγνώριση και παρουσίαση ξένων αντιγόνων (ιών, μικροβίων, μοσχευμάτων) καθώς και τροποποιημένων ίδιων αντιγόνων, π.χ. νεοπλασματικών κυττάρων, σε συνεργασία με τους διάφορους υποπληθυσμούς λεμφοκυττάρων και στην ενεργοποίηση των κυτταροτοξικών μηχανισμών.

Ο βασικός βιολογικός ρόλος των HLA αντιγόνων συνίσταται στον σχηματισμό συμπλέγματος με τα πεπτίδια των επεξεργασμένων πρωτεϊνικών αντιγόνων και στην παρουσίασή τους στην επιφάνεια των αντιγονοπαρουσιαστικών κυττάρων (APCs - Antigen Presenting Cells). Ο υποδοχέας των T-κυττάρων αναγνωρίζει το αντιγονικό πεπτίδιο μόνον όταν αυτό παρουσιάζεται ως σύμπλεγμα με το HLA αντιγόνο και επάγει την ενεργοποίηση των T-κυττάρων. Συγκεκριμένα, τα  $CD8^+$  κυτταρολυτικά T-κύτταρα αναγνωρίζουν τα ξένα αντιγόνα μόνον όταν παρουσιάζονται στην επιφάνεια των APCs συνδεδεμένα με τάξης I HLA μόρια, ενώ τα  $CD4^+$  βοηθητικά T-κύτταρα αναγνωρίζουν μόνον αντιγόνα συνδεδεμένα με τάξης II HLA μόρια. Ο πληθυσμός των T-λεμφοκυττάρων κατά κανόνα αγνοεί τα φυσικά αντιγόνα και αντιδρά είτε μόνο με μετουσιωμένο αντιγόνο ή με τμήματά του, συνδεδεμένα με HLA μόρια. Το φαινόμενο αυτό ονομάζεται **HLA περιορισμός**, που σημαίνει απλά ότι τα T-λεμφοκύτταρα δεν ενεργοποιούνται αν δεν αναγνωρίσουν ένα αντιγόνο συνδεδεμένο με HLA μόρια.



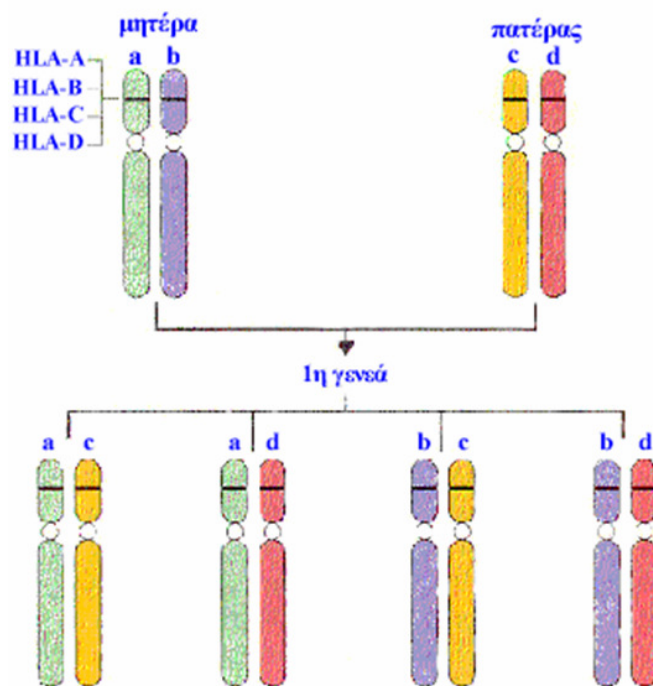
**ΕΙΚΟΝΑ 2.8** Σύνδεση αντιγόνου, MHC μορίου & υποδοχέα TCR

Ο μηχανισμός με τον οποίο ο υποδοχέας του T-λεμφοκυττάρου (TCR - T-Cell Receptor) αναγνωρίζει αντιγόνα συνδεδεμένα με HLA μόρια, έχει διαλευκανθεί μετά την ανακάλυψη της σχισμοειδούς θήκης. Πιο συγκεκριμένα, στα  $CD8^+$  λεμφοκύτταρα παρουσιάζονται ενδογενή αντιγονικά πεπτίδια (ικκά, νεοπλασματικά) προσαρμοσμένα

στη σχισμοειδή θήκη των HLA μορίων τάξης I, ενώ στα CD4<sup>+</sup> λεμφοκύτταρα παρουσιάζονται εξωγενή αντιγονικά πεπτίδια (μικροβιακά) συνδεδεμένα με HLA μόρια τάξεως II. Γίνεται λοιπόν φανερό ότι η απουσία HLA μορίων εμποδίζει την αναγνώριση και την καταστροφή ξένων αντιγόνων από τα CD4 και CD8 λεμφοκύτταρα, την παραγωγή αντισωμάτων από τα B-λεμφοκύτταρα καθώς και άλλες φάσεις συνεργασίας κυττάρων στην ανοσολογική απόκριση.

### 2.3.6 Κληρονομική Μεταβίβαση

Κάθε άτομο κληρονομεί δύο χρωμοσώματα 6, ένα από τον πατέρα του και ένα από την μητέρα του και κάθε χρωμόσωμα ένα μόνο από τα αλληλόμορφα γονίδια που αντιστοιχούν σε ένα συγκεκριμένο γονίδιο του HLA συστήματος. Τα αλληλόμορφα αυτά γονίδια κληρονομούνται ως δύο σύνολα, ένα από κάθε γονέα, χωρίς διάσπαση και ανασυνδυασμό μεταξύ των διαφόρων θέσεων, ενώ η κληρονομική μεταβίβαση υλοποιείται σύμφωνα με τους νόμους της κληρονομικότητας του Mendel. Κάθε σύνολο αλληλόμορφων γονιδίων ως συνδυασμός των HLA γονιδίων που συνδέονται στενά πάνω στο ίδιο χρωμόσωμα και κληρονομούνται από τον ένα γονέα ονομάζεται **απλότυπος (haplotype)**. Έτσι, ένα άτομο κληρονομεί έναν απλότυπο από τη μητέρα και έναν από τον πατέρα. Δύο απλότυποι, ένας από κάθε γονέα, αποτελούν τον **γονότυπο (genotype)** του ατόμου, ενώ οι δύο απλότυποι ενός ατόμου σύμφωνα με τη σειρά των γονιδίων πάνω στο χρωμόσωμα, όπως αυτά προκύπτουν από την HLA τυποποίηση παριστά τον **φαινότυπο (phenotype)**.



**ΕΙΚΟΝΑ 2.9** Κληρονομική μεταβίβαση HLA γονιδίων

Λόγω του γεγονότος ότι κάθε άτομο έχει δύο χρωμοσώματα 6 σε κάθε κύτταρο του οργανισμού του, ένα πατρικής και ένα μητρικής προέλευσης, έχει κατά συνέπεια 14 HLA γονίδια, δύο για κάθε γονιδιακή θέση. Ένα άτομο μπορεί να φέρει δύο ίδια αλληλόμορφα γονίδια (ομόζυγο) ή δύο διαφορετικά (ετερόζυγο). Σύμφωνα με τον τρόπο της κληρονομικής μεταβίβασης των HLA γονιδίων στους απογόνους, για δύο αδέρφια ισχύουν οι εξής πιθανότητες :

- 25% πιθανότητα να έχουν κοινούς και του δύο απλοτύπους (πατρικό και μητρικό), οπότε ονομάζονται ταυτόσημα.
- 50% πιθανότητα να έχουν κοινό μόνον έναν απλότυπο, οπότε ονομάζονται απλοταυτόσημα.
- 25% πιθανότητα να μην έχουν κοινό κανέναν απλότυπο.

Σε σπάνιες περιπτώσεις μπορεί να συμβεί ανασυνδυασμός του γονιδίου, οπότε και σχηματίζεται ένας καινούργιος απλότυπος. Χαρακτηριστικό του HLA συστήματος είναι η αυξημένη συχνότητα εμφάνισης ορισμένων συνδυασμών αλληλόμορφων γονιδίων εντός ενός πληθυσμού.

### **2.3.7 Κλινικές Εφαρμογές HLA Συστήματος**

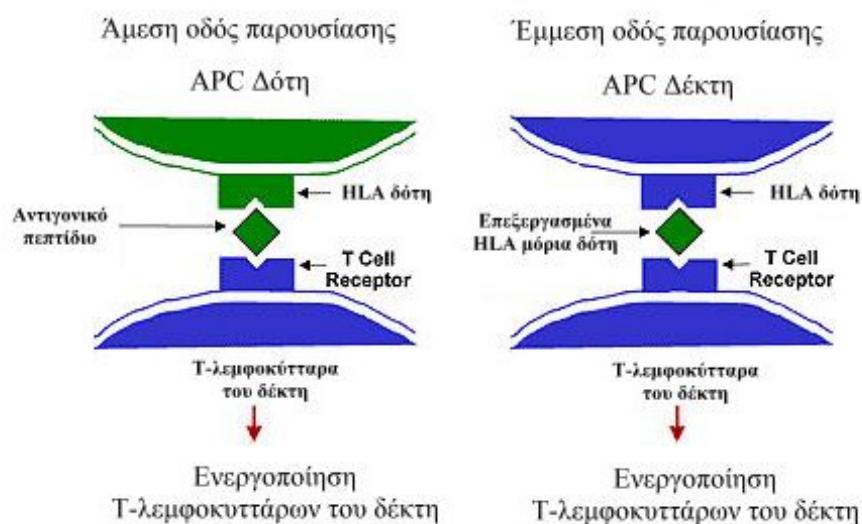
Οι κύριες εφαρμογές του HLA συστήματος είναι οι εξής:

- α) Στις μεταμοσχεύσεις ιστών και οργάνων.
- β) Στον έλεγχο αμφισβητούμενης πατρότητας.
- γ) Σε ανθρωπολογικές μελέτες και φυλογενετική ανάλυση πληθυσμών.
- δ) Σε συσχετισμούς με ασθένειες.

#### **2.3.7.1 Μεταμοσχεύσεις Ιστών και Οργάνων**

Είναι γνωστό ότι η μεταμόσχευση ιστών ή οργάνων από άτομο σε άτομο, ακόμα και του ίδιου είδους δεν είναι πάντα επιτυχής. Ο μεταμοσχευθής ιστός ή όργανο καταστρέφεται και απορρίπτεται από τον οργανισμό του λήπτη με εξαίρεση στα μονογενή δίδυμα. Η ασυμβατότητα αυτή και η αδυναμία μεταμόσχευσης οφείλεται στην ύπαρξη διαφορετικών αντιγόνων ιστοσυμβατότητας στα κύτταρα του δότη και του λήπτη. Τα αντιγόνα αυτά προκαλούν ειδική ανοσολογική αντίδραση, κυτταρική και χημική, στην οποία οφείλεται η απόρριψη του μοσχεύματος. Κάθε άτομο στα κύτταρά του φέρει τέσσερα διαφορετικά αντιγόνα του HLA συστήματος, ένα ζεύγος πατρικής και ένα ζεύγος μητρικής προέλευσης.

Τα αντιγόνα ιστοσυμβατότητας βρίσκονται σε κάθε εμπύρνηνο κύτταρο και αναγνωρίζονται ειδικά από το λήπτη του μοσχεύματος, που έχει διαφορετικούς δείκτες ατομικότητας, δηλαδή αντιγόνα ιστοσυμβατότητας. Συγκεκριμένα, τα Τ-λεμφοκύτταρα του δέκτη αναγνωρίζουν τα ξένα HLA του δότη είτε άμεσα ως άθικτα μόρια είτε έμμεσα ως πεπτίδια, τα οποία προκύπτουν από την αντιγονική επεξεργασία. Κατά την άμεση οδό παρουσίασης, τα αντιγονοπαρουσιαστικά (APC) κύτταρα του δότη παρουσιάζουν τα HLA αντιγόνα του δότη κατευθείαν στον Τ-κυτταρικό υποδοχέα ενός Τ-κυττάρου του λήπτη. Η άμεση οδός παρουσίασης θεωρείται υπεύθυνη για την οξεία απόρριψη και γι' αυτόν τον λόγο πριν τη μεταμόσχευση χρησιμοποιούνται διάφορες αγωγές στο μόσχευμα. Κατά την έμμεση οδό παρουσίασης, τα HLA μόρια του δότη, αφού πρώτα υποβληθούν σε κατάλληλη επεξεργασία, μετατραπούν δηλαδή σε μικρότερα πεπτίδια από τα APC του λήπτη, παρουσιάζονται από τα αντιγονοπαρουσιαστικά κύτταρα του λήπτη και αναγνωρίζονται με κλασικές διαδικασίες αναγνώρισης από τα Τ-κύτταρα του λήπτη.



**ΕΙΚΟΝΑ 2.10** Ο ρόλος των HLA μορίων στις κύριες οδούς παρουσίασης των αντιγόνων του μοσχεύματος του δότη στα Τ-λεμφοκύτταρα του λήπτη

Η επιτυχία μιας μεταμόσχευσης βασίζεται σε λεπτομερειακό ανοσολογικό έλεγχο στον δότη και τον λήπτη, που προηγείται, ώστε να τεκμηριωθεί η συμβατότητα του μοσχεύματος. Μετά τη μεταμόσχευση θα πρέπει να ελέγχεται στο λήπτη η ανοσοβιολογική κατάσταση καθώς και η απόδοση της ανοσοκατασταλτικής αγωγής, έτσι ώστε να αποτραπεί η απόρριψη του μοσχεύματος.

### **2.3.7.2 Έλεγχος Αμφισβητούμενης Πατρότητας**

Η συμβολή των HLA αντιγόνων στον αποκλεισμό της πατρότητας ήρθε να προστεθεί σε ήδη χρησιμοποιούμενα συστήματα ελέγχου πατρότητας (αντιγόνα και ένζυμα ερυθροκυττάρων, πρωτεΐνες ορού). Η προσθήκη αυτή συνέβαλε στην αύξηση της επιτυχίας από 70% σε 99%. Πρέπει να σημειωθεί ότι μπορούμε να μιλήσουμε για αποκλεισμό πατρότητας με βεβαιότητα 100% και όχι για πιθανή πατρότητα.

Τα αντιγόνα HLA είναι κατάλληλα για αποκλεισμό για τους εξής λόγους:

- α) Κληρονομούνται με γνωστό και σταθερό τρόπο.
- β) Παρουσιάζουν υψηλό ποσοστό πολυμορφισμού.
- γ) Παραμένουν σταθερά σε όλη τη ζωή.
- δ) Ελέγχονται με μεθόδους ακριβείς και επαναλήψιμες.

Εφόσον ένα παιδί έχει κληρονομήσει έναν HLA απλότυπο από κάθε γονέα, η σύγκριση του απλοτύπου, που είναι αποδεδειγμένα πατρικής προέλευσης, με τους απλότυπους του αμφισβητούμενου πατέρα είναι αυτή, που θα μας δώσει τα ασφαλή συμπεράσματα. Είναι ευνόητο ότι όσο πιο σπάνια είναι τα κοινά αντιγόνα πατέρα και παιδιού τόσο περισσότερες είναι οι πιθανότητες να είναι ο φυσικός πατέρας. Αντίθετα, αν δεν υπάρχει κοινός απλότυπος, μπορούμε να πούμε με βεβαιότητα ότι αποκλείεται η πατρότητα.

### **2.3.7.3 Ανθρωπολογικές Μελέτες & Φυλογενετική Ανάλυση Πληθυσμών**

Το HLA σύστημα κατέχει σημαντική θέση στις βιολογικές μεθόδους για τον προσδιορισμό των φυλετικών τύπων ενός λαού, επειδή τα HLA αντιγόνα απαντώνται σε διαφορετικές συχνότητες στις διάφορες φυλές, γι' αυτό υποβοηθούν πολύ στον καθορισμό της γενετικής απόστασης μεταξύ τους. Από παρατηρήσεις που έγιναν φαίνεται ότι π.χ. το HLA-A1 απαντάται συχνά στη λευκή φυλή, ενώ αντίθετα βρίσκεται σε χαμηλότερη συχνότητα στην κίτρινη και τη μαύρη φυλή.

#### **2.3.7.4 Συσχετισμοί με Ασθένειες**

Πολλές ασθένειες έχουν συσχετιστεί με τα HLA αντιγόνα. Οι πιο πολλές συσχετίσεις με HLA μόρια αφορούν αυτοάνοσα νοσήματα που συνδέονται κυρίως με HLA αντιγόνα τάξης II και έχουν τα ακόλουθα χαρακτηριστικά:

1. Είναι άγνωστης αιτιολογίας.
2. Κληρονομούνται συχνά, αλλά με χαμηλή διεισδυτικότητα.
3. Δεν μεταβιβάζονται με τους νόμους κληρονομικότητας του Mendel.
4. Συνοδεύονται από ανοσολογικές ανωμαλίες.
5. Έχουν χαμηλή επίδραση στην αναπαραγωγή.

Όπως αναφέρθηκε, οι διάφορες σχετιζόμενες παθήσεις είναι αυτοάνοσης αιτιολογίας. Οι μηχανισμοί, οι οποίοι ενέχονται στην ανοσοπαθογένεια των νοσημάτων δεν έχουν γίνει ακόμα γνωστοί. Ένας από τους μηχανισμούς που έχουν προταθεί είναι η θεωρία του «μοριακού μιμητισμού», σύμφωνα με την οποία το αντιγόνο ιστοσυμβατότητας που σχετίζεται με τη νόσο είναι από ανοσογονικής πλευράς όμοιο με τον αιτιολογικό παράγοντα της νόσου, και για αυτό δεν επάγεται ανοσοβιολογική απόκριση, με αποτέλεσμα να προκαλεί αναπόφευκτα σοβαρές βλάβες. Επίσης, είναι ενδεχόμενη και η εμφάνιση ισχυρής ανοσοβιολογικής απόκρισης έναντι του αιτιολογικού παράγοντα, οπότε λόγω της ομοιότητας η ανοσοβιολογική απόκριση στρέφεται και κατά του HLA «εαυτού». Η συγκεκριμένη ανοσοαπόκριση καταλήγει προφανώς στην εμφάνιση της νόσου. Η έρευνα συσχέτισης των ασθενειών με το HLA σύστημα γίνεται με πληθυσμιακές και οικογενειακές μελέτες.

Η συσχέτιση προσδιορίζεται με δύο δείκτες:

- α) τον σχετικό κίνδυνο (RR) και
- β) τον απόλυτο κίνδυνο (AR).

Ο σχετικός κίνδυνος δείχνει τον κίνδυνο που έχει κάποιος να αναπτύξει τη νόσο όταν φέρει το αντιγόνο σε σχέση με κάποιον που δεν το φέρει. Ο απόλυτος κίνδυνος αναφέρεται στο βαθμό κινδύνου να νοσήσει ένα άτομο που φέρει το αντιγόνο σε σχέση με τη συχνότητα της νόσου στον γενικό πληθυσμό. Η διερεύνηση των συσχετίσεων HLA παραγόντων με νοσήματα μπορούν να έχουν χρησιμότητα στην κλινική πράξη, συμβάλλοντας στη διάγνωση νοσήματος και στην εφαρμογή πιθανών προληπτικών μέτρων. Επίσης, οι HLA δείκτες διαφοροποιούν μορφές και υποομάδες νοσημάτων και συσχετίζονται με κλινική πορεία ή/και πρόγνωση.

**ΠΙΝΑΚΑΣ Α.** Συσχετίσεις HLA αντιγόνων τάξης I & II με νοσήματα.

| <b>ΝΟΣΗΜΑ</b>                             | <b>ΣΥΣΧΕΤΙΖΟΜΕΝΑ<br/>HLA ΑΝΤΙΓΟΝΑ</b> |
|-------------------------------------------|---------------------------------------|
| Ρευματοειδής Αρθρίτιδα                    | DR4                                   |
| Συστηματικός Ερυθηματώδης Λύκος           | DR3                                   |
| Μυασθένεια Graves                         | B8, DR3                               |
| Νόσος Graves                              | B8, DR3                               |
| Θυρεοειδίτιδα Hashimoto                   | DR5                                   |
| Θυρεοειδίτιδα μετά τον τοκετό             | DR4                                   |
| Σακχαρώδης Διαβήτης Ινσουλινοεξαρτώμενος  | DR3, DR4                              |
| Νόσος Addison                             | DR3                                   |
| Κακοήθης Αναιμία                          | DR5                                   |
| Πέμφιγα                                   | DR4                                   |
| Πέμφιγα κύησης                            | DR2, DR4                              |
| Ερπητοειδείς Δερματίτιδα                  | DR3                                   |
| Ανεπάρκεια IgA                            | DR3                                   |
| Σύνδρομο Goodpasture                      | DR2                                   |
| Ιδιοπαθής Μεμβρανώδης Σπειραματονεφρίτιδα | DR3                                   |
| Σκλήρυνση κατά πλάκας                     | DR2,B7                                |
| Αγκυλοποιητική Σπονδυλίτιδα               | B27                                   |
| Υποξεία Θυρεοειδίτιδα                     | B35                                   |
| Σύνδρομο Reiter                           | B27                                   |
| Ψωρίαση                                   | C6                                    |

### **2.3.8 Πολυμορφισμός HLA Συστήματος**

Πολυμορφισμός είναι το φαινόμενο κατά το οποίο διαφορετικά αλληλόμορφα γονίδια μπορούν να ανιχνευθούν σε έναν μόνο γονιδιακό τόπο.

#### **2.3.8.1 Χαρακτηριστικά Πολυμορφισμού**

Τα **βασικά χαρακτηριστικά** του πολυμορφισμού στο MHC είναι τα εξής:

- Υπάρχουν πολλά αλληλόμορφα γονίδια σε κάθε λειτουργικό τόπο του συστήματος και κάθε ένα από αυτά εμφανίζεται με αξιόλογη συχνότητα στον πληθυσμό. Αυτό έρχεται σε αντίθεση με το φαινόμενο το οποίο παρατηρείται σε άλλα γενετικά συστήματα, στα οποία κάθε τύπος έχει, είτε ένα μόνο αλληλόμορφο, είτε λίγα αλληλόμορφα από τα οποία μόνο ένα είναι το πιο συχνό. Επειδή υπάρχουν πολλοί HLA γονιδιακοί τόποι και κάθε τύπος έχει πολλά αλληλόμορφα, είναι θεωρητικά δυνατοί πολλοί συνδυασμοί αλληλομόρφων ( $10^{12}$ ). Ωστόσο, λόγω της στενής σύνδεσης των γονιδίων του HLA συστήματος και της κληρονομικότητάς τους, μόνο ένα μέρος αυτών των πιθανών συνδυασμών εντοπίζεται στην πραγματικότητα.

- Οι διαφορές ορισμένων αλληλομόρφων γονιδίων στο HLA σύστημα μπορεί να οφείλονται σε 100 ή περισσότερες αντικαταστάσεις νουκλεοτιδίων.
- Ο πολυμορφισμός εντοπίζεται στα α1 και α2 πεδία των τάξης I HLA αντιγόνων, στα α1 και β1 πεδία των τάξης II HLA-αντιγόνων και ιδιαίτερα στο έδαφος της σχισμοειδούς θήκης, καθώς και τις εσωτερικές και τις ανώτερες επιφάνειες των α-ελίκων.

### **2.3.8.2 Βιολογική Σημασία Πολυμορφισμού**

Ο πολυμορφισμός πιστεύεται ότι καθιστά δυνατή την αντιμετώπιση μεγάλου εύρους παθογόνων μικροοργανισμών, τα οποία απειλούν τον οργανισμό και ευθύνεται για την ειδικότητα της αναγνώρισης του αντιγόνου από τα CD4<sup>+</sup> T-λεμφοκύτταρα, καθώς και για την παρατηρούμενη ποικιλία στον τρόπο ανοσιακής απάντησης διαφόρων ατόμων που εκτίθενται στα ίδια αντιγόνα. Έτσι, ο πολυμορφισμός μπορεί να συμμετέχει στην επιδεκτικότητα του ατόμου σε διάφορες ασθένειες. Επίσης, αυξάνει την ποικιλία και τον αριθμό των αντιγονικών πεπτιδίων που παρουσιάζονται από τα HLA αντιγόνα. Τέλος, η ύπαρξη τόσο μεγάλου πολυμορφισμού έχει ως αποτέλεσμα την μείωση της πιθανότητας ανεύρεσης μη συγγενών ατόμων, πλήρως συμβατών με τον υποψήφιο για μεταμόσχευση λήπτη (περίπου 1:100.000 ή μικρότερη).

### **2.3.9 Τυποποίηση HLA Συστήματος**

Η **ανίχνευση και τυποποίηση** των HLA αντιγόνων που φέρει ένα άτομο είναι δυνατόν να πραγματοποιηθεί με δύο βασικές τεχνικές :

- 1) Σε επίπεδο γονιδίου με μοριακές τεχνικές.
- 2) Σε επίπεδο κυττάρου με ορολογικές μεθόδους.

Η παραδοσιακή μέθοδος HLA τυποποίησης, δηλαδή προσδιορισμού των HLA αντιγόνων, είναι η **ορολογική μέθοδος (Serology – Microlymphocytotoxicity method)**. Η μέθοδος αυτή χρησιμοποιείται ακόμα και σήμερα από εργαστήρια σε παγκόσμια κλίμακα. Όμως, η εκθετική ανάπτυξη της τεχνολογίας του DNA κατά τα τελευταία χρόνια δεν άφησε ανεπηρέαστο και το πεδίο της ιστοσυμβατότητας. Ακολουθώντας τις σύγχρονες ιατρικές εξελίξεις, όλο και περισσότερα εργαστήρια τείνουν να αντικαταστήσουν την παραδοσιακή ορολογική μέθοδο με αυτές που βασίζονται στο DNA (DNA based methods) όσον αφορά στην τυποποίηση των HLA αντιγόνων και αλληλίων, οι οποίες υπερτερούν έναντι της παραδοσιακής μεθόδου.

Ιδιαίτερα στο πεδίο των μεταμοσχεύσεων μυελού των οστών, η ορολογική μέθοδος κρίνεται μάλλον μη ικανοποιητική σήμερα και έχει εγκαταλειφθεί από τα περισσότερα εργαστήρια τυποποίησης για μεταμοσχεύσεις μυελού των οστών. Τα HLA αντιγόνα διαφέρουν το ένα από το άλλο στις πολυπεπτιδικές τους αλυσίδες μόνο σε μεμονωμένα αμινοξέα. Η αναγνώριση αυτών των ταυτώσεων σε μεγάλο βαθμό δομών είναι σχεδόν αδύνατη ορολογικά. Για αυτόν το λόγο, η ευκρίνεια της ορολογικής μεθόδου HLA τυποποίησης είναι περιορισμένη. Εφόσον οι DNA αλληλουχίες των πιο σημαντικών HLA αλληλίων είναι πλέον γνωστές, παραλλαγές στις αλληλουχίες μπορούν να αναγνωριστούν στο επίπεδο του DNA με την βοήθεια των συνθετικών ολιγονουκλεοτιδίων. Χρησιμοποιώντας την ενίσχυση του γενομικού DNA (PCR - Polymerase Chain Reaction, Αλυσιδωτή Αντίδραση Πολυμεράσης) μαζί με ζεύγη ειδικών εκκινητών (SSP - Sequence Specific Primers), είναι δυνατόν να ταυτοποιηθεί ένας μεγάλος αριθμός των HLA αλληλίων με τις μοριακές μεθόδους. Παρόλα αυτά όμως, σε παγκόσμια κλίμακα, συνεχίζουν να υφίστανται δυσκολίες που αφορούν αφενός την «επικοινωνία» των αποτελεσμάτων της ορολογικής και της **βασισόμενης στο DNA τυποποίησης** και αφετέρου την «επικοινωνία» των αποτελεσμάτων μεταξύ των επί μέρους διαφορετικών τεχνικών της μοριακής τυποποίησης που χρησιμοποιούνται.

### **ΕΠΙΠΕΔΑ HLA ΤΥΠΟΠΟΙΗΣΗΣ**

Με βάση την ποιότητα της HLA τυποποίησης μπορούμε να λάβουμε αποτελέσματα σε πέντε διαφορετικά επίπεδα:

1. Εκτεταμένα ορολογικά αντιγόνα (Broad)
2. Αποκομμένα ορολογικά αντιγόνα (Split)
3. Χαμηλής Ανάλυσης (LR – Low Resolution) DNA τυποποίηση
4. Ενδιάμεσης Ανάλυσης (IR - Intermediate) DNA τυποποίηση
5. Υψηλής Ανάλυσης (HR – High Resolution) DNA τυποποίηση

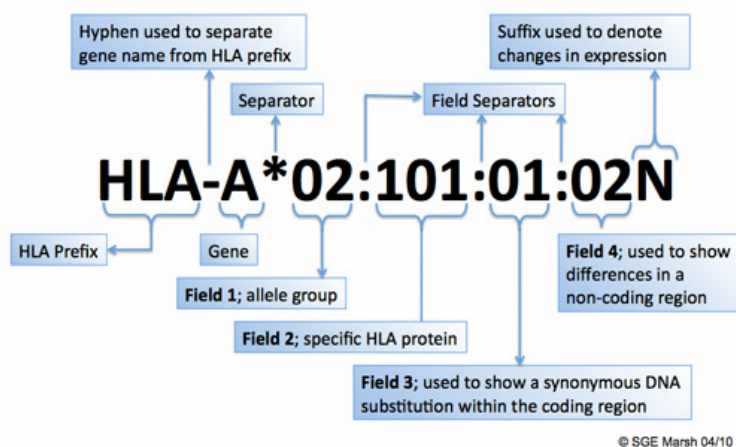
Τα αποτελέσματα τυποποίησης εκτεταμένων και αποκομμένων ορολογικών αντιγόνων βασίζονται σε ορολογικές μεθόδους τυποποίησης σε κυτταρικό επίπεδο, ενώ τα υπόλοιπα αποτελέσματα βασίζονται σε μεθόδους τυποποίησης μοριακής βιολογίας σε γονιδιακό επίπεδο. Η **χαμηλή ανάλυση** αναφέρεται στην αναγνώριση για τα δύο πρώτα ψηφία της HLA ονοματολογίας, δηλαδή όλων των αλληλόμορφων με το ίδιο πρώτο πεδίο. Η **ενδιάμεση ανάλυση** αναφέρεται στην επιλογή μεταξύ τουλάχιστον δύο διαφορετικών κωδικών αλληλόμορφων, τα οποία ανήκουν στην ίδια ομάδα ορολογικού αντιγόνου. Η **υψηλή ανάλυση** συνήθως αναφέρεται σε ένα

αλληλόμορφο με δύο ή περισσότερα πεδία (τουλάχιστον τέσσερα ψηφία). Σε ορισμένες χώρες (π.χ. Γερμανία), οι πολλαπλοί κώδικες αλληλόμορφου εξακολουθούν να θεωρούνται ως υψηλής ανάλυσης αποτέλεσμα, εάν όλα τα αλληλόμορφα είναι ταυτόσημα επί των εξονίων 2 και 3 για τα HLA τάξης I γονίδια ή επί του συνόλου του εξονίου 2 για το HLA τάξης II γονίδια. Η HLA ονοματολογία διαφέρει ως προς το είδος της μεθόδου της HLA τυποποίησης που χρησιμοποιείται (ορολογική ή μοριακή).

### **2.3.10 Ονοματολογία HLA Συστήματος**

Η ονοματολογία (nomenclature) του HLA συστήματος είναι υπό την ευθύνη της Επιτροπής Ονοματολογίας του **Παγκόσμιου Οργανισμού Υγείας (WHO – World Health Organization)**. Επίσης, τα ανά τον κόσμο μητρώα αιμοποιητικών βλαστικών κυττάρων εθελοντών δοτών μυελού των οστών ακολουθούν τα πρότυπα και τις συστάσεις της **Παγκόσμιας Ένωσης Δοτών Μυελού των Οστών (WMDA - World Marrow Donor Association)**.

Η ονομασία κάθε αλληλόμορφου **HLA γονιδίου** έχει δύο, τρία ή τέσσερα μοναδικά ονόματα πεδίων. Τα πεδία διαχωρίζονται μεταξύ τους με το σύμβολο της άνω και κάτω τελείας (« : »). Το μήκος της ονομασίας του αλληλόμορφου εξαρτάται από την αλληλουχία του αλληλόμορφου και του πλησιέστερου συγγενή του. Όλα τα αλληλόμορφα λαμβάνουν τουλάχιστον ένα όνομα δύο πεδίων, ενώ ονόματα τριών και τεσσάρων πεδίων αποδίδονται μόνον όταν αυτό είναι απαραίτητο. Το πρώτο πεδίο (αριθμός) περιγράφει τον τύπο, ο οποίος συχνά αντιστοιχεί στο ορολογικό αντιγόνο. Το δεύτερο πεδίο (αριθμός) χρησιμοποιείται για την απαρίθμηση των υποτύπων, από αριθμούς που έχουν ανατεθεί με τη σειρά με την οποία έχουν καθοριστεί οι αλληλουχίες του DNA. Αλληλόμορφα των οποίων οι αριθμοί διαφέρουν στα δύο πρώτα πεδία πρέπει να διαφέρουν σε μία ή περισσότερες νουκλεοτιδικές υποκαταστάσεις που αλλάζουν την αλληλουχία των αμινοξέων της κωδικοποιημένης πρωτεΐνης.



**EIKONA 2.11** Ονοματολογία αλληλόμορφων HLA γονιδίων

Από το 1989 η ονοματολογία των παραγόντων του HLA συστήματος αλλάζει γρήγορα. Το φαινόμενο αυτό οφείλεται στο γεγονός ότι συνέχεια ανακαλύπτονται καινούργια αλληλόμορφα HLA γονίδια. Επιπλέον, ο διαφορετικός τρόπος γραφής των γονιδίων του HLA συστήματος, ανάλογα με τη μέθοδο (ορολογική ή μοριακή) που χρησιμοποιήθηκε για τον προσδιορισμό τους, αποτελεί συχνό πρόβλημα για τα άτομα που αρχίζουν να μελετούν το μείζον σύμπλεγμα ιστοσυμβατότητας. Αρχικά ο προσδιορισμός των αντιγόνων του HLA συστήματος στηρίχθηκε σε ορολογικές μεθόδους με την χρησιμοποίηση αντι-ορών (οροί που περιέχουν αντισώματα έναντι κάποιων συγκεκριμένων HLA αντιγόνων). Είναι όμως γνωστό ότι τα αντισώματα μπορούν να αναγνωρίσουν επιτόπους που είναι κοινοί σε περισσότερα HLA αντιγόνα. Έτσι, είναι δύσκολο, πολλές φορές, να αναγνωριστεί ποια είναι τα HLA αντιγόνα που βρίσκονται στα λεμφοκύτταρα του εξεταζόμενου ατόμου. Επιπλέον, μπορεί, εξαιτίας του μεγάλου πολυμορφισμού του HLA συστήματος, HLA αντιγόνα που φαίνονται, με την ορολογική μέθοδο ίδια, να διαφέρουν, σε μοριακό επίπεδο, στην αλληλουχία των αμινοξέων τους στην πρωτεϊνική αλυσίδα.

Κάθε αντιγόνο του μείζονος συμπλέγματος ιστοσυμβατότητας που προσδιορίζεται με **ορολογικές μεθόδους** χαρακτηρίζεται ως εξής :

1. Με το σύμβολο *HLA* που υποδηλώνει το σύμπλεγμα ιστοσυμβατότητας.
2. Ένα γράμμα που αντιπροσωπεύει τον γονιδιακό τύπο (*HLA-A*, *-DR*).
3. Ένα αριθμό που χαρακτηρίζει το αντιγόνο (*HLA* ή *HLA-A1*, *-DR4*, *-DQ5*).
4. Μερικά αντιγόνα (εξαιρουμένων των *HLA-C* και των *HLA* τάξης II) συνοδεύονται από το γράμμα «w» το οποίο γράφεται μπροστά από τον αριθμό (από την Αγγλική λέξη workshop) και σημαίνει ότι είναι υπό έρευνα και η ονομασία τους δεν έχει ακόμα οριστικοποιηθεί.

Σε **μοριακό επίπεδο**, τα αλληλόμορφα των γονιδίων κάθε τύπου χαρακτηρίζονται από τετραψήφιους αριθμούς, οι οποίοι χωρίζονται με έναν αστερίσκο από την ονομασία του γονιδιακού τύπου.

**ΠΙΝΑΚΑΣ Β.** Γενική αρχή ονοματολογίας HLA συστήματος.

| ΓΕΝΙΚΗ ΑΡΧΗ ΟΝΟΜΑΤΟΛΟΓΙΑΣ ΗΛΑ ΣΥΣΤΗΜΑΤΟΣ |                                 |          |
|------------------------------------------|---------------------------------|----------|
| ΓΟΝΙΔΙΑΚΟΣ ΤΟΠΟΣ                         | ΤΥΠΟΣ ΑΛΛΗΛΟΜΟΡΦΟΥ              | ΥΠΟΤΥΠΟΣ |
| (HLA) ΗΛΑ - Χ*                           | 00                              | 00       |
| ΑΛΛΗΛΟΜΟΡΦΟ                              | ΟΡΟΛΟΓΙΚΗ ΟΝΟΜΑΣΙΑ ΑΛΛΗΛΟΜΟΡΦΟΥ |          |
| ΤΑΞΗΣ Ι ΗΛΑ ΓΟΝΙΔΙΑ                      |                                 |          |
| (ΗΛΑ) ΗΛΑ - Α * 0101                     | (ΗΛΑ) ΗΛΑ - Α1                  |          |
| (ΗΛΑ) ΗΛΑ - Β * 51:CVAE                  | (ΗΛΑ) ΗΛΑ - Β8                  |          |
| ΤΑΞΗΣ ΙΙ ΗΛΑ ΓΟΝΙΔΙΑ                     |                                 |          |
| (ΗΛΑ) ΗΛΑ - DRΒ1 * 0101                  | (ΗΛΑ) ΗΛΑ - DR1                 |          |
| (ΗΛΑ) ΗΛΑ - DRΒ1 * 0102                  | (ΗΛΑ) ΗΛΑ - DR1                 |          |
| (ΗΛΑ) ΗΛΑ - DRΒ1 * 11:04:01              | (ΗΛΑ) ΗΛΑ - DR1                 |          |

**ΠΙΝΑΚΑΣ Γ.** Παραδείγματα ονοματολογίας & τυποποίησης HLA παραγόντων.

| ΠΛΗΡΗΣ ΟΝΟΜΑΣΙΑ   | ΠΕΡΙΓΡΑΦΗ                                                                                                                    | ΜΕΘΟΔΟΣ ΤΥΠΟΠΟΙΗΣΗΣ                                        |
|-------------------|------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|
| HLA-A             | Γονιδιακός τύπος A                                                                                                           | Ορολογική                                                  |
| HLA-A2            | Ορολογικό αντιγόνο A2                                                                                                        | Ορολογική                                                  |
| HLA-A23           | Διασπασμένο ορολογικό αντιγόνο A23, προερχόμενο από το ορολογικό αντιγόνο A9                                                 | Ορολογική                                                  |
| HLA-A*02:XX       | Ομάδα αλληλόμορφων (υποτύποι του αντιγόνου A2)                                                                               | Μοριακή Χαμηλής Ανάλυσης (LR – Low Resolution)             |
| HLA-A*02:01       | Αλληλόμορφο A*02:01                                                                                                          | Μοριακή Υψηλής Ανάλυσης (HR – High Resolution)             |
| HLA-A*02:01/02:02 | Ομάδα Αλληλόμορφων A*02:01 και A*02:02                                                                                       | Μοριακή Ενδιάμεσης Ανάλυσης (IR – Intermediate Resolution) |
| HLA-A*02:01/02:AB | Ομάδα Αλληλόμορφων A*02:01 και A*02:AB<br><i>Ο κωδικός «A*02:AB» είναι ο NMDP πολλαπλός κωδικός αναπαράστασης της ομάδας</i> | Μοριακή Μέσης Ανάλυσης (IR – Intermediate Resolution)      |

Σε γενικές γραμμές, η σχέση μεταξύ των αποτελεσμάτων της HLA τυποποίησης και των διαφορετικών επιπέδων της τυποποίησης είναι αρκετά περίπλοκη.

- Η τυποποίηση εκτεταμένου (broad) ορολογικού αντιγόνου αντιπροσωπεύει πάντα μια ομάδα διασπασμένων ορολογικών αντιγόνων, έτσι ώστε η σχέση τους να είναι 1:n. , π.χ. το εκτεταμένο ορολογικό αντιγόνο A9 αντιπροσωπεύει την ομάδα {A23, A24}.
- Ο LR DNA κωδικός αντιπροσωπεύει μια ομάδα HR DNA κωδικών και κάθε HR DNA κωδικός ανήκει ακριβώς σε έναν μόνο LR DNA κωδικό, π.χ. ο κωδικός A\*01:XX αντιπροσωπεύει την ομάδα {A\*01:01, A\*01:01:01, A\*01:01:02, ..., A\*01:02, A\*01:03, ..., A\*01:20, ...}.

Άλλες σχέσεις είναι πιο περίπλοκες (m:n).

- Ο IR DNA κωδικός (που ονομάζεται επίσης NMDP κωδικός ή κωδικός πολλαπλών αλληλόμορφων) αντιπροσωπεύει μια ομάδα HR DNA κωδικών, π.χ. ο κωδικός A\*01:AAXP αντιπροσωπεύει την ομάδα {A\*01:02, A\*01:08, A\*01:14}. Ωστόσο, ο HR DNA κωδικός μπορεί να ανήκει σε πολλούς IR DNA κωδικούς, π.χ. ο κωδικός A\*01:01 ανήκει στους κωδικούς A\*01:AB, A\*01:AC, A\*01:AAJ κλπ.
- Ο HR DNA κωδικός DRB1\*11:16 μπορεί να αντιστοιχεί στα ορολογικά αντιγόνα DR11 ή DR13, αλλά το αντιγόνο DR11 αντιστοιχεί σε πολλά από τα αντίστοιχα αποτελέσματα της DNA τυποποίησης (DRB1\*11:01, DRB1\*11:02, ..., DRB1\*11:16, ..., DRB1\*11:60).

**ΠΙΝΑΚΑΣ Δ.** Επίπεδα ανάλυσης HLA τυποποίησης.

| <b>ΕΠΙΠΕΔΟ ΑΝΑΛΥΣΗΣ HLA ΤΥΠΟΠΟΙΗΣΗΣ</b>     |                                 |                                   |
|---------------------------------------------|---------------------------------|-----------------------------------|
| <b>ΧΑΜΗΛΗ ΑΝΑΛΥΣΗ<br/>LR</b>                | <b>ΕΝΔΙΑΜΕΣΗ ΑΝΑΛΥΣΗ<br/>IR</b> | <b>ΥΨΗΛΗ ΑΝΑΛΥΣΗ<br/>HR</b>       |
| A*02                                        | A*02:01/02:09                   | A*02:01                           |
| B*44                                        | B*44:02/44:05                   | B*44:02                           |
| DRB1*04                                     | DRB1*04:03/04:07                | DRB1*04:03                        |
| <b>ΕΠΙΠΕΔΟ HLA ΤΥΠΟΠΟΙΗΣΗΣ</b>              |                                 |                                   |
| Επίπεδο Ορολογικής<br>Ειδικότητας (2-digit) | Επίπεδο Ομάδας<br>Αλληλόμορφων  | Επίπεδο Αλληλόμορφου<br>(4-digit) |
| <b>ΕΠΙΠΕΔΟ ΜΕΤΑΜΟΣΧΕΥΣΗΣ</b>                |                                 |                                   |
| Συμπαγή Όργανα                              |                                 | Αιμοποιητικά Κύτταρα              |

# Κ Ε Φ Α Λ Α Ι Ο 3<sup>ο</sup>

## ΣΤΟΙΧΕΙΑ ΒΙΟΣΤΑΤΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ ΗΛΑ ΔΕΔΟΜΕΝΩΝ

### 3.1 Έλεγχος γενετικής ισορροπίας Hardy-Weinberg

Η θεμελιώδης αρχή Hardy-Weinberg αναλύει πως οι γονότυποι κατανέμονται ανάλογα με τις συχνότητες των ατομικών αλληλομόρφων γονιδίων σε έναν πληθυσμό και παραμένουν σταθεροί από γενεά σε γενεά και αποτελεί ακρογωνιαίό λίθο της πληθυσμιακής γενετικής. Ο νόμος αυτός πήρε το όνομά του προς τιμήν του George Hardy, ενός Άγγλου μαθηματικού, και του Wilhelm Weinberg, ενός Γερμανού γιατρού, οι οποίοι διατύπωσαν τη συγκεκριμένη θεωρία ανεξάρτητα ο ένας από τον άλλο, το 1908. Πιο συγκεκριμένα, σύμφωνα με το νόμο Hardy-Weinberg, εάν  $p$  είναι η συχνότητα του ενός αλληλόμορφου και  $q$  η συχνότητα του άλλου, τότε οι συχνότητες των τριών συνδυασμών των αλληλομόρφων είναι  $p^2 + 2pq + q^2$ , αναλογίες που δεν μεταβάλλονται από γενεά σε γενεά. Στην γενεά που θα προέλθει από τυχαίες διασταυρώσεις σε έναν πληθυσμό στον οποίο οι γονότυποι βρίσκονται σε αναλογίες  $p^2 : 2pq : q^2$ , οι γονότυποι εμφανίζονται με τις ίδιες σχετικές αναλογίες. Εκείνη την εποχή θεωρήθηκε από ορισμένους παράξενο το ότι ένα επικρατές αλληλόμορφο δεν αυξάνει αυτομάτως τη συχνότητά του έως ότου αντικαταστήσει το αντίστοιχο υπολειπόμενο, ή το ότι ένα επιβλαβές υπολειπόμενο αλληλόμορφο, δε μειώνει αυτόματα τη συχνότητά του μέχρι να εξαλειφθεί τελείως από τον πληθυσμό. Ο νόμος Hardy-Weinberg ερμήνευσε τα αίτια της σταθερότητας των γονιδιακών συχνοτήτων. Ένας πληθυσμός που διέπεται από τα βασικά χαρακτηριστικά του νόμου, θεωρείται ότι βρίσκεται σε ισορροπία Hardy-Weinberg (HWE – Hardy-Weinberg Equilibrium). Η αρχή των Hardy-Weinberg αποτελεί μια απλουστευμένη προσέγγιση, επειδή ισχύει απόλυτα μόνο κάτω από ορισμένες ειδικές καταστάσεις που σπάνια συναντώνται στους ανθρώπινους πληθυσμούς. Ο νόμος ισχύει επακριβώς μόνο για μεγάλους πληθυσμούς, στους οποίους υπάρχει τυχαία σύζευξη και στους οποίους δεν υπάρχει επιλογή, μετάλλαξη ή μετανάστευση.

Συγκεκριμένα, οι παραδοχές που πρέπει να ισχύουν ώστε να ισχύει και η αρχή Hardy-Weinberg είναι οι εξής:

- 1) Ο πληθυσμός χαρακτηρίζεται από τυχαίο ζευγάρωμα, με μικρή ή καθόλου ομοιοτυπική σύζευξη ή ενδογαμία.
- 2) Ο υπό μελέτη γενετικός τύπος εμφανίζει σταθερό ρυθμό μετάλλαξης, και όσα μεταλλαγμένα αλληλόμορφα χάνονται λόγω θανάτου, αντικαθίστανται από νέες μεταλλάξεις.
- 3) Δεν υπάρχει επιλογή υπέρ ή κατά ενός συγκεκριμένου φαινοτύπου. Όλοι οι γονότυποι ενός γενετικού τύπου είναι εξίσου βιώσιμοι.
- 4) Ο πληθυσμός είναι αρκετά μεγάλος ώστε να μην λαμβάνει χώρα τυχαία διακύμανση των συχνοτήτων λόγω τυχαίας μεταβίβασης κάποιου γονοτύπου.
- 5) Δεν έχει αλλάξει τίποτα στη δομή του πληθυσμού λόγω μετανάστευσης, η οποία μπορεί βαθμιαία να μεταβάλλει τις γονιδιακές συχνότητες, αυξάνοντας ή μειώνοντας τον αριθμό των ατόμων με κάποιο συγκεκριμένο γονότυπο.

Απόκλιση από οποιαδήποτε από τις συνθήκες αυτές μπορεί να μεταβάλλει τη συχνότητα των αλληλομόρφων σε ένα πληθυσμό και να οδηγήσει σε αύξηση ή μείωση των συχνοτήτων των αλληλομόρφων από τη μία γενεά στην άλλη.

Η **ισορροπία Hardy-Weinberg** αποτελεί θεμελιώδη έννοια στη γενετική ενός πληθυσμού. Είναι μια μαθηματική εξίσωση που περιγράφει την κατανομή και έκφραση των αλληλομόρφων γονιδίων σε έναν πληθυσμό, και εκφράζει τις συνθήκες υπό τις οποίες οι συχνότητες του αλληλόμορφου αναμένεται να αλλάξουν. Η γενετική του Μέντελ απέδειξε ότι η φαινοτυπική (παρατηρήσιμη) έκφραση ορισμένων χαρακτηριστικών βασίζεται σε μια απλή σχέση κυρίαρχου-υπολειπόμενου μεταξύ των αλληλίων που κωδικοποιούν το χαρακτηριστικό. Στην αρχική εργασία του Μέντελ, για παράδειγμα, το πράσινο μπιζέλι ήταν κυρίαρχο σε σχέση με το κίτρινο, που σημαίνει ότι ένας ετεροζυγώτης (ένα άτομο με ένα αλληλόμορφο για το πράσινο και ένα για το κίτρινο) θα δείξει το πράσινο χαρακτηριστικό. Ένα σημαντικό ζήτημα στη γενετική ενός πληθυσμού, ως εκ τούτου, είναι ο καθορισμός της συχνότητας των κυρίαρχων και υπολειπόμενων αλληλομόρφων σε ένα πληθυσμό (για παράδειγμα, η συχνότητα του αλληλόμορφου της ομάδας αίματος O στις Ηνωμένες Πολιτείες), δεδομένης της συχνότητας των φαινοτύπων. Οι φαινοτυπικές συχνότητες και οι συχνότητες των αλληλομόρφων είναι συναφείς, αλλά δεν είναι ίσες. Οι ετεροζυγώτες εμφανίζουν το κυρίαρχο στον φαινότυπο, αλλά φέρουν ένα υπολειπόμενο

αλληλόμορφο. Ως εκ τούτου, η συχνότητα για το υπολειπόμενο αλληλόμορφο είναι υψηλότερη από την συχνότητα του υπολειπόμενου φαινοτύπου.

Στις αρχές του 20<sup>ου</sup> αιώνα ο μαθηματικός Godfrey Hardy και ο γιατρός Wilhelm Weinberg ανέπτυξαν ανεξάρτητα ένα μοντέλο που περιγράφει τη σχέση μεταξύ της συχνότητας των κυρίαρχων και των υπολειπόμενων αλληλόμορφων (εφεξής,  $p$  και  $q$ ), σε έναν πληθυσμό. Αιτιολόγησαν ότι οι συνδυασμένες συχνότητες  $p$  και  $q$  πρέπει να ισούνται προς 1, επειδή μαζί αντιπροσωπεύουν όλα τα αλληλόμορφα για το εν λόγω χαρακτηριστικό στον πληθυσμό. Οι Hardy και Weinberg παρουσίασαν ένα τυχαίο ζευγάρωμα στον πληθυσμό ως το προϊόν  $(p + q)(p + q)$ , σχέση η οποία μπορεί να επεκταθεί σε  $p^2 + 2pq + q^2$ . Αυτό αντιστοιχεί στο βιολογικό γεγονός ότι, ως αποτέλεσμα του ζευγαρώματος, ορισμένα νέα άτομα έχουν δυο αλληλόμορφα  $p$ , κάποιοι ένα  $p$  και ένα  $q$ , και άλλοι δύο αλληλόμορφα  $q$ . Ο όρος  $p^2$  παριστάνει τότε το κλάσμα του πληθυσμού που είναι ομόζυγο κυρίαρχο, ενώ οι όροι  $2pq$  και  $q^2$  αντιπροσωπεύουν τα ετερόζυγα και ομόζυγα εκφυλιστικά κλάσματα, αντίστοιχα. Μαθηματικά, δεδομένου ότι  $p + q = 1$ , ο όρος  $(p + q)^2$  πρέπει επίσης να ισούται με 1, και έτσι η χρησιμότητα αυτής της τελικής μορφής είναι ότι το  $q^2$ , το κλάσμα του πληθυσμού που είναι ομόζυγο εκφυλιστικό μπορεί να προσδιοριστεί με σχετική ευκολία, καθώς και από την εν λόγω τιμή όλων των άλλων συχνοτήτων μπορεί να υπολογιστεί. Για παράδειγμα, αν 1% του πληθυσμού βρίσκεται να είναι ομόζυγο εκφυλιστικό, δηλαδή  $q^2 = 0.01$ , τότε  $q = 0.1$ ,  $p = 0.9$ ,  $p^2 = 0.81$  και  $2pq = 0.09$ . Μία τιμή της εξίσωσης ισορροπίας Hardy-Weinberg επιτρέπει στους γενετιστές πληθυσμών να προσδιορίσουν το ποσοστό του κάθε γονότυπου και φαινοτύπου σε έναν πληθυσμό. Αυτό μπορεί να είναι χρήσιμο για τη γενετική συμβουλευτική στην περίπτωση μιας γενετικής ασθένειας, ή για τη μέτρηση της γενετικής ποικιλότητας σε ένα πληθυσμό ζώων που απειλούνται με εξαφάνιση.

## 3.2 Ανισορροπία Σύνδεσης (LD – Linkage Disequilibrium)

Στην πληθυσμιακή γενετική, η **ανισορροπία σύνδεσης** είναι η μη τυχαία σύνδεση (συσχέτιση) των αλληλομόρφων γονιδίων σε δύο ή περισσότερους γονιδιακούς τόπους που προέρχονται από μεμονωμένα, προγονικά χρωμοσώματα. Η ανισορροπία σύνδεσης αποτελεί απαραίτητο εργαλείο για την υλοποίηση μελετών γενετικής συσχέτισης. Η παράμετρος αυτή αναφέρεται και ως ανισορροπία γενετικής φάσης ή απλά γενετική ανισορροπία, διότι περιγράφεται μέσω του ανασυνδυασμού του DNA. Με άλλα λόγια, η ανισορροπία σύνδεσης είναι η εμφάνιση μερικών συνδυασμών αλληλομόρφων ή γενετικών δεικτών σε έναν πληθυσμό πιο συχνά ή λιγότερο συχνά απ' ό,τι θα αναμενόταν από μια τυχαία διαμόρφωση των απλοτύπων από τα αλληλόμορφα, με βάση τις συχνότητές τους. Είναι ένα δεύτερο κατά σειρά φαινόμενο που προέρχεται από τη γενετική σύνδεση, η οποία απαιτεί την παρουσία δύο ή περισσότερων γονιδιακών τόπων σε ένα χρωμόσωμα με περιορισμένο ανασυνδυασμό μεταξύ τους. Το ποσό της ανισορροπίας σύνδεσης εξαρτάται από τη διαφορά μεταξύ των παρατηρούμενων (observed) συχνοτήτων των αλληλόμορφων και εκείνων των αντίστοιχων συχνοτήτων που θεωρητικά αναμένονται να προκύψουν (expected) από ένα ομοιογενές, τυχαίως κατανομημένο μοντέλο. Οι πληθυσμοί όπου συνδυασμοί αλληλόμορφων ή γονοτύπων μπορεί να βρεθούν στις αναμενόμενες αναλογίες, μπορεί να θεωρηθεί ότι βρίσκονται σε ισορροπία σύνδεσης. Το επίπεδο της ανισορροπίας σύνδεσης επηρεάζεται από ένα σύνολο παραγόντων, συμπεριλαμβανομένης της γενετικής σύνδεσης, της φυσικής επιλογής, το ποσοστό του ανασυνδυασμού, το ποσοστό της μετάλλαξης, της γενετικής παρέκκλισης, του μη-τυχαίου ζευγαρώματος, και της δομής του πληθυσμού.

### 3.2.1 Ποσοτικοποίηση Ανισορροπίας Σύνδεσης (LD)

Η απόκλιση της τυχαίας σύνδεσης των αλληλόμορφων  $A_i$  και  $B_j$  σε δύο διαφορετικούς γονιδιακούς τόπους (two loci), υπολογίζεται κατά βάση χρησιμοποιώντας τον συντελεστή LD ως μέτρο ανισορροπίας σύνδεσης.

$$D_{A_i B_j} = f_{A_i B_j} - p_{A_i} p_{B_j} \Rightarrow D_{ij} = f_{ij} - p_i q_j \quad (1)$$

Ο Lewontin (1964) ανέπτυξε μια κανονικοποιημένη μετρική

$$D'_{ij} = \frac{D_{ij}}{D_{MAX}} \quad (2)$$

ώστε να δηλώσει το γεγονός ότι το εύρος της τιμής  $D_{ij}$  δεν είναι ανεξάρτητο από τις συχνότητες των αλληλόμορφων. Οι τιμές της μετρικής  $D'_{ij}$  έχουν εύρος από -1 έως +1.

Υπάρχουν διάφορες στατιστικές μέθοδοι για τον υπολογισμό της συνολικής ανισορροπίας σύνδεσης LD μεταξύ συνόλων αλληλόμορφων σε δυο διαφορετικούς γονιδιακούς τόπους. Η στατιστική παράμετρος  $D'$  που ανέπτυξε ο Hedrick (1987) αθροίζει την συνεισφορά όλων των απλοτύπων  $D'_{ij}$  των ατόμων ενός δείγματος πληθυσμού, σε ένα πολύ-αλληλομορφικό σύστημα δυο γονιδιακών τόπων, χρησιμοποιώντας σαν βάρη τις συχνότητες των αλληλόμορφων  $p_i$  και  $q_j$ , με τις παραμέτρους  $I$  και  $J$  να αναπαριστούν το πλήθος των αλληλόμορφων στους γονιδιακούς τόπους  $A$  και  $B$  αντίστοιχα.

$$D' = \sum_{i=1}^I \sum_{j=1}^J p_i q_j |D'_{ij}| \quad (3)$$

Η μετρική  $W_n$  (Cramer's V Statistic - Cramer 1946, Cohen 1988), αποτελεί μια δεύτερη στατιστική παράμετρο υπολογισμού του μέτρου της ανισορροπίας σύνδεσης LD μεταξύ δυο γονιδιακών τόπων. Η συγκεκριμένη μετρική βασίζεται με έναν σταθμισμένο μέσο όρο των τιμών ανισορροπίας σύνδεσης  $D_{ij}$ . Η μετρική  $W_n$  μπορεί να γραφεί ως επανέκφραση του συνόλου των Chi-square στατιστικών παραμέτρων  $X_{LD}^2$ , κανονικοποιημένων κατά τρόπο ώστε η τιμή της να έχει εύρος από 0 έως 1.

$$W_n = W_{AB} = \left[ \frac{\sum_{i=1}^I \sum_{j=1}^J \frac{D_{ij}^2}{p_i q_j}}{\min(I-1, J-1)} \right]^{1/2} = \left[ \frac{\frac{X_{LD}^2}{2N}}{\min(I-1, J-1)} \right]^{1/2} \quad (4)$$

Όταν υπάρχουν δυο αλληλόμορφα ανά γονιδιακό τόπο, η μετρική  $W_n$  ισούται με την τιμή της συσχέτισης ανάμεσα στους δυο γονιδιακούς τόπους, η οποία ορίζεται ως εξής :

$$r^2 = \frac{D_{11}}{p_1 p_2 q_1 q_2} \quad (5)$$

### 3.3 Γενετική Απόσταση (Genetic Distance)

Προκειμένου να περιγραφεί ποσοτικά η διαδικασία παραγωγής γενετικών διαφορών μεταξύ δύο πληθυσμών λόγω της φυσικής επιλογής, της μετάλλαξης, της μετανάστευσης και της τυχαίας μετατόπισης, εισήχθη η έννοια της γενετικής απόστασης. Η γενετική απόσταση αποτελεί ένα μέτρο της διαφοράς των γονιδιακών συχνοτήτων μεταξύ διαφορετικών πληθυσμών και χρησιμοποιείται για να περιγράψει την αντίστοιχη ενδοπληθυσμιακή διακύμανση.

Το 1977, οι Cavalli-Sforza και Bodmer όρισαν τη γενετική απόσταση ( $d$ ) ως εξής:

$$d = \sqrt{1 - \sum_i \sqrt{p_{i1}p_{i2}}} \quad (6)$$

όπου  $p_{i1}$  και  $p_{i2}$  είναι οι συχνότητες του γονιδίου  $i$  στους πληθυσμούς 1 και 2 αντίστοιχα.

### 3.4 Εκτίμηση Συχνοτήτων HLA Απλοτύπων

Κατά την προσπάθεια της εκτίμησης της συχνότητας των HLA απλοτύπων εντός ενός πληθυσμού εφαρμόζονται διάφορες μεθοδολογίες, μεταξύ των οποίων ιδιαίτερη έμφαση δίνεται στην **συνάρτηση μέγιστης πιθανότητας** και της λύσης που προκύπτει μέσω της εκτέλεσης του επαναληπτικού αλγορίθμου μεγιστοποίησης προσδοκίας (EM – Expectation-Maximization).

Το **πλήθος των γονοτύπων** ( $c_j$ ) που συνθέτουν τον  $j$ -οστό φαινότυπο είναι μια συνάρτηση του πλήθους ( $s_j$ ) των ετερόζυγων των ετερόζυγων γονιδιακών τόπων (loci).

$$c_j = \begin{cases} 2^{s_j-1}, & \alpha \nu \ s_j > 0 \\ 1, & \alpha \nu \ s_j = 0 \end{cases} \quad (7)$$

### **Μοντελοποίηση του Προβλήματος**

Οι τεχνικές HLA τυποποίησης επιτρέπουν την έρευνα σε πολλούς πολυμορφικούς γονιδιακούς τύπους, χωρίς ωστόσο να διακρίνονται οι γαμετικές φάσεις των απλοτύπων. Για ετερόζυγους διπλοειδείς οργανισμούς η άμεση PCR αλληλούχηση παράγει αποτελέσματα ως προς την ενίσχυση αλληλόμορφων αλλά δεν τυποποιεί τους απλοτύπους όταν το διπλοειδές άτομο είναι ετερόζυγο σε περισσότερους από έναν γονιδιακούς τύπους. Έτσι, το σύνολο των δεδομένων αποτελείται από αποτελέσματα HLA τυποποίησης με άγνωστη γαμετική φάση απλοτύπων σε έναν ή περισσότερους γονιδιακούς τύπους. Ο στόχος είναι να βρεθεί η καλύτερη εκτίμηση των συχνοτήτων των HLA απλοτύπων εντός του πληθυσμού, δεδομένων ορισμένων πληροφοριών που περιλαμβάνονται στα δεδομένα των φαινοτύπων (unphased γονοτύπων) του δείγματος.

### **Προσέγγιση Μέγιστης Πιθανότητας (Maximum Likelihood)**

Βάσει της υπόθεσης της ισορροπίας Hardy-Weinberg, η πιθανότητα  $P_j$  του  $j$ -οστού φαινοτύπου δίνεται από το άθροισμα των πιθανοτήτων κάθε πιθανού  $c_j$  γονοτύπου.

$$P_j = \sum_{i=1}^{c_j} P(\text{γονότυπος } i) = \sum_{i=1}^{c_j} P(h_k h_l) \quad (8)$$

όπου  $P(h_k h_l)$  η πιθανότητα ο  $i$ -οστός γονότυπος να προκύπτει από τους απλοτύπους  $k$  και  $l$ .

$$P(h_k h_l) = \begin{cases} p_k^2, & \alpha \nu \quad k = l \\ 2p_k p_l, & \alpha \nu \quad k \neq l \end{cases} \quad (9)$$

και το  $p_i$  υποδηλώνει την συχνότητα του  $i$ -οστού απλοτύπου  $h_i$  εντός του πληθυσμού.

Η πιθανότητα ένα δείγμα πληθυσμού, αποτελούμενο από  $n$  άτομα να προσδιορίζεται από συχνότητες φαινοτύπων  $P_1, P_2, \dots, P_m$ , δίνεται από την εξής πολυωνυμική έκφραση:

$$P(\text{δείγμα} \mid P_1, P_2, \dots, P_m) = \frac{n!}{n_1! n_2! \dots n_m!} \times P_1^{n_1} \times P_2^{n_2} \times \dots \times P_m^{n_m} \quad (10)$$

όπου το  $m$  υποδηλώνει το συνολικό πλήθος των φαινοτύπων και το  $n_j$  αποτελεί το πλήθος των ατόμων που φέρουν τον  $j$ -οστό φαινότυπο που παρατηρείται στο δείγμα.

$$\sum_{j=1}^m n_j = n \quad (11)$$

Βάσει των εξισώσεων (8) και (10) , προκύπτει η πιθανότητα του δείγματος ως μια συνάρτηση των άγνωστων συχνοτήτων των απλοτύπων.

Συνεπώς, οι πιθανότητες των συχνοτήτων των απλοτύπων είναι :

$$L(p_1, p_2, \dots, p_h) = \frac{n!}{n_1! n_2! \dots n_m!} \times \prod_{j=1}^m \left( \sum_{i=1}^{c_j} P(h_{ik} h_{il}) \right)^{n_j} \quad (12)$$

Ισχύει :

$$\sum_{i=1}^h p_i = 1 \quad (13)$$

### 3.5 Ο αλγόριθμος μεγιστοποίησης προσδοκίας (Αλγόριθμος EM – Expectation-Maximization)

Ο αλγόριθμος EM βασίζεται στην προσπάθεια επίτευξης της μεγιστοποίησης της προσδοκίας και αποτελεί μια μέθοδο επαναληπτικής εκτίμησης συχνοτήτων των απλοτύπων ενός πληθυσμού, λαμβάνοντας υπόψη και την επεξεργασία ελλιπών στοιχείων του δείγματος. Ο όρος EM καθιερώθηκε από τους Dempster, Laird και Rubin στην δημοσιευμένη εργασία τους που σχετιζόταν με αυτή την μέθοδο το έτος 1977. Σύμφωνα με τους Little και Rubin (1987), ο αλγόριθμος EM τυποποιεί μια σχετικά παλιά «ad hoc» ιδέα για το χειρισμό των ελλιπών στοιχείων ως εξής :

- 1) Αντικαθιστά τις ελλιπείς τιμές με τις κατ' εκτίμηση τιμές.
- 2) Εκτιμά τις παραμέτρους.
- 3) Επανεκτιμά τις ελλιπείς τιμές υποθέτοντας ότι οι νέες εκτιμήσεις των παραμέτρων είναι σωστές.
- 4) Επανεκτιμά τις παραμέτρους και ούτω καθ' εξής, επαναλαμβάνοντας την προαναφερθείσα διαδικασία έως ότου να προκύψει σύγκλιση.

Κάθε επανάληψη του αλγορίθμου EM αποτελείται από δύο βήματα :

- 1) Βήμα E (Expectation – Προσδοκία) : η αναμενόμενη τιμή του λογαρίθμου της πιθανοφάνειας του πλήρους συνόλου δεδομένων προκύπτει λαμβάνοντας υπόψη τα παρατηρηθέντα στοιχεία και τις κατ' εκτίμηση παραμέτρους από μια προηγούμενη επανάληψη.
- 2) Βήμα M (Maximization – Μεγιστοποίηση) : η δεσμευμένη αναμενόμενη τιμή του λογαρίθμου της πιθανοφάνειας του πλήρους συνόλου δεδομένων μεγιστοποιείται. Η τιμή αυτή αυξάνεται έως ότου επιτευχθεί ένα στάσιμο σημείο (Dempster et Al, 1977). Με άλλα λόγια, ο αλγόριθμος συνεχίζει να εκτελείται επαναληπτικά έως ότου η παρατηρηθείσα πιθανοφάνεια που παράγεται σε δύο διαδοχικές επαναλήψεις να είναι σχεδόν ίδια.

### 3.5.1 Περιγραφή Αλγορίθμου EM

Ο αλγόριθμος EM αποτελεί μια διαδραστική μέθοδο υπολογισμού συνόλων συχνοτήτων απλοτύπων  $p_1, p_2, \dots, p_h$ , ξεκινώντας από αντίστοιχες αυθαίρετες αρχικές τιμές  $p_1^{(0)}, p_2^{(0)}, \dots, p_h^{(0)}$ . Αυτές οι αρχικές τιμές χρησιμοποιούνται για την εκτίμηση των συχνοτήτων των γονοτύπων  $\tilde{P}(h_k h_l)$  σαν αυτές να ήταν άγνωστες πραγματικές συχνότητες (ΒΗΜΑ 1<sup>ο</sup> : Προσδοκία). Αυτές οι αναμενόμενες (expected) συχνότητες γονοτύπων προτυποποιούνται και στη συνέχεια χρησιμοποιούνται για την εκτίμηση των συχνοτήτων των απλοτύπων  $\hat{p}$  στην επόμενη επανάληψη της εκτέλεσης του αλγορίθμου EM (ΒΗΜΑ 2<sup>ο</sup> : Μεγιστοποίηση). Η παραπάνω διαδικασία συνεχίζει να υλοποιείται κατ' επανάληψη έως ότου επιτευχθεί η επιθυμητή σύγκλιση.

#### Αρχικές Συνθήκες

Υπάρχουν διάφορες δυνατότητες αρχικοποίησης των τιμών των συχνοτήτων των απλοτύπων  $p_1^{(0)}, p_2^{(0)}, \dots, p_h^{(0)}$  με την προϋπόθεση της ισχύος της σχέσης  $\sum_{i=1}^h p_i = 1$ .

Η αρχικοποίηση υλοποιείται με βάση την παρακάτω μοντελοποίηση :

- 1) Οι αρχικές τιμές των συχνοτήτων όλων των απλοτύπων δίνονται ως εξής :

$$p_t^{(0)} = \frac{1}{n_h}, \quad t = 1, 2, \dots, n_h \quad (14)$$

- 2) Οι αρχικές τιμές των συχνοτήτων όλων των πιθανών γονοτύπων κάθε φαινοτύπου δίνονται ως εξής :  $P_j(h_k h_l)^{(0)} = \frac{1}{c_j}, j=1,2,...,m$  (15)
- 3) Οι αρχικές τιμές των συχνοτήτων των απλοτύπων επιλέγονται τυχαία.
- 4) Οι αρχικές τιμές όλων των συχνοτήτων των απλοτύπων είναι ίσες τις αντίστοιχες τιμές των συχνοτήτων των αλληλόμορφων του σχετικού γονιδιακού τόπου (πλήρης ισορροπία γενετικής σύνδεσης).
- 5) Τα αρχικά δεδομένα επηρεάζουν τις αρχικές τιμές των συχνοτήτων των απλοτύπων.

### **ΒΗΜΑ 1<sup>ο</sup> : Προσδοκία (Expectation)**

Η εκτίμηση των συχνοτήτων των γονοτύπων, δεδομένων των συχνοτήτων των απλοτύπων δίνεται ως εξής :

$$\tilde{P}(h_k h_l)^{(g)} = \begin{cases} (p_k^{(g)})^2, & \alpha\nu \ k=l \\ 2p_k^{(g)} p_l^{(g)}, & \alpha\nu \ k \neq l \end{cases} \quad (16)$$

### **ΒΗΜΑ 2<sup>ο</sup> : Μεγιστοποίηση (Maximization)**

Η εκτίμηση των συχνοτήτων των φαινοτύπων, δεδομένων των συχνοτήτων των γονοτύπων δίνεται ως εξής :

$$P_j^{(g)} = \sum_{i=1}^{c_j} \tilde{P}(\gamma\omicron\nu\acute{o}\tau\upsilon\pi\omicron\varsigma \ i)^{(g)} = \sum_{i=1}^{c_j} \tilde{P}(h_k h_l)^{(g)} \quad (17)$$

### **Προτυποποίηση των συχνοτήτων των γονοτύπων**

$$P(h_k h_l)^{(g)} = \frac{n_j}{n} \frac{\tilde{P}(h_k h_l)^{(g)}}{P_j^{(g)}} \quad (18)$$

Ένας γονότυπος έχει έναν ή δύο ιδιαίτερους απλοτύπους, ώστε να υπάρχει η δυνατότητα οι συχνότητες των γονοτύπων να μπορούν να χρησιμοποιηθούν για την εκτίμηση των συχνοτήτων των απλοτύπων μέσω απευθείας υπολογισμού όλων των παρατηρούμενων απλοτύπων εντός ενός δείγματος γονοτύπων.

$$\hat{p}_t^{(g+1)} = \frac{1}{2} \sum_{j=1}^m \sum_{i=1}^{c_j} \delta_{it} P_j(h_k h_l)^{(g)} \quad (19)$$

Όπου  $\delta_{it}$  είναι μια μεταβλητή-δείκτης που ισούται με το πλήθος των φορών που ο απλότυπος  $t$  συμμετέχει στον γονότυπο  $i$ .

$$\delta_{it} = \begin{cases} 0, & \alpha \vee (t \neq k) \wedge (t \neq l) \\ 1, & \alpha \vee ((t \neq k) \wedge (t = l)) \vee (((t = k) \wedge (t \neq l))) \\ 2, & \alpha \vee t = k = l \end{cases} \quad (20)$$

### **Κριτήριο Τερματισμού**

Το κριτήριο τερματισμού (σύγκλισης) του αλγορίθμου EM μπορεί να καθοριστεί ως εξής :

- 1) Η σχετική διαφορά μεταξύ των διαδοχικών τιμών της συνάρτησης ML (Maximum Likelihood), ως αποτέλεσμα της επαναληπτικής εκτέλεσης του αλγορίθμου είναι μικρότερη από ένα προκαθορισμένο κατώφλι  $\varepsilon > 0$ .
- 2) Η απόλυτη τιμή της διαφοράς μεταξύ των διαδοχικών τιμών της συνάρτησης ML (Maximum Likelihood), ως αποτέλεσμα της επαναληπτικής εκτέλεσης του αλγορίθμου είναι μικρότερη από ένα προκαθορισμένο κατώφλι  $\varepsilon > 0$ .
- 3) Όταν οι αλλαγές στις διαδοχικές τιμές των συχνοτήτων των απλοτύπων κατά την επαναληπτική εκτέλεση του αλγορίθμου είναι μικρότερες από ένα προκαθορισμένο κατώφλι  $\varepsilon > 0$

$$\left| \hat{p}_t^{(g+1)} - \hat{p}_t^{(g)} \right| \leq \varepsilon \quad t = 1, 2, \dots, h \quad (21)$$

### **3.5.2 Ιδιότητες Αλγορίθμου EM**

#### **1. Το μέγεθος του δείγματος**

Όπως είναι αναμενόμενο, ο αλγόριθμος εκτελείται αποδοτικότερα για μεγαλύτερα μεγέθη δειγμάτων, διότι όσο μεγαλύτερο πλήθος δεδομένων διαθέτει, τόσο εγκυρότερες είναι οι εκτιμήσεις στις οποίες καταλήγει.

#### **2. Πολλαπλά τοπικά μέγιστα**

Ο αλγόριθμος EM προσπαθεί να φτάσει σε ένα ολικό μέγιστο της συνάρτησης πιθανότητας, ωστόσο δεν υπάρχει καμία εγγύηση ότι η επιφάνεια της πιθανότητας την οποία διατρέχει είναι κυρτή, δηλαδή δεν υπάρχει απόδειξη για την μοναδικότητα της μέγιστης πιθανότητας, έτσι ώστε η επιφάνεια πιθανότητας μπορεί να έχει πολλά τοπικά μέγιστα. Για να διασφαλιστεί η εξεύρεση ολικής μέγιστης πιθανότητας, ο αλγόριθμος EM θα πρέπει να ξεκινήσει από διάφορες αρχικές συνθήκες.

### **3. Απόκλιση από την ισορροπία Hardy-Weinberg (HWE)**

Η απομάκρυνση από την ισορροπία HWE μπορεί να είναι μια σημαντική πηγή σφαλμάτων, επειδή ο αλγόριθμος βασίζεται σε τιμές ισορροπίας HWE στο βήμα της προσδοκίας. Ωστόσο, ενδεχόμενη απόκλιση από την τιμή HWE δεν θα οδηγήσει σε σημαντική διαφοροποίηση στην εκτίμηση της συχνότητας των απλοτύπων. Επίσης η τιμή της ανισορροπίας σύνδεσης (LD) δεν επηρεάζει σε μεγάλο βαθμό τις τιμές συχνοτήτων των κοινών απλοτύπων.

### **4. Ταχύτητα σύγκλισης**

Οι περισσότερες μελέτες επιβεβαιώνουν την αποδοτικότητα της υψηλής ταχύτητας σύγκλισης του αλγορίθμου EM, π.χ. επίτευξη σύγκλισης σε λιγότερες από 20 επαναλήψεις ή σε λιγότερες από 50 επαναλήψεις.

Άλλοι παράγοντες που θα μπορούσαν να επηρεάσουν την εκτέλεση του αλγορίθμου EM είναι: το σχήμα του γραφήματος λογαριθμικής πιθανότητας, η ευαισθησία στα κριτήρια τερματισμού, η ανισορροπία σύνδεσης (LD) και η ευαισθησία σε διαφορετικές αρχικές συνθήκες.

## **Κ Ε Φ Α Λ Α Ι Ο 4<sup>ο</sup>**

### **ΔΙΑΧΕΙΡΙΣΗ ΜΗΤΡΩΩΝ ΕΘΕΛΟΝΤΩΝ ΔΟΤΩΝ ΜΥΕΛΟΥ ΤΩΝ ΟΣΤΩΝ & ΕΡΓΑΛΕΙΑ ΑΝΑΛΥΣΗΣ HLA ΔΕΔΟΜΕΝΩΝ**

#### **4.1 Μητρώα Εθελοντών Δοτών Μυελού των Οστών**

Οι βάσεις δεδομένων των μητρώων εθελοντών δοτών αιμοποιητικών βλαστικών κυττάρων αποτελούν μοναδικές και πολύτιμες πηγές για την υλοποίηση γενετικών πληθυσμιακών μελετών. Τα περισσότερα από τα αποτελέσματα της HLA τυποποίησης λαμβάνονται σε διαπιστευμένα εργαστήρια HLA τυποποίησης με πρότυπα ελέγχου υψηλής ποιότητας. Τα στοιχεία που εισάγονται στα μητρώα δεν παραμένουν ανεκμετάλλευτα, αλλά αναλύονται καθημερινά και ενημερώνονται συνεχώς προκειμένου να βρεθούν οι καταλληλότεροι συμβατοί μη συγγενείς δότες για μεταμόσχευση αιμοποιητικών βλαστικών κυττάρων. Από την άλλη πλευρά, η εκτίμηση των συχνοτήτων των HLA απλοτύπων των ατόμων που είναι εγγεγραμμένοι στα μητρώα είναι μια απαιτητική διαδικασία για τους ακόλουθους λόγους :

- ❖ Ελλιπή δεδομένα.
- ❖ Τα δεδομένα των μητρώων περιλαμβάνουν αποτελέσματα της HLA ανάλυσης που έχει υλοποιηθεί με διάφορες τεχνικές HLA τυποποίησης και κατά συνέπεια έχουν διαφορετικά επίπεδα ανάλυσης.
- ❖ Το HLA σύστημα είναι εξαιρετικά πολυμορφικό και από την HLA τυποποίηση προκύπτουν πολλά νέα αλληλόμορφα γονίδια.
- ❖ Υπάρχουν πάρα πολλοί HLA γονιδιακοί τόποι για τους οποίους θα ήταν χρήσιμο να εκτιμηθούν οι συχνότητες των αντίστοιχων HLA απλοτύπων. Τέτοιοι γονιδιακοί τόποι είναι οι εξής: A, B, C, DRB1, DRB3, DRB4, DRB5, DQA1, DQB1, DPA1 και DPB1. Αντιθέτως οι γονιδιακοί τόποι DRB3, DRB4, DRB5, DQA1, DPA1 και DPB1 είναι σπάνιοι και ανεπαρκείς για τον υπολογισμό συχνοτήτων HLA απλοτύπων, και ως εκ τούτου, για πρακτικούς λόγους, θα εξετάσουμε μόνο τα γονίδια A, B και DRB1.

#### **4.1.1 Διεθνή Μητρώα Εθελοντών Δοτών Μυελού των Οστών**

##### **Ιστορικά Στοιχεία**

Η αρχική ιδέα της δημιουργίας των δεξαμενών (μητρώων) εθελοντών δοτών μυελού των οστών ανατρέχει ήδη στο 1970, όταν ο Dr. Jon J. van Rood πρότεινε σε μία συνάντηση της Γερμανικής Εταιρείας Μεταγγίσεων στο Μόναχο, να δημιουργηθεί ένα αρχείο, στο οποίο θα καταγράφονταν οι HLA τυποποιημένοι εθελοντές αιμοδότες, ώστε να χρησιμοποιηθούν σε μεταγγίσεις HLA συμβατών αιμοπεταλίων και σε μεταμοσχεύσεις μυελού των οστών. Έτσι δημιουργήθηκε το «**Διεθνές Αρχείο Εθελοντών Δοτών Μυελού των Οστών**», (**BMDW - Bone Marrow Donors Worldwide**), μια βάση δεδομένων από όπου γίνεται η αναζήτηση εθελοντών δοτών και μονάδων ομφαλοπλακουντιακού αίματος. Το 1989 η πρώτη έντυπη έκδοση αυτού του «χρυσού οδηγού» δοτών και μονάδων ομφαλοπλακουντιακού αίματος περιελάμβανε 150.000 δότες από 8 χώρες. Σε κάθε κράτος, οι καταγραφόμενοι εθελοντές συγκεντρώνονται σε ένα κεντρικό αρχείο που αποτελεί την Εθνική δεξαμενή της χώρας. Σήμερα, υπάρχουν 58 δεξαμενές δοτών σε 43 χώρες και 39 τράπεζες φύλαξης μονάδων ομφαλοπλακουντιακού αίματος σε 21 χώρες. Τα στοιχεία δοτών και μονάδων συγκεντρώνονται σε ένα Διεθνές Αρχείο, το οποίο είχε, έως τον Ιανουάριο του 2011, καταγεγραμμένους περισσότερους από 14.000.000 εθελοντές (συμπεριλαμβανομένων 450.000 περίπου μονάδων ομφαλοπλακουντιακού αίματος). Όλες οι συνεργαζόμενες εθνικές δεξαμενές έχουν πρόσβαση στο Διεθνές Αρχείο. Η πρώτη αμιγής δεξαμενή για Εθελοντές Δότες Μυελού των Οστών δημιουργήθηκε στη Μεγάλη Βρετανία το 1974, η «**The Anthony Nolan Bone Marrow Trust**». Η προσπάθεια ξεκίνησε προκειμένου να βρεθεί δότης για τον τετράχρονο Anthony Nolan, ο οποίος έπασχε από το Σύνδρομο Wiskott Aldrich. Δυστυχώς ο μικρός Anthony πέθανε το 1979, χωρίς να μπορέσει να μεταμοσχευτεί, όμως άλλοι ασθενείς επωφελήθηκαν από το κληροδότημά του. Κάτω από την εμπνευσμένη καθοδήγηση της μητέρας του Anthony, Shirley Nolan, η καταγραφή δοτών συνεχίστηκε με τα χρόνια και ήταν η πρόθεσή της να δώσει την ελπίδα ενός δότη σε κάθε ασθενή που χρειαζόταν μεταμόσχευση σε κάθε σημείο του κόσμου. Η καταγραφή των εθελοντών δοτών μυελού των οστών, που εμπνεύστηκε η Shirley Nolan, λειτούργησε ως μοντέλο για την ίδρυση πολυάριθμων δεξαμενών σχεδόν σε κάθε αναπτυγμένη χώρα του κόσμου. Σήμερα, το Διεθνές Αρχείο Εθελοντών Δοτών Μυελού των Οστών, αποτελεί μια διαδικτυακή μηχανή αναζήτησης και βασίζεται στη συνεργασία χωρών από όλο τον κόσμο. Παρά το γεγονός ότι υπάρχουν πλέον μεγάλες πιθανότητες

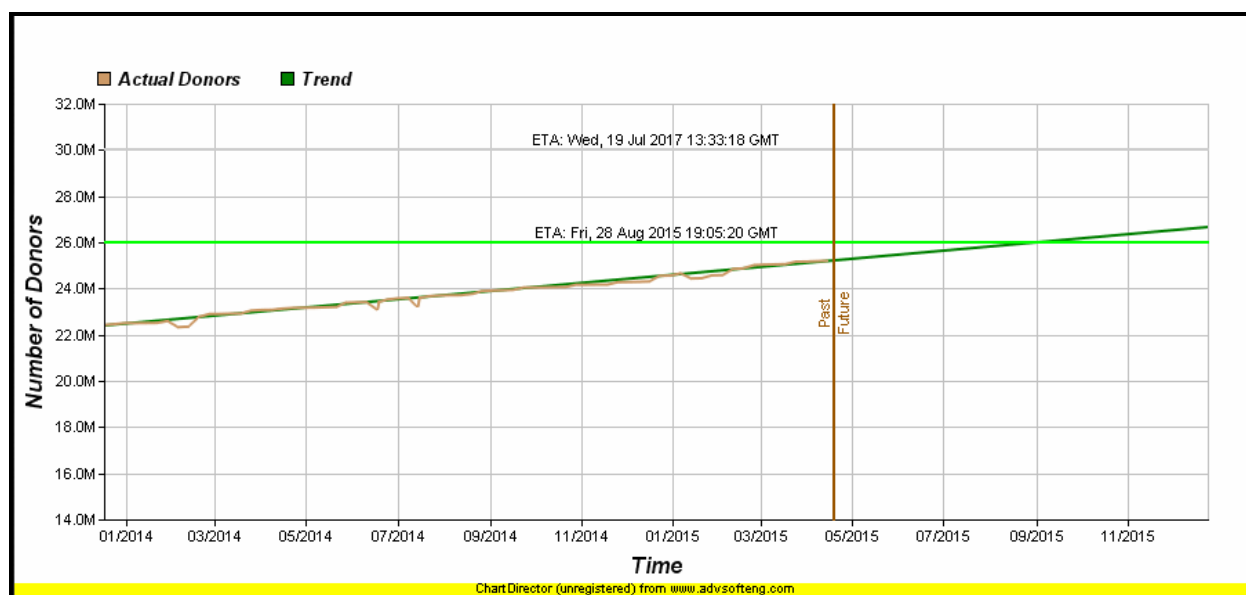
ανεύρεσης δότη, υπάρχουν δυστυχώς ακόμα ασθενείς που δεν είναι δυνατόν να βρουν ικανοποιητικά συμβατό δότη.

#### 4.1.1.1 Διεθνές Αρχείο Εθελοντών Δοτών Μυελού των Οστών (BMDW)

Το **BMDW (Bone Marrow Donors Worldwide)** ξεκίνησε τη λειτουργία του το 1988, με τα γραφεία του να βρίσκονται στο Leiden (Ολλανδία), ως μια πρωτοβουλία του Immunobiology Working Party του European Group of Blood and Marrow Transplantation (EBMT). Η πρώτη έκδοση του προγράμματος ήταν διαθέσιμη τον Φεβρουάριο του 1989 και περιείχε δεδομένα από 8 αρχεία και ένα σύνολο 155.000 δωρεών αιμοποιητικών βλαστικών κυττάρων από εθελοντές δότες. Το πρόγραμμα BMDW αποτελεί μια εθελοντική προσπάθεια που βασίζεται στη συνεργασία μητρώων δοτών αιμοποιητικών βλαστικών κυττάρων αλλά και τραπεζών ομφαλοπλακουντιακού αίματος.

Οι στόχοι της προσπάθειας είναι :

- Η παροχή κεντρικής πληροφόρησης όσον αφορά στους HLA φαινοτύπους, αλλά και σχετικά με τα απαραίτητα δεδομένα των δωρεών αιμοποιητικών βλαστικών κυττάρων και μονάδων ομφαλοπλακουντιακού αίματος από μη συγγενείς δότες.
- Η διευκόλυνση της πρόσβασης των ιατρών των ασθενών που έχουν την ένδειξη μεταμόσχευσης αιμοποιητικών βλαστικών κυττάρων στην πληροφόρηση αυτή.



**EIKONA 4.1** Η εξέλιξη του πλήθους εθελοντών δοτών μυελού των οστών στο χρόνο σε παγκόσμιο επίπεδο, σύμφωνα με το αρχείο BMWD (Πηγή : [www.bmdw.org](http://www.bmdw.org))

#### ***4.1.1.2 Παγκόσμια Ένωση Δοτών Μυελού των Οστών (WMDA)***

Η WMDA (**World Marrow Donor Association**) είναι μια εθελοντική οργάνωση εκπροσώπων των μητρώων εθελοντών δοτών βλαστικών κυττάρων του αίματος, των τραπεζών ομφαλοπλακουντιακού αίματος, και άλλων οργάνωσεων και ατόμων που ενδιαφέρονται για μεταμόσχευση βλαστικών κυττάρων του αίματος. Παρέχει ένα φόρουμ για τη συζήτηση των θεμάτων που αφορούν την κλινική χρήση των βλαστικών κυττάρων του αίματος από μη συγγενείς δότες διεθνών συνόρων και για τη διαμόρφωση των κατευθυντήριων γραμμών σχετικά με την εφοδιαστική, τον ποιοτικό έλεγχο, την ηθική, τα οικονομικά, την πληροφορική και τη διαπίστευση των μητρώων. Η WMDA ξεκίνησε ως ένα πρόγραμμα το 1988 από τρεις πρωτοπόρους στον τομέα των μεταμοσχεύσεων: τον John Goldman (Ηνωμένο Βασίλειο), τον E. Donnell Thomas (ΗΠΑ) και τον Jon J. van Rood (Ολλανδία). Αυτό το πρόγραμμα οδήγησε στην ίδρυση της ένωσης WMDA το 1994. Από τότε, η WMDA έχει εξελιχθεί σε μια παγκόσμια ένωση, που εκπροσωπεί 73 μητρώα δοτών, 158 τράπεζες αίματος ομφάλιου λώρου, 350 κέντρα δοτών και 1.615 νοσοκομεία μεταμόσχευσης από 52 διαφορετικές χώρες.

#### ***4.1.1.3 Εθνικό Πρόγραμμα Δοτών Μυελού των Οστών (NMDP)***

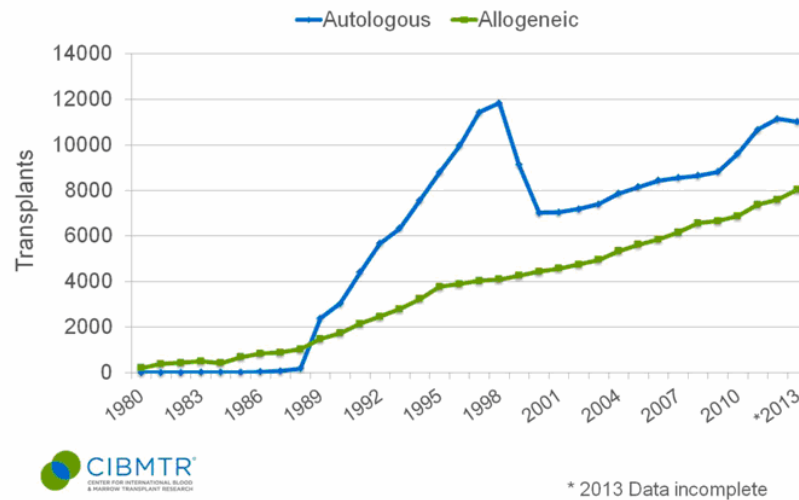
Το NMDP (**National Marrow Donor Program**) είναι ένας μη κερδοσκοπικός οργανισμός που ιδρύθηκε το 1986 και εδρεύει στη Μινεάπολη της Μινεσότα των Η.Π.Α., ο οποίος υλοποιεί το πρόγραμμα «Be The Match Registry» σε εθελοντές δότες αιμοποιητικών βλαστικών κυττάρων και μονάδων ομφαλοπλακουντιακού αίματος. Το «Be The Match Registry» είναι το μεγαλύτερο μητρώο αιμοποιητικών βλαστικών κυττάρων στον κόσμο, απαριθμώντας πάνω από 10,5 εκατομμύρια άτομα και περίπου 185.000 μονάδες αίματος ομφάλιου λώρου. Από τον Ιανουάριο του 2013, το NMDP είχε διευκολύνει περισσότερες από 55.000 μεταμοσχεύσεις σε όλο τον κόσμο.

#### ***4.1.1.4 Κέντρο Έρευνας Διεθνών Μεταμοσχεύσεων Αίματος & Μυελού (CIBMTR)***

Το CIBMTR (**Center for International Blood & Marrow Transplant Research**) , συνεργάζεται με την παγκόσμια επιστημονική κοινότητα με σκοπό να προωθήσει την έρευνα σχετικά με τη μεταμόσχευση αιμοποιητικών κυττάρων και την κυτταρική θεραπεία σε όλο τον κόσμο. Αποτελεί ένα συνδυαστικό ερευνητικό πρόγραμμα του NMDP και του Ιατρικού Κολεγίου του Wisconsin των ΗΠΑ. Το CIBMTR

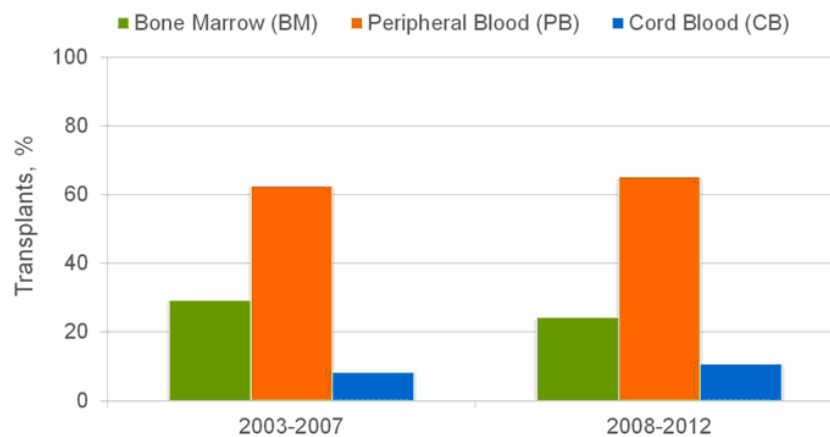
διευκολύνει την ουσιαστική έρευνα που έχει οδηγήσει σε σημαντική αύξηση της επιβίωσης και βελτίωση της ποιότητας της ζωής για χιλιάδες ασθενείς.

#### ΕΤΗΣΙΑ ΕΚΤΙΜΗΣΗ ΠΛΗΘΟΥΣ ΜΕΤΑΜΟΣΧΕΥΣΕΩΝ ΣΤΙΣ Η.Π.Α.



**ΕΙΚΟΝΑ 4.2** Ετήσια εκτίμηση πλήθους αυτόλογων και αλλογενών μεταμοσχεύσεων στις Η.Π.Α, σύμφωνα με το κέντρο CIBMTR (Πηγή : [www.cibmtr.org](http://www.cibmtr.org))

#### ΠΗΓΕΣ ΑΛΛΟΓΕΝΩΝ ΜΟΣΧΕΥΜΑΤΩΝ ΑΝΑ ΧΡΟΝΙΚΕΣ ΠΕΡΙΟΔΟΥΣ ΣΤΙΣ Η.Π.Α



**ΕΙΚΟΝΑ 4.3** Πηγές αλλογενών μοσχευμάτων ανά χρονικές περιόδους στις Η.Π.Α., σύμφωνα με το κέντρο CIBMTR (Πηγή : [www.cibmtr.org](http://www.cibmtr.org))

#### **4.1.2 Ελληνικά Μητρώα Εθελοντών Δοτών Μυελού των Οστών**

Ελληνικά μεταμοσχευτικά κέντρα που έχουν διεθνή άδεια λειτουργίας για πραγματοποίηση μη συγγενικών μεταμοσχεύσεων μυελού των οστών λειτουργούν στα Νοσοκομεία «Ο Ευαγγελισμός» και «Η Αγία Σοφία» στην Αθήνα, στο Νοσοκομείο «Παπανικολάου» στη Θεσσαλονίκη και στο Πανεπιστημιακό Νοσοκομείο Πατρών στο Ρίο. Στην Ελλάδα, κατά τη δεκαετία του 1980, λειτούργησαν δύο δεξαμενές δοτών, μία στην Αθήνα και μία στη Θεσσαλονίκη, στις οποίες αναφέρονταν οι εθελοντές που στρατολογούνταν και τυποποιούνταν σε 5 Κέντρα Δοτών. Από το τέλος του 2002, λειτουργεί μια κεντρική δεξαμενή στον Εθνικό Οργανισμό Μεταμοσχεύσεων (Ε.Ο.Μ.), ο οποίος συντονίζει όλες τις μεταμοσχεύσεις στη χώρα μας.

Οι καταχωρημένοι Έλληνες εθελοντές δότες, στις 25.01.2011, ήταν 29.608, αριθμός εξαιρετικά μικρός, αν αναλογιστεί κανείς τον συνεχώς αυξανόμενο αριθμό των ασθενών που χρειάζονται μεταμόσχευση. Μέχρι σήμερα, για το 95% των ασθενών που έχουν μεταμοσχευθεί, οι δότες προέρχονταν από δεξαμενές του Εξωτερικού. Είναι δεδομένο ότι η αναζήτηση δοτών ανάμεσα σε άτομα με κοινή καταγωγή είναι αποτελεσματικότερη λόγω γενετικών ομοιοτήτων. Επομένως, είναι επιτακτική η ανάγκη αύξησης των Ελλήνων δοτών προς όφελος των Ελλήνων ασθενών, γεγονός που συμβάλλει και στην ουσιαστικότερη και αξιοπρεπέστερη συμμετοχή της χώρας μας στη διεθνή προσπάθεια.

Μοσχεύματα που έχουν δώσει Έλληνες δότες σε ασθενείς:

- 2003 Έλληνας σε Ιταλό ασθενή
- 2004 Έλληνας σε Αμερικανό ασθενή
- 2005 Έλληνας σε Έλληνα ασθενή (Παίδων Αγ. Σοφία)
- 2007 Έλληνας σε Έλληνα ασθενή (Παπανικολάου Θεσσαλονίκης)
- 2008 Έλληνας σε Ιταλό ασθενή
- 2010 Έλληνας σε Ιταλό ασθενή

(Πηγή : Στατιστικά Στοιχεία Εθνικού Οργανισμού Μεταμοσχεύσεων , [www.eom.gr](http://www.eom.gr))

**ΠΙΝΑΚΑΣ Ε.** Στατιστικά στοιχεία μεταμοσχεύσεων στην Ελλάδα  
(Πηγή : Εθνικός Οργανισμός Μεταμοσχεύσεων , [www.eom.gr](http://www.eom.gr))

| ΜΕΤΑΜΟΣΧΕΥΣΕΙΣ ΑΠΟ ΜΗ ΣΥΓΓΕΝΕΙΣ ΕΘΕΛΟΝΤΕΣ ΔΟΤΕΣ |              |                              |                     |                  |        |
|-------------------------------------------------|--------------|------------------------------|---------------------|------------------|--------|
|                                                 | ΕΥΑΓΓΕΛΙΣΜΟΣ | ΠΑΠΑΝΙΚΟΛΑΟΥ<br>ΘΕΣΣΑΛΟΝΙΚΗΣ | ΠΑΙΔΩΝ<br>ΑΓ. ΣΟΦΙΑ | ΡΙΟ<br>ΠΑΤΡΑΣ    | ΣΥΝΟΛΟ |
| 2002                                            | 4            | 0                            | 0                   | 0                | 4      |
| 2003                                            | 3            | 5                            | 2                   | 0                | 10     |
| 2004                                            | 14           | 1                            | 7                   | 0                | 22     |
| 2005                                            | 13           | 8                            | 9                   | 0                | 30     |
| 2006                                            | 19           | 4                            | 13                  | 3                | 39     |
| 2007                                            | 19           | 17                           | 11                  | 4                | 51     |
| 2008                                            | 31           | 16                           | 12                  | 7                | 66     |
| 2009                                            | 29           | 14                           | 12                  | 7                | 62     |
| 2010                                            | 33           | 22                           | 21                  | 10               | 86     |
| 2011                                            | 21           | 22                           | 19                  | 8                | 70     |
| ΣΥΝΟΛΟ                                          | 186          | 109                          | 106                 | 39               | 440    |
| ΠΗΓΕΣ ΜΟΣΧΕΥΜΑΤΩΝ                               |              |                              |                     |                  |        |
|                                                 | ΜΥΕΛΟΣ       | ΠΕΡΙΦΕΡΙΚΑ<br>ΚΥΤΤΑΡΑ        | ΟΠ.Α                | ΛΕΜΦΟ<br>ΚΥΤΤΑΡΑ |        |
| 2007                                            | 10           | 26                           | 15                  | 0                |        |
| 2008                                            | 8            | 46                           | 10                  | 2                |        |
| 2009                                            | 1            | 58                           | 8                   | 0                |        |
| 2010                                            | 13           | 57                           | 13                  | 3                |        |
| 2011                                            | 10           | 42                           | 4                   | 2                |        |
| ΣΥΝΟΛΟ                                          | 42           | 229                          | 50                  | 7                |        |
| τελευταία ενημέρωση 21/10/2011                  |              |                              |                     |                  |        |

Για τους Έλληνες ασθενείς οι **χώρες προέλευσης** των μοσχευμάτων είναι οι εξής : Γερμανία, Αμερική, Κύπρος, Γαλλία, Ισπανία, Αγγλία, Ιταλία, Πορτογαλία, Βέλγιο, Αυστραλία, Καναδάς, Πολωνία, Δανία, Αρμενία, Ισραήλ, Σιγκαπούρη.

Συνηθέστερες χώρες από τις οποίες λαμβάνονται μοσχεύματα για Έλληνες ασθενείς είναι η **Γερμανία**, η **Κύπρος** και η **Αμερική** λόγω της άριστης οργάνωσης των Κέντρων Δοτών τους και της υψηλής ανάλυσης τυποποίησής τους.

#### **4.1.2.1 Κέντρο Ενημέρωσης & Προσέλκυσης Εθελοντών Δοτών Μυελού των Οστών Πανεπιστημίου Πατρών (ΚΕΔΜΟΠ)**

Η ανάγκη ενός παιδιού το 2006, που έπασχε από λευχαιμία, για μεταμόσχευση, ευαισθητοποίησε τους πολίτες της Πάτρας και τη διοίκηση του Πανεπιστημιακού Γενικού Νοσοκομείου Πατρών (ΠΓΝΠ). Ο διευθυντής της Αιματολογικής Κλινικής του ΠΓΝΠ Καθ. κ. Ζούμπος Νικόλαος και ο υπεύθυνος της Μονάδας Μεταμόσχευσης Αιμοποιητικών Κυττάρων Επικ. Καθ. κ. Σπυριδωνίδης Αλέξανδρος γνωρίζοντας ότι η συγκεκριμένη περίπτωση, δυστυχώς, δεν ήταν η μοναδική και η ανάγκη για εθελοντές δότες είναι πάντοτε μεγάλη, πρότειναν, με την υποστήριξη του Διοικητή του Πανεπιστημιακού Γενικού Νοσοκομείου της Πάτρας (ΠΓΝΠ), κ. Δημάρατο, να δημιουργηθεί μια νέα δομή που θα βοηθούσε στην προσέλκυση εθελοντών δοτών. Αυτή συμπεριλάμβανε εκτός από γραφείο ενημέρωσης δοτών και την σύσταση εργαστηρίου τυποποίησης (HLA τυποποίηση).

Στην προσπάθεια αυτή πολύ σημαντική ήταν η υποστήριξη του κ. Κωστέα, διευθυντή του Καραϊσκάκειου Ιδρύματος Κύπρου που αποτελεί το μεγαλύτερο κέντρο δοτών ανά πληθυσμό (>130.000 Κύπριοι εθελοντές δότες). Η υποστήριξή του ήταν τόσο συμβουλευτική όσο και υλική με την παροχή τεχνογνωσίας σε θέματα τυποποίησης και προσέλκυσης δοτών. Στο κέντρο του εκπαιδεύτηκαν η τεχνολόγος του εργαστηρίου Μεταμόσχευσης Μυελού των Οστών του Αιματολογικού Τμήματος του ΠΓΝΠ κ. Οικονομοπούλου Γεωργία σε θέματα HLA τυποποίησης και αργότερα η βιολόγος Dr. Παπάζογλου Άννα σε θέματα προσέλκυσης δοτών. Ως επισφράγιση της συνεργασίας του Πανεπιστημίου Πατρών με το Καραϊσκάκειο Ίδρυμα υπογράφηκε το Μάιο του 2008 επίσημο Πρωτόκολλο Συνεργασίας.

Καθοριστική στη δημιουργία του κέντρου δοτών ήταν η συμβολή του τότε Αντιπρύτανη Στρατηγικού Σχεδιασμού και Ανάπτυξης της Έρευνας του Πανεπιστημίου Πατρών κ. Αναστασόπουλου. Στις αρχές του 2009, εγκρίθηκε η οικονομική υποστήριξη του Κέντρου από το Πρυτανικό Συμβούλιο του Πανεπιστημίου Πατρών (Πρύτανης 2006-2010 - κ. Κουμπιάς) μέσω του προγράμματος της Επιτροπής Ερευνών C654 «Κέντρο Προώθησης Μεταμόσχευσης Μυελού των Οστών και Κυτταρικών Θεραπειών». Έτσι, ιδρύεται το Κέντρο Ενημέρωσης και Προσέλκυσης Εθελοντών Δοτών Μυελού των Οστών του Πανεπιστημίου Πατρών (ΚΕΔΜΟΠ) με ιδρυτικά μέλη τον κ. Β. Αναστασόπουλο, τον κ. Ν. Ζούμπο, την κ. Α. Μουζάκη και τον κ. Α. Σπυριδωνίδη. Τον Ιανουάριο του 2010 το ΚΕΔΜΟΠ λαμβάνει επίσημη έγκριση λειτουργίας από τον Εθνικό

Οργανισμό Μεταμοσχεύσεων (Ε.Ο.Μ.) και ξεκινά την προσέλκυση και εγγραφή των πρώτων εθελοντών δοτών Μυελού των Οστών. Το Οκτώβριο 2010 εγκαινιάζονται τα γραφεία του ΚΕΔΜΟΠ στο κέντρο της Πάτρας, Υψηλών Αλωνίων 11. Το 2010 έγινε η πρώτη καταγραφή δοτών του ΚΕΔΜΟΠ στην Ελληνική και Παγκόσμια Δεξαμενή εθελοντών Δοτών.

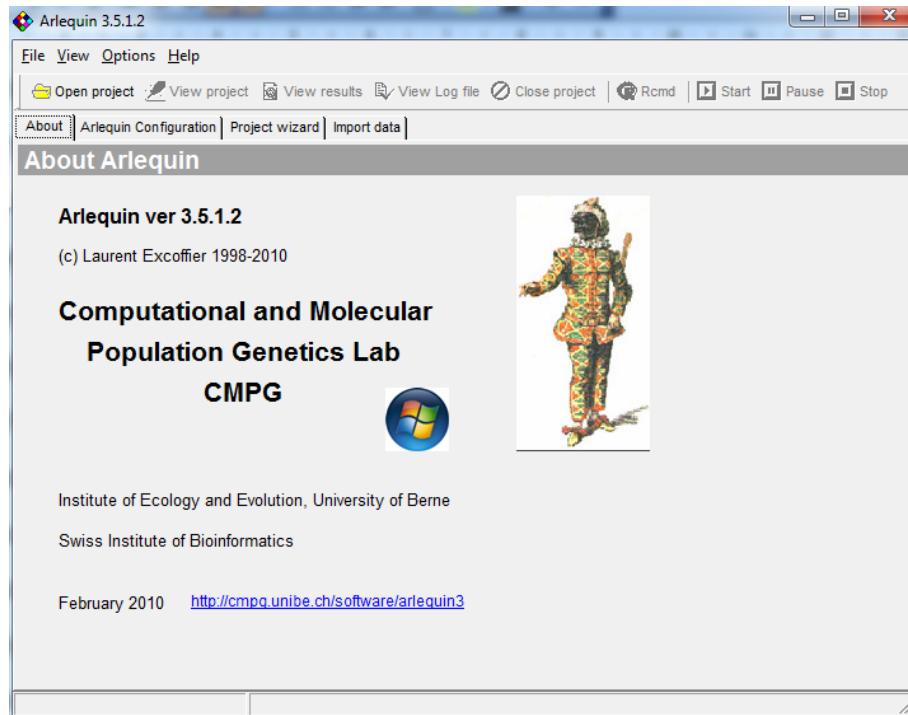
## **4.2 Εργαλεία Διαχείρισης HLA Δεδομένων**

Υπάρχουν αρκετά εργαλεία τα οποία χρησιμοποιούνται για την έγκυρη και αποτελεσματική στατιστική επεξεργασία και διαχείριση των HLA δεδομένων που καταγράφηκαν με τις μεθοδολογίες της HLA τυποποίησης και αποθηκεύτηκαν στα μητρώα εθελοντών δοτών μυελού των οστών. Τα εργαλεία αυτά αποτελούν επαγγελματικά υπολογιστικά προγράμματα και εφαρμογές λογισμικού για την υλοποίηση βιοστατιστικής επεξεργασίας και γενετικής ανάλυσης επί των δεδομένων των μητρώων. Τέτοιες εφαρμογές περιλαμβάνουν μεθοδολογίες για την πραγματοποίηση εκτιμήσεων των συχνοτήτων αλληλόμορφων, αντιγόνων και απλοτύπων, τον υπολογισμό μετρικών όπως η ανισορροπία σύνδεσης (LD), την ανάλυση της ισορροπίας Hardy-Weinberg, στατιστικούς ελέγχους υποθέσεων, φυλογενετική ανάλυση πληθυσμών, υπολογισμό γενετικών αποστάσεων κλπ.

### **4.2.1 Ανάλυση HLA δεδομένων με χρήση του λογισμικού Arlequin**

Η ονομασία Arlequin προέρχεται από την γαλλική μετάφραση του «Arlecchino», ενός διάσημου Ιταλού χαρακτήρα από το «Commedia dell'arte». Πρόκειται για έναν πολύπλευρο χαρακτήρα, που έχει τη δυνατότητα να αλλάζει τα διάφορα στοιχεία του χαρακτήρα του πολύ εύκολα, ανάλογα με τις απαιτήσεις και τις ανάγκες του. Αυτή η πολυμορφική ικανότητα συμβολίζεται με το πολύχρωμο κοστούμι του, από το οποίο προέκυψε και ο σχεδιασμός του εικονιδίου του πακέτου Arlequin, δημιουργοί του οποίου είναι οι Laurent Excoffier, Guillaume Laval, Stephan Schneider, και Heidi Lischer από το Εργαστήριο Υπολογιστικής & Μοριακής Γενετικής Πληθυσμών (CMPG – Computational & Molecular Population Genetics Lab) του Ινστιτούτου Οικολογίας & Εξέλιξης του Πανεπιστημίου της Βέρνης. Το εργαστήριο CMPG είναι μέλος του Ελβετικού Ινστιτούτου Βιοπληροφορικής (Swiss Institute of Bioinformatics). Ο στόχος του Arlequin είναι να παρέχει στο μέσο χρήστη μια

ποικιλία βασικών μεθόδων και στατιστικών ελέγχων γενετικής πληθυσμού, ώστε να προκύψουν πληροφορίες για τα γενετικά και δημογραφικά χαρακτηριστικά της συλλογής των δειγμάτων ενός πληθυσμού. Το Arlequin μπορεί να χειριστεί διάφορους τύπους δεδομένων, είτε σε απλοτυπική, είτε γονοτυπική μορφή, όπως αλληλουχίες DNA, RFLP δεδομένα, μικροδορυφορικά (microsatellite) δεδομένα, τυπικά (standard) δεδομένα, δεδομένα συχνοτήτων αλληλόμορφων.



**ΕΙΚΟΝΑ 4.4** Η γενική καρτέλα πληροφοριών του πακέτου Arlequin 3.5

Τα HLA δεδομένα ανήκουν στην κατηγορία των τυπικών δεδομένων (STANDARD), στην οποία η μοριακή βάση ενός πολυμορφισμού δεν ορίζεται ειδικώς, ή όταν διαφορετικά αλληλόμορφα θεωρούνται ότι ισαπέχουν μεταξύ τους ως προς τη μετάλλαξη. Ως εκ τούτου, τα τυπικά απλοτυπικά δεδομένα συγκρίνονται για το περιεχόμενό τους σε κάθε διαφορετικό γονιδιακό τόπο, χωρίς να λαμβάνεται υπόψη η φύση των αλληλομόρφων, τα οποία μπορούν είτε να είναι όμοια ή διαφορετικά.

#### ***Μέθοδοι που υλοποιούνται στο Arlequin***

Οι μέθοδοι που υλοποιούνται στο Arlequin εκτελούν αναλύσεις επί των δεδομένων εισόδου, οι οποίες εμπίπτουν σε δύο βασικές κατηγορίες:

- A. Ενδοπληθυσμιακές (intra-population),** όπου οι στατιστικές πληροφορίες εξάγονται ξεχωριστά για κάθε πληθυσμό.
- B. Διαπληθυσμιακές (inter-population),** όπου τα δείγματα συγκρίνονται μεταξύ τους, από τα αποτελέσματα που προκύπτουν.

#### **A. Ενδοπληθυσμιακές Αναλύσεις**

- 1. Πρότυποι Δείκτες :** Ορισμένα μέτρα πολυμορφίας, όπως ο αριθμός των πολυμορφικών τόπων, η γονιδιακή ποικιλομορφία.
- 2. Μοριακή Ποικιλότητα :** Υπολογίζει αρκετούς δεικτών ποικιλότητας όπως η διαφορετικότητα των νουκλεοτιδίων, διαφορετικούς εκτιμητές της παραμέτρου  $\theta$  του πληθυσμού.
- 3. Ασυμφωνία Κατανομής :** Η κατανομή του πλήθους των κατά ζεύγη διαφορών μεταξύ των απλοτύπων, με βάση τις οποίες μπορούν να εκτιμηθούν οι παράμετροι μιας δημογραφικής ή χωρικής επέκτασης του πληθυσμού
- 4. Εκτίμηση Απλοτυπικών Συχνοτήτων :** Η εκτίμηση της συχνότητας των απλοτύπων που υπάρχουν στον πληθυσμό μπορεί να υλοποιηθεί με την εφαρμογή στατιστικών μεθόδων μέγιστης πιθανότητας
- 5. Εκτίμηση Γαμετικής Φάσης :** Εκτιμά ότι την πιθανότερη γαμετική φάση των γονοτύπων πολλαπλών γονιδιακών τόπων, χρησιμοποιώντας μια ψευδο-Bayesian προσέγγιση (αλγόριθμος ELB).
- 6. Ανισορροπία Σύνδεσης :** Έλεγχος μη τυχαίας συσχέτιση αλληλόμορφων σε διαφορετικούς γονιδιακούς τόπους.
- 7. Ισορροπία Hardy-Weinberg :** Έλεγχος μη-τυχαίας συσχέτισης των αλληλομόρφων σε διπλοειδή άτομα.
- 8. Έλεγχος ουδετερότητας Tajima :** Έλεγχος της επιλεκτικής ουδετερότητας ενός τυχαίου δείγματος αλληλουχιών DNA ή RFLP απλοτύπων
- 9. Έλεγχος ουδετερότητας Fu's FS :** Έλεγχος της επιλεκτικής ουδετερότητας ενός τυχαίου δείγματος αλληλουχιών DNA ή RFLP
- 10. Έλεγχος ουδετερότητας Ewens-Watterson :** Έλεγχος επιλεκτικής ουδετερότητας που βασίζεται στην θεωρία δειγματοληψίας Ewens σύμφωνα με το μοντέλο άπειρων αλληλόμορφων.
- 11. Έλεγχος συγχώνευσης Chakraborty :** Έλεγχος επιλεκτικής ουδετερότητας και ομοιογένειας του πληθυσμού. Ο συγκεκριμένος έλεγχος μπορεί να χρησιμοποιηθεί όταν υπάρχει υποψία ετερογένειας του δείγματος.
- 12. Ελάχιστο Γενετικό Δένδρο (MSN – Minimum Spanning Tree) :** Υπολογίζει ένα ελάχιστο γενετικό δέντρο (MST) και δίκτυο (MSN) μεταξύ των απλοτύπων. Αυτό το δένδρο μπορεί επίσης να υπολογίζεται για όλους τους απλοτύπους που εντοπίζονται σε διαφορετικούς πληθυσμούς εάν ενεργοποιηθεί στο τμήμα AMOVA.

## **B. Διαπληθυσμιακές Αναλύσεις**

- 1. Αναζήτηση κοινών απλοτύπων μεταξύ των πληθυσμών :** Σύγκριση των δειγμάτων των πληθυσμών ως προς το περιεχόμενο των απλοτύπων τους. Όλα τα αποτελέσματα συνοψίζονται στη συνέχεια σε έναν πίνακα.
- 2. AMOVA :** Διαφορετικές ιεραρχικές αναλύσεις μοριακής μεταβλητότητας για να υπολογιστεί η γενετική δομή του πληθυσμού.
- 3. Ζεύγη Γενετικών Αποστάσεων :** Υπολογισμός FST γενετικών αποστάσεων για αποκλίσεις σύντομου χρονικού διαστήματος.
- 4. Ακριβής έλεγχος της διαφοροποίησης του πληθυσμού :** Έλεγχος μη-τυχαίας κατανομής των απλοτύπων σε δείγματα πληθυσμού κάτω από την υπόθεση της παμμυξίας.
- 5. Έλεγχος Μεταβίβασης Γονοτύπων :** Μεταβίβαση των ατομικών γονοτύπων σε συγκεκριμένους πληθυσμούς σύμφωνα με τις εκτιμώμενες συχνότητες αλληλόμορφων.
- 6. Ανίχνευση γονιδιακών θέσεων υπό επιλογή μέσω F-στατιστικών :** Ανίχνευση γονιδιακών υπό επιλογή, μέσω της εξέτασης της από κοινού κατανομής των FST και ετεροζυγωτίας στο πλαίσιο ενός ιεραρχικού μοντέλου.
- 7. Έλεγχος Μέντελ (Mantel Test) :** Οι συσχετίσεις ή οι μερικές συσχετίσεις μεταξύ ενός συνόλου δύο ή τριών μητρώων μπορούν να χρησιμοποιηθούν για να ελεγχθεί η παρουσία της απομόνωσης από απόσταση.

# Κ Ε Φ Α Λ Α Ι Ο 5<sup>ο</sup>

## ΕΠΕΞΕΡΓΑΣΙΑ & ΑΝΑΛΥΣΗ ΗΛΑ ΔΕΔΟΜΕΝΩΝ ΕΛΛΗΝΙΚΟΥ ΜΗΤΡΩΟΥ 5.131 ΕΘΕΛΟΝΤΩΝ ΔΟΤΩΝ ΜΕΣΩ ΤΟΥ ΠΑΚΕΤΟΥ Arlequin 3.5

### 5.1 Ανάπτυξη Μητρώου Ελλήνων Εθελοντών Δοτών ΚΕΔΜΟΠ

Το Κέντρο Ενημέρωσης & Προσέλευσης Εθελοντών Δοτών Μυελού των Οστών του Πανεπιστημίου Πατρών (ΚΕΔΜΟΠ) έχει αρχίσει να δημιουργεί μητρώο εθελοντών δοτών μυελού των οστών μέσω δράσεων ενημέρωσης και προσέλευσης που αναπτύσσει. Από τον Οκτώβριο του 2009 εγγράφονται στο μητρώο εθελοντές δότες μέσω της συμπλήρωσης ειδικής αίτησης εγγραφής δότη μυελού των οστών.

ΚΕΝΤΡΟ ΕΝΗΜΕΡΩΣΗΣ & ΠΡΟΣΕΛΕΥΣΗΣ ΕΘΕΛΟΝΤΩΝ ΔΟΤΩΝ ΜΥΕΛΟΥ ΤΩΝ ΟΣΤΩΝ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΠΑΤΡΩΝ (ΚΕΔΜΟΠ)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ ΣΤΡΩΤΗΝ ΕΡΕΥΡΗΜΩΝ - ΤΜΗΜΑ ΑΙΜΑΤΟΛΟΓΙΑΣ ΠΑΝΕΠΙΣΤΗΜΙΟ ΤΕΧΝΟΛΟΓΙΚΟΥ ΠΑΤΡΩΝ

#### ΑΙΤΗΣΗ ΕΓΓΡΑΦΗΣ ΔΟΤΗ ΜΥΕΛΟΥ ΤΩΝ ΟΣΤΩΝ

Τα προσωπικά στοιχεία που θα μας δώσετε είναι εμπιστευτικά, δεν δίδονται σε τρίτους και θα επεξεργαστούν με απόλυτη εμπιστοσύνη. Το ΚΕΔΜΟΠ λειτουργεί σύμφωνα με την υπ' αριθμό 2201/90-12-2009 απόφαση του Διοικητικού Συμβουλίου του Ελληνικού Οργανισμού Μεταμοσχεύσεων (ΕΟΜ).

Παρακαλούμε να συμπληρώσετε την αίτηση με ΚΕΦΑΛΑΙΑ ΓΡΑΜΜΑΤΑ (Διχράστε δότες ηλικίας 18-50 ετών)

|                                                                                                                                                                                                                                                                                                                          |                                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| ΟΝΟΜΑΤΕΠΩΝΥΜΟ:                                                                                                                                                                                                                                                                                                           |                                |
| ΟΝΟΜΑΤΕΠΩΝΥΜΟ ΠΑΤΕΡΑ:                                                                                                                                                                                                                                                                                                    |                                |
| Άνδρας <input type="checkbox"/>                                                                                                                                                                                                                                                                                          | Ημερ. Γέννησης: HH / MM / ΕΤΟΣ |
| Γυναίκα <input type="checkbox"/>                                                                                                                                                                                                                                                                                         | Βόρος (σε ιαλά):               |
|                                                                                                                                                                                                                                                                                                                          | Υψος (σε εκατοστά):            |
|                                                                                                                                                                                                                                                                                                                          | Ομάδα Αίματος:                 |
| Τόπος Γέννησης:                                                                                                                                                                                                                                                                                                          | Εθνικότητα:                    |
| Τόπος Κατοικίας (Χώρα, Νομός, Δήμος, Χωριά)                                                                                                                                                                                                                                                                              |                                |
| Δευτερεύουσα κατοικία:                                                                                                                                                                                                                                                                                                   | Τηλέφωνο Κατοικίας:            |
| Οδός:                                                                                                                                                                                                                                                                                                                    | Τηλέφωνο Κινητό:               |
| ΤΚ, Περιφέρεια:                                                                                                                                                                                                                                                                                                          | Τηλέφωνο Εργασίας:             |
| Επάγγελμα:                                                                                                                                                                                                                                                                                                               | Ασφαλιστικός Φορέας:           |
| Αρ. Δελτίου Ταυτότητας / Διαβατηρίου:                                                                                                                                                                                                                                                                                    | Ημερ. Έκδοσης HH / MM / ΕΤΟΣ   |
| Αρχή Έκδοσης:                                                                                                                                                                                                                                                                                                            |                                |
| email:                                                                                                                                                                                                                                                                                                                   |                                |
| Είμαι απαραίτητο να γνωρίζουμε τα στοιχεία επικοινωνίας ενός στενού συγγενή ή φίλου σας για να μπορούμε να έρθουμε σε επαφή μαζί σας στην περίπτωση που τα δικά σας στοιχεία έχουν διαφοροποιηθεί. Παρακαλούμε συμπληρώστε.                                                                                              |                                |
| ΟΝΟΜΑΤΕΠΩΝΥΜΟ:                                                                                                                                                                                                                                                                                                           |                                |
| Σχέση:                                                                                                                                                                                                                                                                                                                   |                                |
| Άνδρας <input type="checkbox"/>                                                                                                                                                                                                                                                                                          | Τηλέφωνο Κατοικίας:            |
| Γυναίκα <input type="checkbox"/>                                                                                                                                                                                                                                                                                         | Τηλέφωνο Κινητό:               |
|                                                                                                                                                                                                                                                                                                                          | Τηλέφωνο Εργασίας:             |
| email:                                                                                                                                                                                                                                                                                                                   |                                |
| Έχω διαβάσει το ενημερωτικό φυλλάδιο για τον Δότη Μυελού των Οστών. Μου δόθηκε η ευκαιρία να υποβάλω ερωτήσεις και πήρα τις απαραίτητες δυνατότες απαντήσεων. Επιθυμώ να καταχωρηθώ στο Εθνικό και Πανεπιστημιακό Μητρώο Εθελοντών Δοτών Μυελού των Οστών. Συμφωνώ το δείγμα αίματός μου να τυποποιηθεί για ΗΛΑ ανιχνών. |                                |
| Ημερομηνία                                                                                                                                                                                                                                                                                                               | Υπογραφή                       |

ΚΕΝΤΡΟ ΕΝΗΜΕΡΩΣΗΣ & ΠΡΟΣΕΛΕΥΣΗΣ ΕΘΕΛΟΝΤΩΝ ΔΟΤΩΝ ΜΥΕΛΟΥ ΤΩΝ ΟΣΤΩΝ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΠΑΤΡΩΝ (ΚΕΔΜΟΠ)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ ΣΤΡΩΤΗΝ ΕΡΕΥΡΗΜΩΝ - ΤΜΗΜΑ ΑΙΜΑΤΟΛΟΓΙΑΣ ΠΑΝΕΠΙΣΤΗΜΙΟ ΤΕΧΝΟΛΟΓΙΚΟΥ ΠΑΤΡΩΝ

#### ΙΑΤΡΙΚΟ ΙΣΤΟΡΙΚΟ

Τα προσωπικά στοιχεία που θα μας δώσετε είναι εμπιστευτικά, δεν δίδονται σε τρίτους και θα επεξεργαστούν με απόλυτη εμπιστοσύνη. Το ΚΕΔΜΟΠ λειτουργεί σύμφωνα με την υπ' αριθμό 2201/90-12-2009 απόφαση του Διοικητικού Συμβουλίου του Ελληνικού Οργανισμού Μεταμοσχεύσεων (ΕΟΜ).

Σκοπός της συλλογής αυτής του ιστορικού είναι να μετριάσει τον κίνδυνο τόσο για τον ασθενή όσο και για το δότη. Παρακαλούμε οι πληροφορίες που δίνονται να είναι ακριβείς.

|                                                                                                               |                                 |                                  |                                 |
|---------------------------------------------------------------------------------------------------------------|---------------------------------|----------------------------------|---------------------------------|
| ΟΝΟΜΑΤΕΠΩΝΥΜΟ ΔΟΤΗ:                                                                                           |                                 |                                  |                                 |
| Ημερ. Γέννησης HH / MM / ΕΤΟΣ                                                                                 | ΑΝΔΡΑΣ <input type="checkbox"/> | ΓΥΝΑΙΚΑ <input type="checkbox"/> |                                 |
| 1. Έχετε διαγνωστεί με:                                                                                       | ΗΠΑΤΙΤΙΔΑ Β ή C                 | ΝΑΙ <input type="checkbox"/>     | ΟΧΙ <input type="checkbox"/>    |
|                                                                                                               | ΚΑΡΚΙΝΟ                         | ΝΑΙ <input type="checkbox"/>     | ΟΧΙ <input type="checkbox"/>    |
|                                                                                                               | AIDS                            | ΝΑΙ <input type="checkbox"/>     | ΟΧΙ <input type="checkbox"/>    |
| Αν απαντήσετε ΝΑΙ στην πιο πάνω ερώτηση δεν μπορείτε να εγγραφείτε δότης. Ευχαριστούμε για το ενδιαφέρον σας. |                                 |                                  |                                 |
| 2. Πως θα περιγράφατε την κατάσταση της υγείας σας;                                                           | Καλή <input type="checkbox"/>   | Μέτρια <input type="checkbox"/>  | Χαμηλή <input type="checkbox"/> |
| 3. Προσπαθήσατε ή υποφέρατε ποτέ από:                                                                         |                                 |                                  |                                 |
| ΗΠΑΤΙΤΙΔΑ Α                                                                                                   | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΙΚΤΕΡΟ                                                                                                        | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΔΙΑΒΗΤΗ                                                                                                       | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΡΕΥΜΑΤΙΚΑ                                                                                                     | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΕΠΙΛΗΨΙΑ                                                                                                      | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΦΥΜΑΤΙΩΣΗ                                                                                                     | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΑΣΘΜΑ                                                                                                         | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΚΑΡΔΙΑΚΕΣ ΠΑΘΗΣΕΙΣ                                                                                            | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΠΝΕΥΜΑΤΙΚΕΣ ΔΙΑΤΑΡΑΧΕΣ / ΚΛΙΝΙΚΗ ΚΑΤΑΘΛΙΨΗ                                                                    | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΣΤΙΟΛΑΝ/ΚΗ ΠΑΘΗΣΗ / ΔΙΣΧΟΛΙΑΣΙΑ                                                                               | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| ΑΛΛΗ ΣΟΒΑΡΗ ΑΣΘΕΝΕΙΑ (π.χ. υψηλή πίεση, διαβήτης, αλλεργίες)                                                  | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| Αν απαντήσετε ΝΑΙ στις πιο πάνω επιλογές, παρακαλώ εξηγήστε περιληπτικά.                                      |                                 |                                  |                                 |
| 4. Πιάνετε οποιαδήποτε φάρμακα (περιλαμβανομένων αντισυλληπτικών χαπιών)                                      | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| Αν ΝΑΙ, ποια και γιατί:                                                                                       |                                 |                                  |                                 |
| 5. Έχετε υποστεί χειρουργική επέμβαση τον τελευταίο χρόνο;                                                    | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| Αν ΝΑΙ, ποια και γιατί:                                                                                       |                                 |                                  |                                 |
| 6. Είστε αιμοδότης;                                                                                           | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| 7. Αποκλειστικάτε ποτέ από αιμοδότη;                                                                          | ΝΑΙ <input type="checkbox"/>    | ΟΧΙ <input type="checkbox"/>     |                                 |
| Αν ΝΑΙ, γιατί:                                                                                                |                                 |                                  |                                 |
| 8. Για Γυναίκες μόνο                                                                                          | Έχετε παιδιά:                   | ΝΑΙ <input type="checkbox"/>     | ΟΧΙ <input type="checkbox"/>    |
|                                                                                                               |                                 |                                  | Αριθμός παιδιών:                |
| Ημερομηνία                                                                                                    | Υπογραφή                        |                                  |                                 |

ΚΕΔΜΟΠ : Τηλέφωνο: 2610 986872 - 2613 603255, Fax: 2610 983850, email: cbmdpatras@upatras.gr, website: www.xarisezoi.gr

ΚΕΔΜΟΠ : Τηλέφωνο: 2610 986872 - 2613 603255, Fax: 2610 983850, email: cbmdpatras@upatras.gr, website: www.xarisezoi.gr

ΕΙΚΟΝΑ 5.1 Η ειδική αίτηση εγγραφής εθελοντή δότη μυελού των οστών του ΚΕΔΜΟΠ

Στη συνέχεια, εφόσον ο υποψήφιος εθελοντής δότης πληροί τις απαιτούμενες ιατρικές προϋποθέσεις και κριτήρια για να μπορεί να γίνει δότης μυελού των οστών (να μην πάσχει από συγκεκριμένες χρόνιες παθήσεις, μεταδοτικές ασθένειες και αυτοάνοσα νοσήματα), καταγράφονται στο αρχείο του ΚΕΔΜΟΠ τα δεδομένα που βρίσκονται στην αίτηση εγγραφής, δηλ. τα προσωπικά του στοιχεία, τα στοιχεία επικοινωνίας του ιδίου και συγγενών του, πληροφορίες του ιατρικού ιστορικού του κλπ. και στη συνέχεια γίνεται λήψη επιχρίσματος από το στόμα (SWAB - δείγμα σάλιου με μπατονέτα) ή αίματος (BLOOD - 5-6 ml αίματος). Ακολουθεί η εργαστηριακή ανάλυση με την διαδικασία της HLA τυποποίησης του δείγματος του εθελοντή δότη, η οποία πραγματοποιείται σε κάποιο εξειδικευμένο εργαστήριο HLA τυποποίησης (π.χ. Histogenetics Lab στις ΗΠΑ). Τα HLA δεδομένα που προκύπτουν από την υψηλής ανάλυσης HLA τυποποίηση για τα γονίδια HLA-A, HLA-B και HLA-DRB1 καταχωρούνται στο μητρώο του ΚΕΔΜΟΠ, αλλά επιπλέον και στην Ελληνική (Ε.Ο.Μ.) και στην Παγκόσμια (BMDW) δεξαμενή εθελοντών δοτών μυελού των οστών.

Αν κάποιος εθελοντής δότης βρεθεί συμβατός με έναν ασθενή:

1. Το ΚΕΔΜΟΠ τον ενημερώνει ότι βρέθηκε συμβατός με έναν ασθενή που χρειάζεται μεταμόσχευση και τον ρωτά εάν συνεχίζει να μπορεί και να επιθυμεί να κάνει δωρεά.
2. Καλείται για τις απαραίτητες ιατρικές εξετάσεις.
3. Η δωρεά πραγματοποιείται μέσω της μεταμόσχευσης αιμοποιητικών βλαστικών κυττάρων με μια μέθοδο που ονομάζεται **λευκαφαίρεση**.

| Patras Donor ID | Last Name             | First Name | Last Name(In English) | First Name(In English) | Father's Name | Gender (M: male, F: female) |
|-----------------|-----------------------|------------|-----------------------|------------------------|---------------|-----------------------------|
| CBMDP-0000501   | ΔΑΟΥΛΑΡΗ              | ΝΙΚΟΛΕΤΤΑ  | ΔΑΟΥΛΑΡΙ              | ΝΙΚΟΛΕΤΤΑ              | ΔΙΟΝΥΣΙΟΣ     | F                           |
| CBMDP-0000502   | ΚΑΡΑΜΠΙΝΑ             | STELA      | KARABINA              | STELA                  | FATMIR        | F                           |
| CBMDP-0000503   | ΚΑΡΑΓΙΑΝΝΗ            | ΦΩΤΙΟΣ     | KARAGIANNI            | FOTIOS                 | ΑΡΙΣΤΕΙΔΗΣ    | M                           |
| CBMDP-0000504   | ΚΑΛΥΒΙΩΤΗ             | ΕΛΕΝΗ      | KALYVIOTI             | ELENI                  | ΣΩΤΗΡΙΟΣ      | F                           |
| CBMDP-0000505   | ΘΕΜΕΛΗ                | MARIA      | THEMELI               | MARIA                  | ΠΑΝΑΓΙΩΤΗΣ    | F                           |
| CBMDP-0000506   | ΚΟΚΚΙΝΟΥ              | ΔΗΜΗΤΡΑ    | KOKKINOY              | DIMITRA                | ΒΑΣΙΛΕΙΟΣ     | F                           |
| CBMDP-0000507   | ΚΑΒΑΣΗ                | MARIA      | KAVASI                | MARIA                  | ΝΙΚΟΛΑΟΣ      | F                           |
| CBMDP-0000508   | ΛΑΖΑΡΗΣ               | ΒΑΣΙΛΕΙΟΣ  | LAZARIS               | VASILEIOS              | ΜΙΧΑΗΛ        | M                           |
| CBMDP-0000509   | ΚΑΡΑΓΙΑΝΝΗ            | ΕΥΓΕΝΙΑ    | KARAGIANNI            | EVGENIA                | ΔΗΜΗΤΡΙΟΣ     | F                           |
| CBMDP-0000510   | ΚΑΡΑΓΙΑΝΝΗ            | ΑΙΚΑΤΕΡΙΝΗ | KARAGIANNI            | AIKATERINI             | ΦΩΤΙΟΣ        | F                           |
| CBMDP-0000511   | ΒΓΕΝΗ-ΠΑΝΑΓΙΩΤΟΠΟΥΛΟΥ | MARIA      | VGENI-PANAGIOTOPOULOU | MARIA                  | ΑΝΔΡΕΑΣ       | F                           |
| CBMDP-0000512   | ΚΑΤΣΟΜΑΛΙΑΡΗΣ         | ΓΕΩΡΓΙΟΣ   | KATSOMALIARIS         | GEORGIOS               | ΙΩΑΝΝΗΣ       | M                           |
| CBMDP-0000513   | ΚΑΡΑΓΙΑΝΝΗΣ           | ΑΘΑΝΑΣΙΟΣ  | KARAGIANNIS           | ATHANASIOS             | ΦΩΤΙΟΣ        | M                           |

**ΕΙΚΟΝΑ 5.2** Τμήμα δεδομένων ταυτότητας εγγεγραμμένων δοτών μητρώου ΚΕΔΜΟΠ

| A or A*(1) | A or A*(2) | B or B*(1) | B or B*(2) | DR or DRB1*(1) | DR or DRB1*(2) |
|------------|------------|------------|------------|----------------|----------------|
| 02:DFKP    | 02:DFKP    | 18:RRG     | 51:CVAE    | 16:01:01       | 11:04:01       |
| 01:DCEJ    | 24:DRZV    | 08:BYXR    | 27:BHXW    | 03:01:01       | 04:04          |
| 02:DFKP    | 30:AUX     | 35:BEZX    | 51:CVAE    | 16:01:01       | 04:03:01       |
| 01:CNJK    | 32:01      | 44:WRJ     | 52:AH      | 15:02:01       | 11:04:01       |
| 24:DUHT    | 32:UVF     | 35:BEZX    | 44:WRJ     | 15:01:01       | 16:02:01       |
| 01:CPZT    | 68:AUW     | 08:XKT     | 53:01:01   | 03:01:01       | 11:01:01       |
| 02:DFKP    | 26:GAX     | 14:01      | 39:06:02   | 16:01:01       | 07:01:01       |
| 11:BEYE    | 26:BTHM    | 35:BEZX    | 55:01      | 07:01:01       | 10:01:01       |
| 02:DFKP    | 68:01      | 15:17:01   | 18:RRG     | 14:BCAD        | 07:01:01       |
| 33:01:01   | 68:01      | 15:17:01   | 35:02:01   | 11:04:01       | 14:BCAD        |
| 03:ECAM    | 23:ENWE    | 13:02      | 49:01:01   | 11:01          | 11:04:01       |
| 31:DUKU    | 32:01      | 08:XKT     | 15:08      | 03:01:01       | 04:05:01       |
| 02:DFKP    | 24:CWFP    | 18:RRG     | 49:01:01   | 15:01:01       | 07:01:01       |

**ΕΙΚΟΝΑ 5.3** Τμήμα δεδομένων HLA τυποποίησης εγγεγραμμένων δοτών μητρώου ΚΕΔΜΟΠ

## 5.2 Αντικείμενα & Μέθοδοι HLA Ανάλυσης

### 5.2.1 Στοιχεία Μητρώου Δοτών ΚΕΔΜΟΠ

Βάσει των δεδομένων που περιλαμβάνονται στο μητρώο του ΚΕΔΜΟΠ, εκτιμήσαμε τις συχνότητες των HLA αλληλόμορφων γονιδίων, των HLA αντιγόνων και των HLA απλοτύπων του μητρώου Ελλήνων εθελοντών δοτών αιμοποιητικών βλαστικών κυττάρων για τα γονίδια HLA-A, HLA-B και HLA-DRB1. Οι αναλύσεις που υλοποιήθηκαν βασίζονται σε δεδομένα που συλλέχθηκαν από δότες που εγγράφηκαν στο μητρώο Ελλήνων εθελοντών δοτών μυελού των οστών του ΚΕΔΜΟΠ κατά το διάστημα μεταξύ Οκτωβρίου 2009 και Απριλίου 2014, μέσω των τμημάτων αιμοδοσίας διαφόρων νοσοκομείων της Ελλάδας ή μέσω εκδηλώσεων ενημέρωσης και προσέλευσης εθελοντών δοτών που διοργάνωνε το ΚΕΔΜΟΠ και διάφοροι αιμοδοτικοί σύλλογοι. Το μητρώο του ΚΕΔΜΟΠ αποτελείται από 5.131 Έλληνες εθελοντές δότες μυελού των οστών και περιλαμβάνει πληροφορίες σχετικές με στοιχεία ταυτότητας (φύλο, ηλικία, τηλέφωνα επικοινωνίας, διευθύνσεις κατοικίας, τόπος καταγωγής, στοιχεία συγγενών), στοιχεία εγγραφής (ημερομηνία εγγραφής, τόπος εγγραφής, τρόπος ενημέρωσης), βιολογικά στοιχεία (τύπος δείγματος, ύψος, βάρος, αριθμός εγκυμοσύνων) και στοιχεία ιατρικού ιστορικού. Τα αποτελέσματα που προέκυψαν χρησιμοποιούνται για να εκτιμηθεί η πιθανότητα εύρεσης πλήρως

συμβατών δοτών από το διαθέσιμο μητρώο εθελοντών δοτών. Ειδικότερα, έχουμε αναλύσει τον τρόπο που οι πιθανότητες εξαρτώνται από διάφορες αντίστοιχες απαιτήσεις.

### **5.2.2 Στοιχεία Δείγματος Πληθυσμού**

Για καθέναν από τους δότες, τυποποιήθηκαν τα γονίδια HLA-A, HLA-B, και HLA-DRB1 σε υψηλή ανάλυση, από το εργαστήριο Histogenetics Lab των ΗΠΑ.

**ΦΥΛΟ:** Από τους 5.131 δότες, το 52,37% είναι γυναίκες και το υπόλοιπο 47,63% άνδρες.

**ΗΛΙΚΙΑ :** Από τους 5.131 δότες, το 35,59% (1.826) είναι ηλικίας 30-40 ετών, το 28,65% (1.470) είναι ηλικίας 40-50 ετών, το 27,30% (1.401) είναι ηλικίας 20-30 ετών, το 5,5% (280) είναι ηλικίας άνω των 50 ετών και το 2,96% (154) είναι ηλικίας κάτω των 20 ετών.

Το σύνολο των δοτών δήλωσαν ότι είναι Ελληνικής καταγωγής και ένα πολύ μικρό ποσοστό της τάξης του 1% δήλωσαν αλλοδαποί, οι οποίοι κατοικούν στην Ελλάδα. Οι τόποι καταγωγής και κατοικίας του συνόλου των δοτών κατανέμονται σε όλες τις διοικητικές περιφέρειες της ελληνικής επικράτειας (ηπειρωτική και νησιωτική χώρα). Λόγω της έδρας του ΚΕΔΜΟΠ στην Πάτρα αλλά και των δράσεων ενημέρωσης και προσέλκυσης εθελοντών δοτών που αυτό πραγματοποιούσε κατά καιρούς σε διάφορες περιοχές της Ελλάδας, το πλήθος των δοτών που έχουν εγγραφεί στο μητρώο κατανέμεται από 17 στην Ήπειρο έως 1.882 στην Δυτική Ελλάδα.

Πιο συγκεκριμένα η κατανομή έχει ως εξής :

Το 36,68% (1.882) των δοτών προέρχονται από την Δυτική Ελλάδα, το 20,46% (1.050) από το Νότιο Αιγαίο, το 18,22% (935) από την Αττική, το 7,7% (394) από την Πελοπόννησο, το 5,67% (291) από την Στερεά Ελλάδα, το 5% (257) από την Μακεδονία, το 3,9% (200) από την Κρήτη και το υπόλοιπο 2,37% (122) από τη Θεσσαλία, τα Ιόνια Νησιά, την Ήπειρο και την Κύπρο.

### **5.2.3 HLA τυποποίηση και επεξεργασία HLA δεδομένων**

Η υψηλής ανάλυσης HLA τυποποίηση των δεδομένων του μητρώου του ΚΕΔΜΟΠ διεξήχθη στο διαπιστευμένο από την Αμερικάνικη Εταιρεία Ιστοσυμβατότητας & Ανοσογενετικής (ASHI - American Society for Histocompatibility and Immunogenetics) Εργαστήριο Histogenetics Inc. (Ossining, Νέα Υόρκη). Κατά την HLA τυποποίηση αλληλουχίστηκαν τα εξώνια 2 και 3 για τα HLA γονίδια τάξης I. Για το γονίδιο HLA-DRB1, αλληλουχίστηκε το εξώνιο 2. Οι Cis-trans ασάφειες εντός

των σχετικών εξωνίων επιλύθηκαν. Συνώνυμες μεταλλάξεις δεν ελήφθησαν υπόψη. Αλληλόμορφα με συνώνυμες μεταλλάξεις εντός ή εκτός των σχετικών εξωνίων συγχωνεύονται. Η μορφή της ονοματολογίας των αλληλόμορφων, έτσι όπως προκύπτουν από την υψηλής ανάλυσης HLA τυποποίηση του εργαστηρίου Histogenetics Inc. φαίνεται παρακάτω

(για τον δότη με κωδικό ΚΕΔΜΟΠ : CBMDP-0000501)

| CBMDP-0000501 |            |            |            |                |                |
|---------------|------------|------------|------------|----------------|----------------|
| A or A*(1)    | A or A*(2) | B or B*(1) | B or B*(2) | DR or DRB1*(1) | DR or DRB1*(2) |
| 02:DFKP       | 02:DFKP    | 18:RRG     | 51:CVAE    | 16:01:01       | 11:04:01       |

#### 5.2.3.1 DNA Type Lookup Tool

Οι κωδικοί ονοματολογίας των αλληλόμορφων αποτελούν μια στοιχειώδη σημειογραφία που χρησιμοποιείται από το NMDP για την αναπαράσταση των συνδυασμών των αλληλόμορφων της βάσης δεδομένων των μητρώων των δοτών κατά την μοριακή HLA τυποποίησή τους. Μόλις εντοπιστεί ένας νέος συνδυασμός αλληλόμορφων, ένας κωδικός εκχωρείται στον συνδυασμό αυτό, ο οποίος στη συνέχεια ενεργοποιείται (ACTIVATED) και είναι διαθέσιμος για χρήση κατά την αναπαράσταση του συγκεκριμένου γονιδιακού τόπου και της συγκεκριμένης οικογένειας αλληλόμορφων όπου ανήκει. Το εργαλείο **DNA Type Lookup Tool** χρησιμοποιείται για την μετατροπή των κωδικών κάθε αλληλόμορφου στην συνήθη μορφή των 4 ψηφίων (4-digit). Η μετατροπή των κωδικών πραγματοποιείται μέσω της πρόσβασης που διαθέτει το συγκεκριμένο εργαλείο σε δεδομένα πινάκων με κωδικούς αλληλόμορφων, οι οποίοι ανανεώνονται σε καθημερινή βάση. Το εργαλείο διαθέτει δυνατότητα διαδραστικότητας μέσω επικοινωνίας με ηλεκτρονική αλληλογραφία για την εξυπηρέτηση αιτημάτων απόδοσης κωδικών νέων αλληλόμορφων ή για την ενεργοποίηση ενός υπάρχοντος κωδικού σε διαφορετική γονιδιακή θέση.

Για τον δότη με κωδικό ΚΕΔΜΟΠ : CBMDP-0000501, ο οποίος είναι ομόζυγος ως προς το γονίδιο HLA-A, η μετατροπή των κωδικών των αλληλόμορφων φαίνεται παρακάτω. Για κάθε κωδικό, κατά τη μετατροπή, από το σύνολο των αποτελεσμάτων επιλέγεται ο πρώτος κωδικός αλληλόμορφου, ο οποίος είναι και ο συχνότερος που έχει καταγραφεί, έως εκείνη τη στιγμή. **Μικρότερος αριθμός κωδικού αναπαριστά και μεγαλύτερη πιθανότητα ιστοσυμβατότητας.**

### DNA Type Lookup Tool

[Allele Code Update to Version 3 WHO Nomenclature](#)

[Instructions \(PDF\)](#)

**Locus:**

☒ A     ☐ B     ☐ C  
☐ DRB1   ☐ DRB3   ☐ DRB4   ☐ DRB5  
☐ DQA1   ☐ DQB1   ☐ DPA1   ☐ DPB1   ☐ None

**Nomenclature Version:**

You may request allele codes for strings of alleles in the new Nomenclature or the old Nomenclature.

☐ Old     ☒ New

**Enter a list of alleles and/or allele codes:**

02:DFKP

Υποβολή ερωτήματος

Επαναφορά

**New Nomenclature**

**Allele Code:**

A\*02:DFKP -> 02:01/02:01L/02:09/02:43N/02:66/02:75/02:83N/02:89/02:97/02:132/02:134/02:140

DFKP -> 02:01/02:01L/02:09/02:43N/02:66/02:75/02:83N/02:89/02:97/02:132/02:134/02:140

**ACTIVATED** for use at:

A\*02:DFKP

**ΕΙΚΟΝΑ 5.4** Μετατροπή κωδικού αλληλόμορφου σε 4-digit μορφή μέσω του DNA Type Lookup Tool

| CBMDP-0000501 |            |            |            |                |                |
|---------------|------------|------------|------------|----------------|----------------|
| A or A*(1)    | A or A*(2) | B or B*(1) | B or B*(2) | DR or DRB1*(1) | DR or DRB1*(2) |
| 02:01         | 02:01      | 18:01      | 51:01      | 16:01          | 11:04          |

### **Ελλιπή Δεδομένα (Missing Data)**

Το Arlequin αντιμετωπίζει τα ελλιπή δεδομένα (NULL) για οποιοδήποτε αλληλόμορφο όπως κάθε άλλο καταγεγραμμένο αλληλόμορφο, γεγονός που κατά την ανάλυση των HLA δεδομένων δημιουργεί απλοτύπους που δεν ανταποκρίνονται στην πραγματικότητα. Ωστόσο, η συχνότητα των ελλিপών τιμών λαμβάνεται υπόψη κατά τον υπολογισμό των συχνοτήτων αλληλόμορφων και απλοτύπων.

Ο **κωδικός XX** είναι ένας ειδικός κωδικός αλληλόμορφου, ο οποίος αναπαριστά όλα τα αλληλόμορφα που είναι ενεργοποιημένα εντός του συστήματος σε δεδομένη χρονική στιγμή, για μια συγκεκριμένη οικογένεια αλληλόμορφων και συγκεκριμένο γονιδιακό τόπο.

| CBMDP-0000696 |            |            |            |                |                |
|---------------|------------|------------|------------|----------------|----------------|
| A or A*(1)    | A or A*(2) | B or B*(1) | B or B*(2) | DR or DRB1*(1) | DR or DRB1*(2) |
| 02:AUWF       |            | 39:XX      | 51:XX      | 04:XX          | 16:XX          |
|               |            |            |            |                |                |
| 02:07         |            | 39:XX      | 51:XX      | 04:164         | 16:XX          |

### **ΕΞΕΤΑΖΟΜΕΝΑ ΣΕΝΑΡΙΑ HLA ΑΝΑΛΥΣΗΣ**

Κατά τους υπολογισμούς που διενεργήθηκαν μέσω του Arlequin, εξετάστηκαν τρία διαφορετικά σενάρια ως προς τα επίπεδα της HLA ανάλυσης των τυποποιημένων δεδομένων των δοτών που είναι καταχωρημένα στο μητρώο εθελοντών δοτών του ΚΕΔΜΟΠ.

#### **ΣΕΝΑΡΙΟ A : High Resolution HLA-A, HLA-B & HLA-DRB1**

Το βασικό σενάριο A αναφέρεται στην διαχείριση των δεδομένων, όταν αυτά αναπαρίστανται σε υψηλής ανάλυσης HLA τυποποίηση για τα HLA-A, HLA-B και HLA-DRB1 γονίδια. Η υψηλής ανάλυσης HLA τυποποίηση των αλληλόμορφων αντιστοιχεί στην 4-digit αναπαράσταση (Επίπεδο Αλληλόμορφου).

### **ΣΕΝΑΡΙΟ Β : Low Resolution HLA-A, HLA-B & HLA-DRB1**

Το σενάριο Β αναφέρεται στην διαχείριση των δεδομένων, όταν αυτά αναπαρίστανται σε χαμηλής ανάλυσης HLA τυποποίηση για τα HLA-A, HLA-B και HLA-DRB1 γονίδια. Η χαμηλής ανάλυσης HLA τυποποίηση των αλληλόμορφων αντιστοιχεί στην 2-digit αναπαράσταση (Επίπεδο Αντιγόνου).

### **ΣΕΝΑΡΙΟ Γ : Low Resolution HLA-A, HLA-B & High Resolution HLA-DRB1**

Το σενάριο Γ αναφέρεται στην διαχείριση των δεδομένων, όταν αυτά αναπαρίστανται σε χαμηλής ανάλυσης HLA τυποποίηση για τα HLA-A και HLA-B γονίδια και σε υψηλής ανάλυσης HLA τυποποίηση για το HLA-DRB1 γονίδιο.

#### **5.2.4 Δεδομένα Εισόδου**

Το πρώτο βήμα για την ανάλυση των δεδομένων είναι η προετοιμασία ενός αρχείου δεδομένων εισόδου σε κατάλληλη και επεξεργάσιμη από το Arlequin μορφή (Project File). Το Arlequin αποτελεί ένα αρκετά ευέλικτο πρόγραμμα και είναι σε θέση να αναλύσει διάφορους τύπους δεδομένων. Έτσι, θα πρέπει το αρχείο εισόδου να περιλαμβάνει ορισμένες πληροφορίες σχετικά με τις ιδιότητες των δεδομένων μαζί με τα ανεπεξέργαστα δεδομένα.

##### **5.2.4.1 Δημιουργία Αρχείου Arlequin (Arlequin Project File \*.arp)**

Το project file του Arlequin περιλαμβάνει περιγραφή των ιδιοτήτων των δεδομένων εισόδου, πλέον των ίδιων των πρωτογενών ανεπεξέργαστων δεδομένων εισόδου. Επίσης, το project file μπορεί να αναφέρεται σε ένα ή περισσότερα εξωτερικά αρχεία δεδομένων. Το περιεχόμενο του αρχείου εισόδου διαρθρώνεται σε δυο βασικές ενότητες με επιπλέον υποενότητες με τη σειρά που φαίνεται παρακάτω :

1. **Ενότητα Προφίλ (Profile)** - υποχρεωτική ενότητα
2. **Ενότητα Δεδομένων (Data)** - υποχρεωτική ενότητα
  - 2.α) Κατάλογος Απλοτύπων - προαιρετική υποενότητα
  - 2.β) Πίνακες Γενετικών Αποστάσεων - προαιρετική υποενότητα
  - 2.γ) **Δείγματα (Samples)** - υποχρεωτική υποενότητα
  - 2.δ) Γενετική Δομή - προαιρετική υποενότητα
  - 2.ε) Μεντελικοί Έλεγχοι - προαιρετική υποενότητα

### **Ενότητα Προφίλ (Profile)**

Η αρχή της ενότητας Profile υποδεικνύεται με τη λέξη-κλειδί [Profile].

Στην ενότητα αυτή ορίζουμε τις εξής μεταβλητές και τιμές παραμέτρων :

- 1. Title :** Μια ονομασία του τρέχοντος project.  
*Ορίζουμε τον τίτλο «HLA-A-B-DR Greeks».*
- 2. NbSamples :** Το πλήθος των πληθυσμών που αναλύονται.  
*Ορίζουμε την τιμή 1 γιατί αναλύεται δείγμα του Ελληνικού πληθυσμού μόνο.*
- 3. DataType :** Ο τύπος των δεδομένων που αναλύονται.  
*Ορίζουμε την τιμή STANDARD γιατί αναλύονται HLA δεδομένα.*
- 4. GenotypicData :** Το είδος των δεδομένων (0=Απλοτυπικά ή 1=Γονοτυπικά).  
*Ορίζουμε την τιμή 1 γιατί αναλύονται HLA γονοτυπικά δεδομένα .*
- 5. LocusSeparator :** Ο χαρακτήρας που διαχωρίζει αλληλόμορφα διαφορετικών γονιδιακών τόπων. *Ορίζουμε ως διαχωριστή τον χαρακτήρα κενού διαστήματος WHITESPACE.*
- 6. GameticPhase :** Γνώση γαμετικής φάσης γονοτύπων (0=Άγνωστη ή 1=Γνωστή). *Ορίζουμε την τιμή 0 γιατί η γαμετική φάση είναι άγνωστη.*
- 7. RecessiveData :** Παρουσία υπολοιπόμενων αλληλόμορφων στα γονοτυπικά δεδομένα (0=Κυρίαρχα ή 1=Υπολοιπόμενα).  
*Ορίζουμε την τιμή 1 γιατί περιλαμβάνονται υπολοιπόμενα αλληλόμορφα.*
- 8. RecessiveAllele :** Κώδικας αναπαράστασης υπολοιπόμενων αλληλόμορφων.  
*Ορίζουμε την τιμή «blank» ως αναπαράσταση υπολοιπόμενων αλληλόμορφων.*
- 9. MissingData :** Κώδικας αναπαράστασης ελλιπών δεδομένων.  
*Ορίζουμε τον χαρακτήρα «?» ως αναπαράσταση ελλιπών δεδομένων.*
- 10. FrequencyThreshold :** Πλήθος σημαντικών ψηφίων των αποτελεσμάτων της εκτίμησης των συχνοτήτων των απλοτύπων.  
*Ορίζουμε την τιμή  $10^{-5}$ .*
- 11. EpsilonValue :** Το κριτήριο σύγκλισης για τον τερματισμό της επαναληπτικής εκτέλεσης του αλγορίθμου EM, που χρησιμοποιείται για την εκτίμηση των συχνοτήτων των απλοτύπων και της ανισορροπίας σύνδεσης από τα γονοτυπικά δεδομένα.  
*Ορίζουμε την τιμή  $10^{-7}$ .*

### ***Ενότητα Δεδομένων (Data)***

Η ενότητα αυτή περιλαμβάνει τα πρωτογενή δεδομένα που θα αναλυθούν. Η αρχή της ενότητας Data υποδεικνύεται με τη λέξη-κλειδί [Data].

### ***Υποενότητα Δειγμάτων (Samples)***

Η αρχή της υποενότητας Samples υποδεικνύεται με τη λέξη-κλειδί [[Samples]].

Σε αυτήν την ενότητα ορίζεται το γονοτυπικό περιεχόμενο του δείγματος των δεδομένων του μητρώου του ΚΕΔΜΟΠ των 5.131 Ελλήνων εθελοντών δοτών μυελού των οστών που πρόκειται να αναλυθεί.

Στην υποενότητα αυτή ορίζουμε τις εξής μεταβλητές και τιμές παραμέτρων :

1. **SampleName** : Μια ονομασία του δείγματος.

*Ορίζουμε το όνομα «Greeks».*

2. **SampleSize** : Το μέγεθος του δείγματος. Για γονοτυπικά δεδομένα, το μέγεθος του δείγματος θα πρέπει να είναι ίσο με τον αριθμό των διπλοειδών ατόμων που περιλαμβάνονται στο δείγμα του πληθυσμού που αναλύεται.

*Ορίζουμε την τιμή 5.131 των Ελλήνων δοτών του μητρώου του ΚΕΔΜΟΠ.*

3. **SampleData** : Τα δεδομένα που αναλύονται. Ορίζεται η λίστα με τα γονοτυπικά δεδομένα με τη μορφή του αλληλομορφικού περιεχομένου κάθε ενός από τους τρεις γονιδιακούς τόπους (HLA-A,-B και -DRB1 που περιλαμβάνει το δείγμα (άγνωστη γαμετική φάση γονοτύπων).

#### **[Profile]**

```
Title="HLA-A-B-DR Greeks"
NbSamples=1
DataType=STANDARD
GenotypicData=1
LocusSeparator=WHITESPACE
GameticPhase=0
RecessiveData=1
RecessiveAllele="blank"
MissingData="?"
FrequencyThreshold=0.00001
EpsilonValue=0.000001
```

```
[Data]

[[Samples]]
SampleName="Greeks"
SampleSize=5131
SampleData={

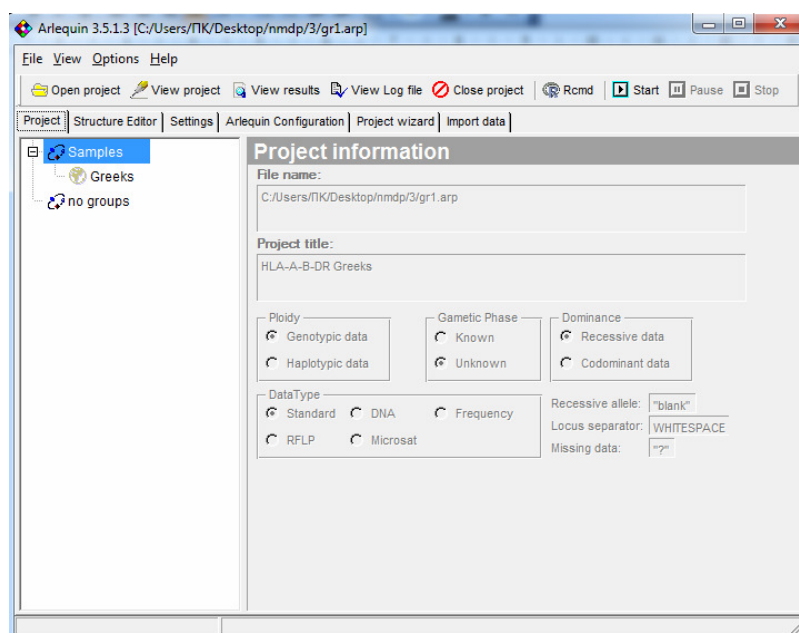
CBMDP-0000501  1      02:01  18:01  16:01
                02:01  51:01  11:04
CBMDP-0000502  1      01:01  08:01  03:01
                24:02  27:05  04:04
CBMDP-0000503  1      02:01  35:01  16:01
                30:01  51:01  04:03
CBMDP-0000504  1      01:01  44:02  15:02
                32:01  52:01  11:04
CBMDP-0000505  1      24:02  35:01  15:01
                32:01  44:02  16:02
CBMDP-0000506  1      01:01  08:01  03:01
                68:02  53:01  11:01
CBMDP-0000507  1      02:01  14:01  16:01
                26:01  39:06  07:01
CBMDP-0000508  1      11:01  35:01  07:01
                26:01  55:01  10:01
CBMDP-0000509  1      02:01  15:17  14:01
                68:01  18:01  07:01
CBMDP-0000510  1      33:01  15:17  11:04
                68:01  35:02  14:01

. . . . .
}

```

#### 5.2.4.2 Τιμές Υπολογιστικών Παραμέτρων

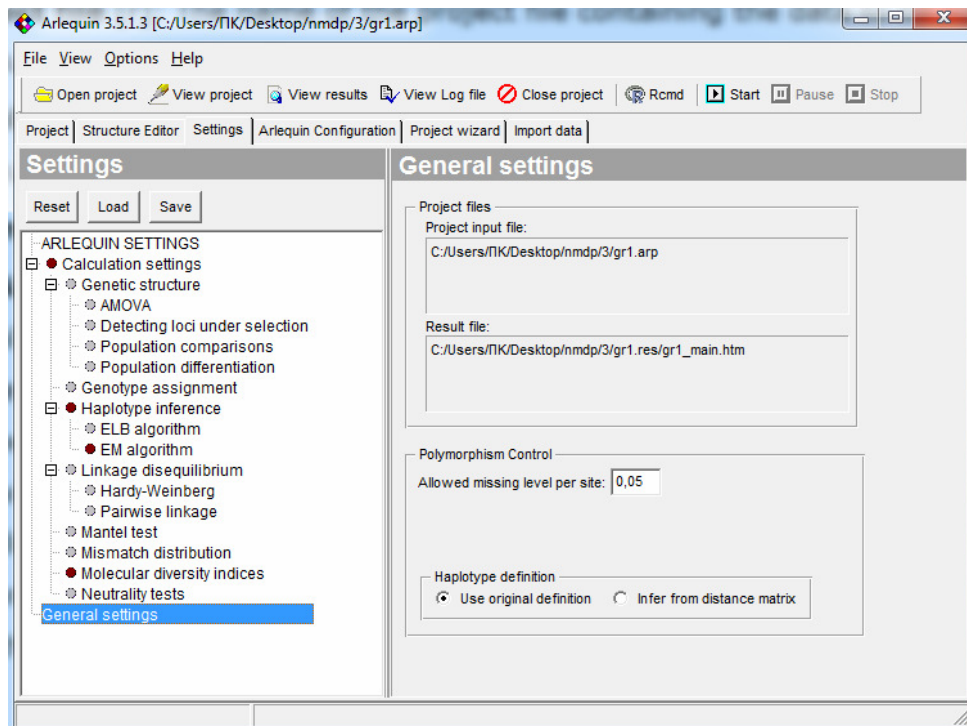
Αρχικά φορτώνουμε το Project File που δημιουργήσαμε στο Arlequin. Στην καρτέλα Project παρατηρούμε όλες τις σχετικές πληροφορίες που ορίσαμε στο αρχείο.



**ΕΙΚΟΝΑ 5.5** Η καρτέλα Project του Arlequin 3.5

Στη συνέχεια, στην καρτέλα Settings, στην επιλογή General Settings, στο πεδίο **Polymorphism Control** ορίζουμε το μέγιστο επιτρεπτό όριο ελλειπόν δεδομένων «**Allowed missing level per site**» για έναν γονιδιακό τόπο ώστε να μπορεί να συμπεριληφθεί με ασφάλεια στις αναλύσεις που θα υλοποιηθούν. Ορίζουμε την τιμή του επιπέδου στο 5%. Αυτό σημαίνει ότι ένας γονιδιακός τόπος με πάνω από 5% ελλιπή δεδομένα, δεν θα συμπεριληφθεί στις αναλύσεις που θα υλοποιηθούν. Αυτή η δυνατότητα είναι χρήσιμη περισσότερο για χειρισμό DNA αλληλουχιών, όταν πολλά άτομα έχουν αλληλουχηθεί σε ελαφρώς διαφορετικά τμήματα του γενετικού υλικού τους.

Στο πεδίο **Haplotype definition** επιλέγουμε την χρήση του πρωτότυπου ορισμού των απλοτύπων «**Use original definition**», έτσι ώστε οι απλότυποι να προσδιορίζονται σύμφωνα με το αρχικό τους αναγνωριστικό, χωρίς να ληφθεί υπόψη το γεγονός ότι ο μοριακός ορισμός τους θα μπορούσε να είναι πανομοιότυπος.



**ΕΙΚΟΝΑ 5.6** Η επιλογή General Settings της καρτέλας Settings του Arlequin 3.5

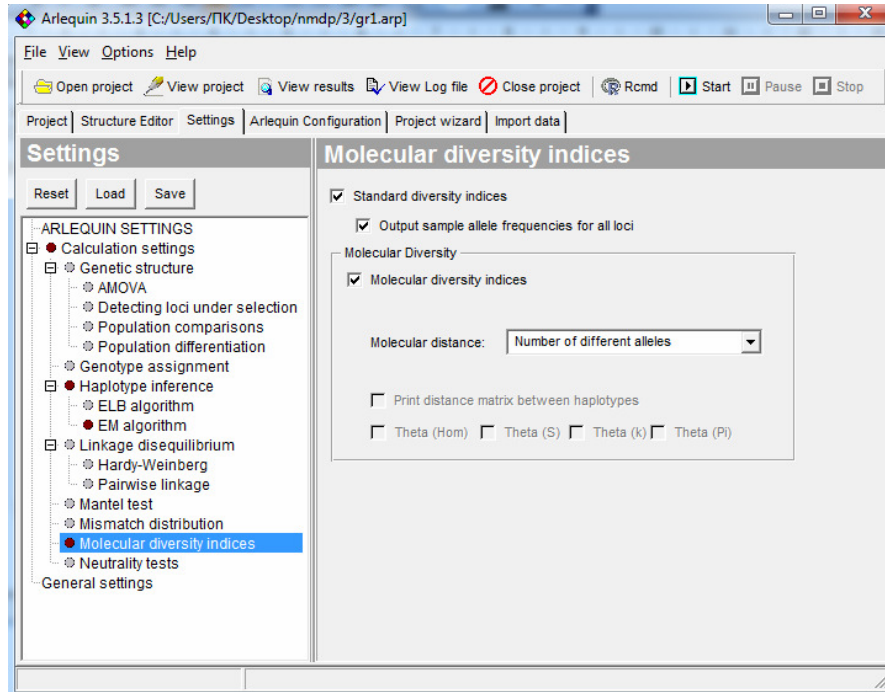
Στην επιλογή **Molecular diversity indices**,

τσεκάρουμε την επιλογή **Standard diversity indices**, ώστε να υπολογιστούν οι κοινοί δείκτες ποικιλομορφίας, όπως ο αριθμός των αλληλομόρφων, ο αριθμός των διαφορετικών γονιδιακών τόπων, το επίπεδο ετεροζυγωτίας κλπ.

τσεκάρουμε την επιλογή **Output sample allele frequencies for all loci**, ώστε να υπολογιστούν οι συχνότητες των αλληλόμορφων για όλους τους HLA γονιδιακούς τόπους του δείγματος.

Στο πεδίο **Molecular diversity**,

τσεκάρουμε την επιλογή **Molecular diversity indices**, ώστε να υπολογιστούν δείκτες ποικιλομορφίας και το εύρος της διαφορετικότητας σε μοριακό επίπεδο. Ως παράμετρος γενετικής απόστασης ορίζεται το πλήθος των διαφορετικών αλληλόμορφων.



**ΕΙΚΟΝΑ 5.7** Η επιλογή Molecular diversity indices της καρτέλας Settings του Arlequin 3.5

Στην επιλογή **Haplotype inference – EM algorithm**,

τσεκάρουμε την επιλογή **Use EM algorithm to estimate ML haplotype frequencies**, ώστε να υλοποιηθεί η εκτίμηση των συχνοτήτων μέγιστης πιθανότητας των απλοτύπων από τα παρατηρούμενα δεδομένα χρησιμοποιώντας τον αλγόριθμο EM για γονοτυπικά δεδομένα πολλαπλών γονιδιακών τόπων (HLA-A, -B και -DRB1) όταν δεν είναι γνωστή η γαμετική φάση, ή όταν παρουσιάζονται υπολειπόμενα αλληλόμορφα.

Στο πεδίο **EM algorithm settings**

τσεκάρουμε την επιλογή **Haplotype and locus levels**, ώστε να εκτιμηθούν οι συχνότητες των αλληλόμορφων για κάθε γονιδιακό τόπο και στη συνέχεια οι συχνότητες των απλοτύπων που προσδιορίζονται από τα αλληλόμορφα όλων των γονιδιακών τόπων.

Ορίζουμε ως **Epsilon value** την τιμή  $10^{-7}$ , η οποία αποτελεί το κατώφλι της παύσης της εκτέλεσης του αλγορίθμου EM (κριτήριο σύγκλισης). Μετά από κάθε επανάληψη, το Arlequin ελέγχει αν οι τρέχουσες συχνότητες απλοτύπων είναι διαφορετικές από εκείνες στην προηγούμενη επανάληψη. Εάν το άθροισμα της διαφοράς είναι μικρότερο από το κατώφλι, ο αλγόριθμος σταματά.

Ορίζουμε ως **Significant digits for output** την τιμή 5, η οποία αποτελεί την απαιτούμενη ακρίβεια του αποτελέσματος της εκτίμησης των συχνοτήτων των απλοτύπων.

Ορίζουμε ως **No. of starting points for EM algorithm** την τιμή 50, η οποία αποτελεί το πλήθος των τυχαίων αρχικών τιμών, από τις οποίες ο αλγόριθμος EM αρχίζει να εκτιμά κατ' επανάληψη τις συχνότητες των απλοτύπων. Οι συχνότητες των απλοτύπων, οι οποίες γενικά μεγιστοποιούν την πιθανότητα του δείγματος θα είναι αυτές που τελικά θα διατηρηθούν κατά την διαδοχή των επαναλήψεων της εκτέλεσης του αλγορίθμου. Συνήθως επιλέγεται πλήθος 50 ή περισσότερων τυχαίων αρχικών τιμών.

Ορίζουμε ως **Maximum number of iterations** την τιμή 5.000, η οποία αποτελεί το μέγιστο όριο στο πλήθος των επαναλήψεων που επιτρέπεται να εκτελεστούν στον αλγόριθμο EM. Η επαναληπτική διαδικασία θα έχει πλέον ένα ανώτατο όριο στο πλήθος των επαναλήψεων, αλλά μπορεί να σταματήσει νωρίτερα, αν η σύγκλιση έχει επιτευχθεί. Η σύγκλιση επιτυγχάνεται όταν το άθροισμα των διαφορών μεταξύ των συχνοτήτων των απλοτύπων μεταξύ δύο διαδοχικών επαναλήψεων είναι μικρότερη από την τιμή του κατωφλίου που ορίστηκε παραπάνω.

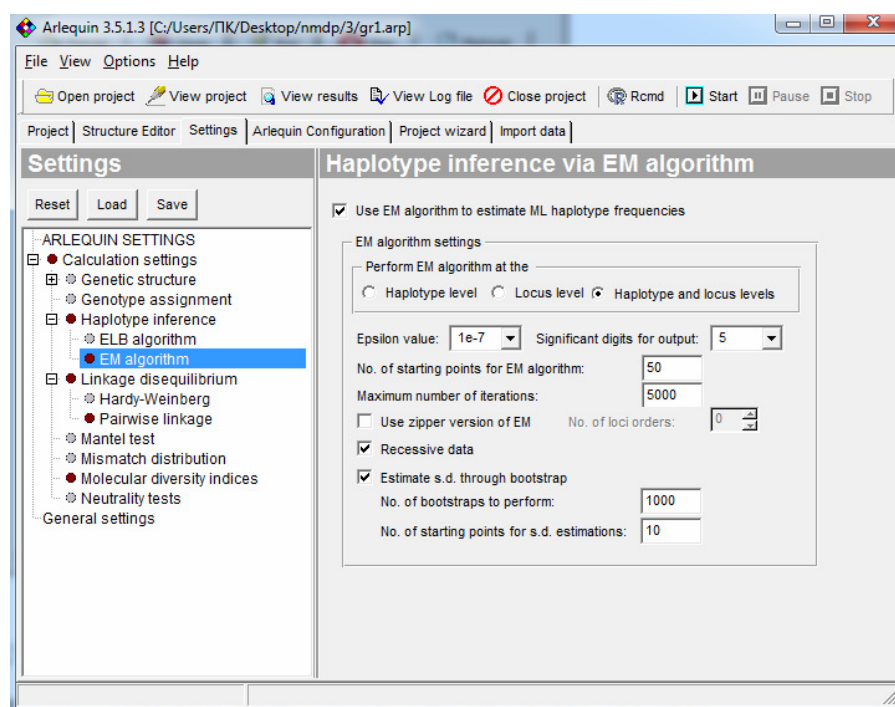
Τσεκάρουμε την επιλογή **Recessive data**, η οποία αναφέρεται στην παρουσία υπολοίπων αλληλόμορφων σε όλους τους γονιδιακούς τόπους, ώστε να ληφθεί υπόψη η παρουσία κενών αλληλόμορφων κατά τους υπολογισμούς του αλγορίθμου.

Τσεκάρουμε την επιλογή **Estimate s.d. through bootstrap**, η οποία αναφέρεται στην εκτίμηση της τυπικής απόκλισης των συχνοτήτων των απλοτύπων, μέσω μιας bootstrap προσέγγισης.

Ορίζουμε ως **No. of bootstraps to perform** την τιμή 1.000, η οποία αποτελεί το πλήθος των παραμετρικών αντιγράφων ενίσχυσης της διαδικασίας εκτίμησης των συχνοτήτων από τον αλγόριθμο EM σε τυχαία δείγματα που παράγονται από έναν

πλασματικό πληθυσμό, που έχει ίσες συχνότητες απλοτύπων με τις εκτιμήσεις των συχνοτήτων ML της προηγούμενης επανάληψης κατά την εκτέλεση του αλγορίθμου. Αυτή η διαδικασία χρησιμοποιείται για να εκτιμηθεί η τυπική απόκλιση των συχνοτήτων των απλοτύπων.

Ορίζουμε ως **No. of starting points for s.d. estimations** την τιμή 10, η οποία αποτελεί το πλήθος των αρχικών σημείων για τη διαδικασία εκτίμησης της τυπικής απόκλισης των συχνοτήτων των απλοτύπων. Η τιμή αυτή μπορεί να είναι μικρότερη από το αντίστοιχο πλήθος των αρχικών σημείων που ορίστηκε προηγουμένως για την εκτίμηση των συχνοτήτων των απλοτύπων, διότι οι bootstrap επαναλήψεις είναι αρκετά χρονοβόρες. Η ρύθμιση της τιμής αυτής διατηρείται σε χαμηλά επίπεδα.



**ΕΙΚΟΝΑ 5.8** Η επιλογή Haplotype inference - EM algorithm της καρτέλας Settings του Arlequin 3.5

Όταν δεν είναι γνωστή η γαμετική φάση, ο έλεγχος της ανισορροπίας της γενετικής σύνδεσης μεταξύ (LD) μεταξύ των ζευγών των γονιδιακών τόπων βασίζεται σε ένα τεστ λόγου πιθανοφάνειας, όπου η πιθανότητα του δείγματος που αξιολογήθηκε σύμφωνα με την υπόθεση της μη ύπαρξης σχέσης μεταξύ των γονιδιακών τόπων (ισορροπία σύνδεσης) συγκρίνεται με την πιθανότητα του δείγματος, όταν η σύνδεση επιτρέπεται (ανισορροπία σύνδεσης). Η σημασία της παρατηρούμενης αναλογίας πιθανότητας προσδιορίζεται με τον υπολογισμό της μηδενικής κατανομής αυτής της αναλογίας σύμφωνα με την υπόθεση της ισορροπίας σύνδεσης, χρησιμοποιώντας μία διαδικασία αντιμεταθέσεων.

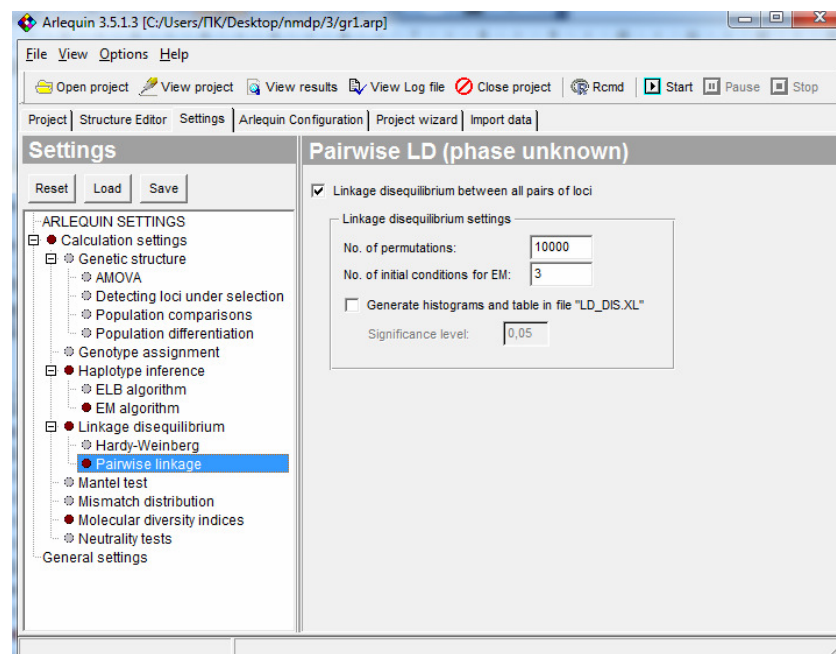
Στην επιλογή **Linkage disequilibrium - Pairwise linkage**,

τσεκάρουμε την επιλογή **Linkage disequilibrium between all pairs of loci**, ώστε να υλοποιηθεί ο έλεγχος ανισορροπίας σύνδεσης μεταξύ των ζευγών των γονιδιακών τόπων.

Στο πεδίο **Linkage disequilibrium settings**

Ορίζουμε ως **No. of permutations** την τιμή 10.000, η οποία αποτελεί το πλήθος των τυχαία αντιμετωπιζόμενων δειγμάτων που θα δημιουργηθούν. Συνήθως επιλέγεται μια τιμή αρκετών χιλιάδων τέτοιων δειγμάτων, με την τιμή των 16.000 παραλλαγών να εγγυάται να έχει μικρότερη από το 1% διαφορά με την ακριβή πιθανότητα στο 99% των περιπτώσεων. Επίσης, υπολογίζεται ένα τυπικό σφάλμα για την εκτιμώμενη P-τιμή χρησιμοποιώντας ένα σύστημα παρτίδων (Guo και Thomson, 1992).

Ορίζουμε ως **No. of initial conditions for EM** την τιμή 2, η οποία αποτελεί το πλήθος των τυχαίων αρχικών συνθηκών από τις οποίες ο αλγόριθμος EM έχει αρχίσει να εκτιμά κατ' επανάληψη την πιθανότητα του δείγματος. Οι συχνότητες των απλοτύπων, οι οποίες γενικά μεγιστοποιούν την πιθανότητα του δείγματος θα είναι αυτές που τελικά θα διατηρηθούν κατά την διαδοχή των επαναλήψεων της εκτέλεσης του αλγορίθμου. Συνήθως επιλέγεται πλήθος 3 ή περισσότερων τυχαίων αρχικών συνθηκών.



**ΕΙΚΟΝΑ 5.9** Η επιλογή Linkage disequilibrium - Pairwise linkage της καρτέλας Settings του Arlequin 3.5

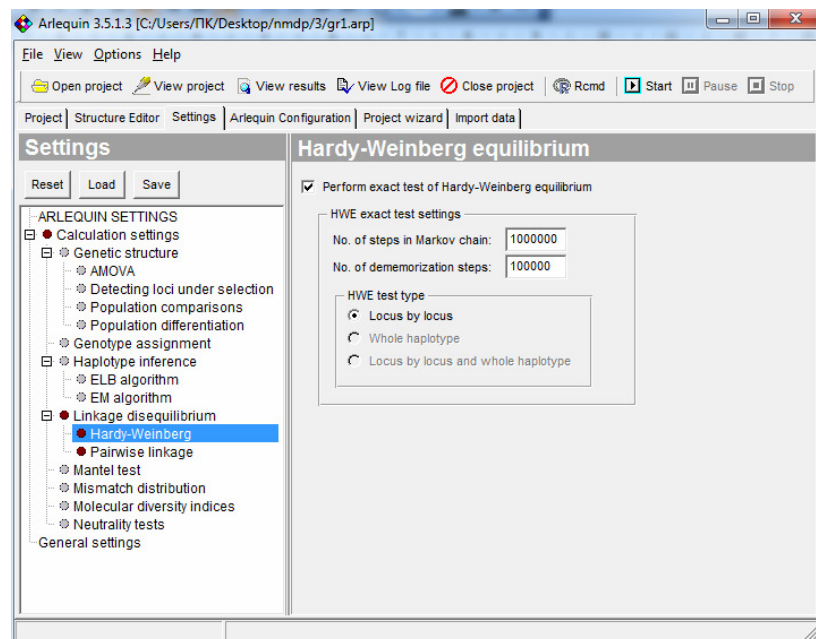
Στην επιλογή **Linkage disequilibrium – Hardy-Weinberg**,

τσεκάρουμε την επιλογή **Perform exact test of Hardy-Weinberg equilibrium**, ώστε να υλοποιηθεί ο έλεγχος της υπόθεσης ότι η παρατηρούμενη διπλοειδείς γονότυποι αποτελούν προϊόν μιας τυχαίας ένωσης των γαμετών. Ο έλεγχος αυτός είναι δυνατόν να υλοποιηθεί μόνο για γονοτυπικά δεδομένα. Ξεχωριστοί έλεγχοι HWE πραγματοποιούνται σε κάθε έναν από τους γονιδιακούς τόπους. Εάν η γαμετική φάση είναι άγνωστη, ο έλεγχος HWE είναι δυνατόν να υλοποιηθεί ανά γονιδιακό τόπο. Για δεδομένα με γνωστή γαμετική φάση, ο έλεγχος HWE είναι δυνατόν να υλοποιηθεί επίσης και σε επίπεδο απλοτύπων των ατόμων του δείγματος.

Στο πεδίο **HWE exact test settings**

Ορίζουμε ως **No. of steps in Markov chain** την τιμή 1.000.000, η οποία αποτελεί το μέγιστο πλήθος των εναλλακτικών πινάκων πιθανοτήτων προς διερεύνηση. Αριθμητικά, συνήθως επιλέγεται η τιμή 100.000 ή μεγαλύτερη.

Ορίζουμε ως **No. of dememorization steps** την τιμή 100.000, η οποία αποτελεί το πλήθος των βημάτων που θα εκτελεστούν πριν από την έναρξη της σύγκρισης των τιμών των εναλλακτικών πινάκων πιθανοτήτων με εκείνες του παρατηρούμενου πίνακα. Λίγες χιλιάδες βήματα είναι αναγκαία για την επίτευξη ενός τυχαίου σημείου εκκίνησης, που αντιστοιχεί σε έναν πίνακα πιθανοτήτων ανεξάρτητο από τον παρατηρούμενο πίνακα.



**ΕΙΚΟΝΑ 5.10** Η επιλογή Linkage disequilibrium – Hardy-Weinberg της καρτέλας Settings του Arlequin 3.5

### 5.2.5 Αποτελέσματα Βιοστατιστικής Ανάλυσης Δεδομένων

#### ΣΕΝΑΡΙΟ Α : High Resolution HLA-A, HLA-B & HLA-DRB1

##### 5.2.5.1 Πλήθος & Συχνότητες Αλληλόμορφων

Στα αποτελέσματα της στατιστικής ανάλυσης των δεδομένων εισόδου, παρατηρούνται οι τιμές του πλήθους των διαφορετικών αλληλόμορφων που προκύπτουν για τους γονιδιακούς τόπους HLA-A, HLA-B και HLA-DRB1 ξεχωριστά, καθώς επίσης και οι τιμές των συχνοτήτων κάθε αλληλόμορφου.

Οι τιμές του πλήθους των παρατηρούμενων (observed) αλληλόμορφων για κάθε γονιδιακό τόπο (συμπεριλαμβανομένων των κενών - blank) φαίνονται στον πίνακα Ζ.

**ΠΙΝΑΚΑΣ Ζ.** Αποτελέσματα πλήθους παρατηρούμενων αλληλόμορφων

| ΓΟΝΙΔΙΑΚΟΣ ΤΟΠΟΣ<br>(LOCUS) | ΠΛΗΘΟΣ<br>UNIQUE ΑΛΛΗΛΟΜΟΡΦΩΝ |
|-----------------------------|-------------------------------|
| HLA-A                       | 77                            |
| HLA-B                       | 125                           |
| HLA-DRB1                    | 73                            |

Όπως είναι φανερό, η μεγαλύτερη ποικιλομορφία αλληλομόρφων εμφανίζεται για το γονίδιο HLA-B στο σύνολο των δεδομένων του μητρώου, και η μικρότερη εμφανίζεται για το γονίδιο HLA-DRB1.

Οι τιμές των συχνοτήτων όλων των αλληλόμορφων των τριών γονιδιακών τόπων HLA-A, HLA-B και HLA-DRB1, όπως αυτές προέκυψαν από την ανάλυση του δείγματος φαίνονται στον πίνακα Η.

**ΠΙΝΑΚΑΣ Η.** Αποτελέσματα συχνοτήτων αλληλόμορφων

| HLA-A       |               | HLA-B       |               |             |               | HLA-DRB1    |               |
|-------------|---------------|-------------|---------------|-------------|---------------|-------------|---------------|
| ΑΛΛΗΛΟΜΟΡΦΟ | ΣΥΧΝΟΤΗΤΑ (%) | ΑΛΛΗΛΟΜΟΡΦΟ | ΣΥΧΝΟΤΗΤΑ (%) | ΑΛΛΗΛΟΜΟΡΦΟ | ΣΥΧΝΟΤΗΤΑ (%) | ΑΛΛΗΛΟΜΟΡΦΟ | ΣΥΧΝΟΤΗΤΑ (%) |
| 02:01       | 23,400%       | 51:01       | 12,442%       | 58:47       | 0,039%        | 11:04       | 15,676%       |
| 24:02       | 14,147%       | 18:01       | 10,644%       | 08:112      | 0,029%        | 16:01       | 12,876%       |
| 01:01       | 9,992%        | 35:01       | 7,816%        | 15:02       | 0,029%        | 11:01       | 8,500%        |
| 03:01       | 8,564%        | 08:01       | 4,531%        | 27:107      | 0,029%        | 03:01       | 7,529%        |
| 11:01       | 7,800%        | 35:03       | 4,290%        | 27:12       | 0,029%        | 07:01       | 6,375%        |
| 32:01       | 7,101%        | 44:02       | 4,025%        | 35:36       | 0,029%        | 14:01       | 5,332%        |
| 26:01       | 4,852%        | 07:02       | 3,109%        | 42:01       | 0,029%        | 01:01       | 4,789%        |
| 68:01       | 3,163%        | 55:01       | 2,951%        | 47:03       | 0,029%        | 15:01       | 4,713%        |
| 23:01       | 2,438%        | 13:02       | 2,923%        | 13:01       | 0,020%        | 13:01       | 4,201%        |
| 30:01       | 1,972%        | 49:01       | 2,812%        | 13:XX       | 0,020%        | 13:02       | 3,074%        |
| 31:01       | 1,589%        | 38:01       | 2,750%        | 14:XX       | 0,020%        | 04:03       | 2,675%        |
| 33:01       | 1,585%        | 52:01       | 2,692%        | 15:10       | 0,020%        | 15:02       | 2,272%        |
| blank       | 1,495%        | 35:02       | 2,637%        | 15:71       | 0,020%        | 04:05       | 2,207%        |
| 02:05       | 1,452%        | 39:01       | 2,594%        | 18:33       | 0,020%        | 16:02       | 2,105%        |
| 29:01       | 1,055%        | 35:08       | 2,383%        | 35:158      | 0,020%        | 10:01       | 1,900%        |
| 03:02       | 0,799%        | 14:02       | 2,162%        | 37:38       | 0,020%        | 04:02       | 1,827%        |
| 25:01       | 0,731%        | 40:02       | 2,026%        | 41:29       | 0,020%        | 01:02       | 1,724%        |
| 33:03       | 0,711%        | 44:03       | 1,831%        | 44:06       | 0,020%        | 12:01       | 1,495%        |
| 29:02       | 0,682%        | 58:01       | 1,479%        | 49:XX       | 0,020%        | 13:03       | 1,265%        |
| 30:04       | 0,624%        | 15:01       | 1,460%        | 50:34       | 0,020%        | 11:03       | 1,085%        |
| 30:02       | 0,604%        | 44:05       | 1,411%        | 54:01       | 0,020%        | blank       | 1,020%        |
| 02:415      | 0,526%        | 37:01       | 1,391%        | 55:02       | 0,020%        | 08:01       | 0,789%        |
| 69:01       | 0,507%        | 41:01       | 1,348%        | 78:XX       | 0,020%        | 04:01       | 0,760%        |
| 24:03       | 0,487%        | 50:01       | 1,339%        | 08:09       | 0,010%        | 04:04       | 0,743%        |
| 66:01       | 0,487%        | 27:05       | 1,277%        | 08:37       | 0,010%        | 04:07       | 0,555%        |
| 68:02       | 0,411%        | 57:01       | 1,277%        | 13:09       | 0,010%        | 14:04       | 0,518%        |
| 02:02       | 0,312%        | 40:01       | 1,218%        | 15:29       | 0,010%        | 08:04       | 0,448%        |
| 24:241      | 0,273%        | blank       | 1,152%        | 15:46       | 0,010%        | 11:148      | 0,429%        |
| 01:02       | 0,224%        | 27:02       | 0,968%        | 18:07       | 0,010%        | 13:05       | 0,322%        |
| 03:175      | 0,195%        | 15:17       | 0,863%        | 18:11       | 0,010%        | 09:01       | 0,313%        |
| 26:08       | 0,166%        | 07:05       | 0,692%        | 31:01       | 0,010%        | 16:XX       | 0,263%        |
| 29:10       | 0,156%        | 51:08       | 0,647%        | 35:05       | 0,010%        | 08:03       | 0,214%        |
| 02:06       | 0,156%        | 41:02       | 0,567%        | 35:11       | 0,010%        | 13:XX       | 0,214%        |
| 32:59       | 0,156%        | 47:01       | 0,507%        | 37:03       | 0,010%        | 15:99       | 0,205%        |
| 01:139      | 0,146%        | 53:01       | 0,497%        | 39:10       | 0,010%        | 03:85       | 0,175%        |
| 11:151      | 0,146%        | 40:06       | 0,469%        | 44:04       | 0,010%        | 04:164      | 0,156%        |
| 26:XX       | 0,088%        | 56:01       | 0,409%        | 44:17       | 0,010%        | 15:03       | 0,146%        |
| 74:01       | 0,088%        | 18:05       | 0,401%        | 49:03       | 0,010%        | 04:08       | 0,127%        |
| 02:11       | 0,049%        | 73:01       | 0,352%        | 49:11       | 0,010%        | 07:27       | 0,097%        |
| 02:85       | 0,049%        | 18:03       | 0,341%        | 50:02       | 0,010%        | 11:02       | 0,088%        |
| 68:103      | 0,049%        | 15:18       | 0,302%        | 51:04       | 0,010%        | 01:60       | 0,078%        |
| 02:17       | 0,039%        | 39:06       | 0,292%        | 51:65       | 0,010%        | 08:02       | 0,078%        |
| 23:XX       | 0,039%        | 35:230      | 0,283%        | 53:05       | 0,010%        | 13:21       | 0,068%        |
| 31:78       | 0,039%        | 39:24       | 0,224%        | 53:XX       | 0,010%        | 01:03       | 0,049%        |
| 02:08       | 0,029%        | 45:01       | 0,224%        | 56:05       | 0,010%        | 14:143      | 0,049%        |
| 24:17       | 0,029%        | 15:03       | 0,214%        | 56:40       | 0,010%        | 08:XX       | 0,039%        |
| 74:03       | 0,029%        | 51:XX       | 0,214%        | 59:XX       | 0,010%        | 04:06       | 0,029%        |
| 01:03       | 0,020%        | 15:08       | 0,166%        | 81:01       | 0,010%        | 11:15       | 0,029%        |
| 02:07       | 0,020%        | 27:07       | 0,166%        |             |               | 16:05       | 0,029%        |
| 02:20       | 0,020%        | 34:24       | 0,166%        |             |               | 04:11       | 0,020%        |
| 03:05       | 0,020%        | 51:07       | 0,156%        |             |               | 04:52       | 0,020%        |
| 25:XX       | 0,020%        | 57:03       | 0,156%        |             |               | 04:96       | 0,020%        |
| 30:XX       | 0,020%        | 14:01       | 0,146%        |             |               | 07:11       | 0,020%        |
| 33:07       | 0,020%        | 18:92       | 0,146%        |             |               | 08:10       | 0,020%        |
| 68:13       | 0,020%        | 39:31       | 0,146%        |             |               | 11:13       | 0,020%        |
| 01:42       | 0,010%        | 15:24       | 0,107%        |             |               | 11:39       | 0,020%        |
| 02:04       | 0,010%        | 18:18       | 0,107%        |             |               | 12:41       | 0,020%        |
| 02:12       | 0,010%        | 44:182      | 0,107%        |             |               | 13:11       | 0,020%        |
| 02:22       | 0,010%        | 07:199      | 0,097%        |             |               | 14:07       | 0,020%        |
| 02:40       | 0,010%        | 48:01       | 0,097%        |             |               | 14:17       | 0,020%        |
| 02:58       | 0,010%        | 39:XX       | 0,088%        |             |               | 15:06       | 0,020%        |
| 02:64       | 0,010%        | 40:XX       | 0,088%        |             |               | 04:10       | 0,010%        |
| 03:65       | 0,010%        | 44:29       | 0,088%        |             |               | 04:13       | 0,010%        |
| 11:131      | 0,010%        | 57:02       | 0,088%        |             |               | 11:06       | 0,010%        |
| 23:53       | 0,010%        | 27:03       | 0,068%        |             |               | 11:10       | 0,010%        |
| 24:07       | 0,010%        | 39:05       | 0,068%        |             |               | 11:12       | 0,010%        |
| 24:34       | 0,010%        | 55:60       | 0,068%        |             |               | 11:29       | 0,010%        |
| 24:52       | 0,010%        | 15:XX       | 0,059%        |             |               | 12:02       | 0,010%        |
| 26:32       | 0,010%        | 44:25       | 0,059%        |             |               | 13:104      | 0,010%        |
| 30:10       | 0,010%        | 51:05       | 0,059%        |             |               | 13:14       | 0,010%        |
| 30:29       | 0,010%        | 15:16       | 0,049%        |             |               | 13:88       | 0,010%        |
| 33:72       | 0,010%        | 52:32       | 0,049%        |             |               | 14:05       | 0,010%        |
| 34:02       | 0,010%        | 07:10       | 0,039%        |             |               | 14:16       | 0,010%        |
| 66:XX       | 0,010%        | 15:09       | 0,039%        |             |               |             |               |
| 68:24       | 0,010%        | 27:09       | 0,039%        |             |               |             |               |
| 68:55       | 0,010%        | 38:XX       | 0,039%        |             |               |             |               |
| 80:01       | 0,010%        | 57:66       | 0,039%        |             |               |             |               |

### Αθροιστικές Συχνότητες

Πολλές φορές μας ενδιαφέρει να μάθουμε το πλήθος ή το ποσοστό των παρατηρήσεων, των οποίων οι τιμές είναι μικρότερες ή ίσες μιας ορισμένης τιμής  $x_j$  μιας ποσοτικής μεταβλητής που ορίζεται ως  $X$ . Για παράδειγμα αν θέλουμε να απαντήσουμε στο ερώτημα «Ποιο ποσοστό του γονιδιακού τύπου HLA-A έχει ποικιλομορφία το πολύ 2 συχνότερων αλληλόμορφων», υπολογίζουμε τη λεγόμενη αθροιστική συχνότητα

$$N_j = \sum_{i=1}^k n_i = n_1 + n_2 + \dots + n_k, 1 \leq j \leq k \quad (22)$$

και τη σχετική αθροιστική συχνότητα

$$F_j = \sum_{i=1}^k f_i = f_1 + f_2 + \dots + f_k, 1 \leq j \leq k \quad (23)$$

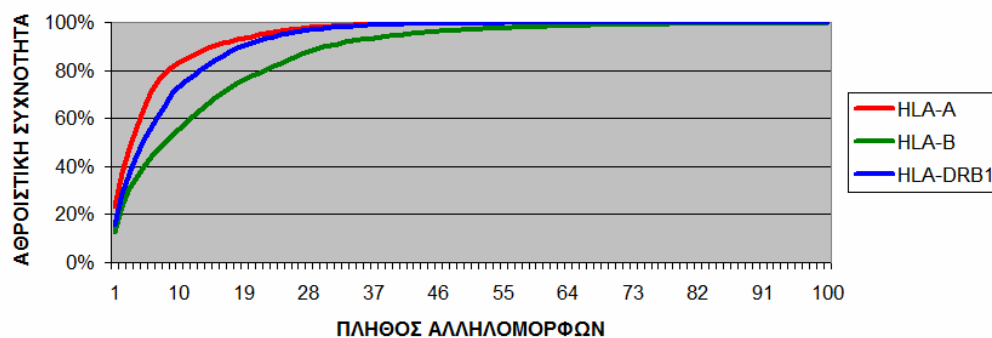
με τις τιμές  $x_1, x_2, \dots, x_k$  της ποσοτικής μεταβλητής  $X$  διατεταγμένες σε αύξουσα τάξη. Η επί τοις εκατό σχετική αθροιστική συχνότητα ισούται με

$$F_j \% = 100F_j \quad (24)$$

**ΠΙΝΑΚΑΣ Θ.** Αθροιστικές Συχνότητες αλληλόμορφων

| $j$ | $F_j$   |         |          |
|-----|---------|---------|----------|
|     | HLA-A   | HLA-B   | HLA-DRB1 |
| 1   | 23,400% | 12,442% | 15,676%  |
| 2   | 37,547% | 23,086% | 28,551%  |
| 3   | 47,539% | 30,901% | 37,051%  |
| 4   | 56,103% | 35,432% | 44,580%  |
| 5   | 63,903% | 39,722% | 50,954%  |
| 6   | 71,004% | 43,747% | 56,286%  |
| 7   | 75,856% | 46,856% | 61,075%  |
| 8   | 79,018% | 49,806% | 65,788%  |
| 9   | 81,456% | 52,729% | 69,989%  |
| 10  | 83,428% | 55,541% | 73,063%  |
| 11  | 85,017% | 58,291% | 75,738%  |
| 12  | 86,602% | 60,982% | 78,009%  |
| 13  | 88,097% | 63,619% | 80,217%  |
| 14  | 89,549% | 66,213% | 82,322%  |
| 15  | 90,604% | 68,596% | 84,222%  |
| 16  | 91,403% | 70,758% | 86,049%  |
| 17  | 92,134% | 72,784% | 87,773%  |
| 18  | 92,846% | 74,615% | 89,268%  |
| 19  | 93,528% | 76,094% | 90,532%  |
| 20  | 94,151% | 77,554% | 91,617%  |

**ΑΘΡΟΙΣΤΙΚΕΣ ΣΥΧΝΟΤΗΤΕΣ ΑΛΛΗΛΟΜΟΡΦΩΝ ΓΟΝΙΔΙΩΝ**



**ΕΙΚΟΝΑ 5.11** Γραφική παράσταση αθροιστικών συχνοτήτων αλληλόμορφων γονιδίων

## ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΑΛΛΗΛΟΜΟΡΦΩΝ ARLEQUIN 3.5

```
Arlequin results
////////////////////////////////////////////////////////////////
RUN NUMBER 1 (08/10/14 at 19:36:19)
////////////////////////////////////////////////////////////////
Project information:
-----
      NbSamples      = 1
      DataType       = STANDARD
      GenotypicData  = 1
      GameticPhase   = 0
      RecessiveData  = 1

=====
Settings used for Calculations
=====
General settings:
-----
      Deletion Weight           = 1
      Transition Weight Weight  = 1
      Tranversion Weight Weight = 1
      Epsilon Value             = 1e-07
      Significant digits for output = 5
      Use original haplotype definition
      Allowed level of missing data = 0.05

Active Tasks:
-----

Standard indices:
-----

Molecular Diversity:
-----
      Molecular Distance :Pairwise differences
      GammaA Value      = 0
      Theta estimators  :

Haplotypic Frequency estimation:
-----
      Make estimations at the:
      Haplotypic level
      Locus level
      Initial Conditions           = 50
      Maximum No. Of Iterations   = 5000
      Bootstrap Replicates        = 1000
      Initial Conditions for Bootstrap = 10
      Recessive Data
      Do not Compact Haplotypes
```



|    |         |          |          |        |
|----|---------|----------|----------|--------|
| 24 | UNKNOWN | 0.007991 | 0.000885 | 03:02  |
| 25 | UNKNOWN | 0.000195 | 0.000139 | 03:05  |
| 26 | UNKNOWN | 0.001949 | 0.000431 | 03:175 |
| 27 | UNKNOWN | 0.000097 | 0.000097 | 03:65  |
| 28 | UNKNOWN | 0.077997 | 0.002661 | 11:01  |
| 29 | UNKNOWN | 0.000097 | 0.000096 | 11:131 |
| 30 | UNKNOWN | 0.001462 | 0.000377 | 11:151 |
| 31 | UNKNOWN | 0.024376 | 0.001492 | 23:01  |
| 32 | UNKNOWN | 0.000097 | 0.000097 | 23:53  |
| 33 | UNKNOWN | 0.000390 | 0.000201 | 23:XX  |
| 34 | UNKNOWN | 0.141470 | 0.003334 | 24:02  |
| 35 | UNKNOWN | 0.004872 | 0.000696 | 24:03  |
| 36 | UNKNOWN | 0.000097 | 0.000098 | 24:07  |
| 37 | UNKNOWN | 0.000292 | 0.000166 | 24:17  |
| 38 | UNKNOWN | 0.002729 | 0.000515 | 24:241 |
| 39 | UNKNOWN | 0.000097 | 0.000095 | 24:34  |
| 40 | UNKNOWN | 0.000097 | 0.000093 | 24:52  |
| 41 | UNKNOWN | 0.007309 | 0.000817 | 25:01  |
| 42 | UNKNOWN | 0.000195 | 0.000140 | 25:XX  |
| 43 | UNKNOWN | 0.048519 | 0.002123 | 26:01  |
| 44 | UNKNOWN | 0.001662 | 0.000418 | 26:08  |
| 45 | UNKNOWN | 0.000097 | 0.000096 | 26:32  |
| 46 | UNKNOWN | 0.000877 | 0.000293 | 26:XX  |
| 47 | UNKNOWN | 0.010550 | 0.001023 | 29:01  |
| 48 | UNKNOWN | 0.006821 | 0.000792 | 29:02  |
| 49 | UNKNOWN | 0.001564 | 0.000414 | 29:10  |
| 50 | UNKNOWN | 0.019722 | 0.001376 | 30:01  |
| 51 | UNKNOWN | 0.006042 | 0.000771 | 30:02  |
| 52 | UNKNOWN | 0.006237 | 0.000767 | 30:04  |
| 53 | UNKNOWN | 0.000097 | 0.000101 | 30:10  |
| 54 | UNKNOWN | 0.000097 | 0.000098 | 30:29  |
| 55 | UNKNOWN | 0.000195 | 0.000143 | 30:XX  |
| 56 | UNKNOWN | 0.015888 | 0.001232 | 31:01  |
| 57 | UNKNOWN | 0.000390 | 0.000196 | 31:78  |
| 58 | UNKNOWN | 0.071010 | 0.002503 | 32:01  |
| 59 | UNKNOWN | 0.001559 | 0.000385 | 32:59  |
| 60 | UNKNOWN | 0.015854 | 0.001203 | 33:01  |
| 61 | UNKNOWN | 0.007114 | 0.000812 | 33:03  |
| 62 | UNKNOWN | 0.000195 | 0.000135 | 33:07  |
| 63 | UNKNOWN | 0.000097 | 0.000098 | 33:72  |
| 64 | UNKNOWN | 0.000097 | 0.000097 | 34:02  |
| 65 | UNKNOWN | 0.004872 | 0.000705 | 66:01  |
| 66 | UNKNOWN | 0.000097 | 0.000095 | 66:XX  |
| 67 | UNKNOWN | 0.031626 | 0.001734 | 68:01  |
| 68 | UNKNOWN | 0.004105 | 0.000656 | 68:02  |
| 69 | UNKNOWN | 0.000487 | 0.000225 | 68:103 |
| 70 | UNKNOWN | 0.000195 | 0.000138 | 68:13  |
| 71 | UNKNOWN | 0.000097 | 0.000094 | 68:24  |
| 72 | UNKNOWN | 0.000097 | 0.000097 | 68:55  |
| 73 | UNKNOWN | 0.005067 | 0.000680 | 69:01  |
| 74 | UNKNOWN | 0.000877 | 0.000283 | 74:01  |
| 75 | UNKNOWN | 0.000292 | 0.000172 | 74:03  |
| 76 | UNKNOWN | 0.000097 | 0.000096 | 80:01  |

|                                                |           |               |          |        |
|------------------------------------------------|-----------|---------------|----------|--------|
| 77                                             | UNKNOWN   | 0.014949      | 0.002150 | blank  |
| -----                                          |           |               |          |        |
| Sum of all 77 haplotype frequencies : 1.000000 |           |               |          |        |
| Sum of 77 listed frequencies : 1.000000        |           |               |          |        |
| Allele frequencies for the locus 2             |           |               |          |        |
| -----                                          |           |               |          |        |
| No. of gene copies in sample                   |           | : 10262       |          |        |
| No. of random initial conditions for EM        |           | : 50          |          |        |
| No. of bootstrap replicates                    |           | : 1000        |          |        |
| No. of different maximum likelihoods found     |           | : 1           |          |        |
| Epsilon value for stopping iterations          |           | : 1.00000e-07 |          |        |
| Logarithm of the sample maximum-likelihood     |           | : -32484.0    |          |        |
| -----                                          |           |               |          |        |
| Maximum-likelihood haplotype frequencies :     |           |               |          |        |
| -----                                          |           |               |          |        |
| Total number of possible haplotypes            |           | :             | 125      |        |
| Minimum frequency to reach for output          |           | :             | 1.00e-05 |        |
| -----                                          |           |               |          |        |
| #                                              | Haplotype | Freq.         | s.d.     |        |
| -----                                          |           |               |          |        |
| 1                                              | UNKNOWN   | 0.031087      | 0.001638 | 07:02  |
| 2                                              | UNKNOWN   | 0.006919      | 0.000812 | 07:05  |
| 3                                              | UNKNOWN   | 0.000390      | 0.000197 | 07:10  |
| 4                                              | UNKNOWN   | 0.000974      | 0.000307 | 07:199 |
| 5                                              | UNKNOWN   | 0.045308      | 0.002080 | 08:01  |
| 6                                              | UNKNOWN   | 0.000097      | 0.000103 | 08:09  |
| 7                                              | UNKNOWN   | 0.000292      | 0.000170 | 08:112 |
| 8                                              | UNKNOWN   | 0.000097      | 0.000098 | 08:37  |
| 9                                              | UNKNOWN   | 0.000195      | 0.000141 | 13:01  |
| 10                                             | UNKNOWN   | 0.029226      | 0.001649 | 13:02  |
| 11                                             | UNKNOWN   | 0.000097      | 0.000096 | 13:09  |
| 12                                             | UNKNOWN   | 0.000195      | 0.000134 | 13:XX  |
| 13                                             | UNKNOWN   | 0.001462      | 0.000365 | 14:01  |
| 14                                             | UNKNOWN   | 0.021624      | 0.001431 | 14:02  |
| 15                                             | UNKNOWN   | 0.000195      | 0.000137 | 14:XX  |
| 16                                             | UNKNOWN   | 0.014595      | 0.001225 | 15:01  |
| 17                                             | UNKNOWN   | 0.000292      | 0.000173 | 15:02  |
| 18                                             | UNKNOWN   | 0.002144      | 0.000458 | 15:03  |
| 19                                             | UNKNOWN   | 0.001657      | 0.000415 | 15:08  |
| 20                                             | UNKNOWN   | 0.000390      | 0.000189 | 15:09  |
| 21                                             | UNKNOWN   | 0.000195      | 0.000138 | 15:10  |
| 22                                             | UNKNOWN   | 0.000487      | 0.000212 | 15:16  |
| 23                                             | UNKNOWN   | 0.008628      | 0.000888 | 15:17  |
| 24                                             | UNKNOWN   | 0.003021      | 0.000540 | 15:18  |
| 25                                             | UNKNOWN   | 0.001072      | 0.000329 | 15:24  |
| 26                                             | UNKNOWN   | 0.000097      | 0.000100 | 15:29  |
| 27                                             | UNKNOWN   | 0.000097      | 0.000100 | 15:46  |

|    |         |          |          |        |
|----|---------|----------|----------|--------|
| 28 | UNKNOWN | 0.000195 | 0.000133 | 15:71  |
| 29 | UNKNOWN | 0.000585 | 0.000235 | 15:XX  |
| 30 | UNKNOWN | 0.106437 | 0.002966 | 18:01  |
| 31 | UNKNOWN | 0.003411 | 0.000557 | 18:03  |
| 32 | UNKNOWN | 0.004010 | 0.000624 | 18:05  |
| 33 | UNKNOWN | 0.000097 | 0.000101 | 18:07  |
| 34 | UNKNOWN | 0.000097 | 0.000103 | 18:11  |
| 35 | UNKNOWN | 0.001072 | 0.000313 | 18:18  |
| 36 | UNKNOWN | 0.000195 | 0.000141 | 18:33  |
| 37 | UNKNOWN | 0.001462 | 0.000367 | 18:92  |
| 38 | UNKNOWN | 0.009676 | 0.000963 | 27:02  |
| 39 | UNKNOWN | 0.000682 | 0.000263 | 27:03  |
| 40 | UNKNOWN | 0.012766 | 0.001117 | 27:05  |
| 41 | UNKNOWN | 0.001657 | 0.000401 | 27:07  |
| 42 | UNKNOWN | 0.000390 | 0.000190 | 27:09  |
| 43 | UNKNOWN | 0.000292 | 0.000168 | 27:107 |
| 44 | UNKNOWN | 0.000292 | 0.000166 | 27:12  |
| 45 | UNKNOWN | 0.000097 | 0.000096 | 31:01  |
| 46 | UNKNOWN | 0.001657 | 0.000409 | 34:24  |
| 47 | UNKNOWN | 0.078155 | 0.002640 | 35:01  |
| 48 | UNKNOWN | 0.026369 | 0.001552 | 35:02  |
| 49 | UNKNOWN | 0.042897 | 0.001899 | 35:03  |
| 50 | UNKNOWN | 0.000097 | 0.000096 | 35:05  |
| 51 | UNKNOWN | 0.023827 | 0.001467 | 35:08  |
| 52 | UNKNOWN | 0.000097 | 0.000103 | 35:11  |
| 53 | UNKNOWN | 0.000195 | 0.000134 | 35:158 |
| 54 | UNKNOWN | 0.002826 | 0.000544 | 35:230 |
| 55 | UNKNOWN | 0.000292 | 0.000172 | 35:36  |
| 56 | UNKNOWN | 0.013911 | 0.001178 | 37:01  |
| 57 | UNKNOWN | 0.000097 | 0.000098 | 37:03  |
| 58 | UNKNOWN | 0.000195 | 0.000136 | 37:38  |
| 59 | UNKNOWN | 0.027497 | 0.001551 | 38:01  |
| 60 | UNKNOWN | 0.000390 | 0.000191 | 38:XX  |
| 61 | UNKNOWN | 0.025938 | 0.001615 | 39:01  |
| 62 | UNKNOWN | 0.000682 | 0.000260 | 39:05  |
| 63 | UNKNOWN | 0.002923 | 0.000531 | 39:06  |
| 64 | UNKNOWN | 0.000097 | 0.000098 | 39:10  |
| 65 | UNKNOWN | 0.002241 | 0.000467 | 39:24  |
| 66 | UNKNOWN | 0.001462 | 0.000369 | 39:31  |
| 67 | UNKNOWN | 0.000877 | 0.000295 | 39:XX  |
| 68 | UNKNOWN | 0.012181 | 0.001040 | 40:01  |
| 69 | UNKNOWN | 0.020263 | 0.001424 | 40:02  |
| 70 | UNKNOWN | 0.004694 | 0.000680 | 40:06  |
| 71 | UNKNOWN | 0.000877 | 0.000288 | 40:XX  |
| 72 | UNKNOWN | 0.013484 | 0.001119 | 41:01  |
| 73 | UNKNOWN | 0.005671 | 0.000704 | 41:02  |
| 74 | UNKNOWN | 0.000195 | 0.000140 | 41:29  |
| 75 | UNKNOWN | 0.000292 | 0.000165 | 42:01  |
| 76 | UNKNOWN | 0.040254 | 0.001954 | 44:02  |
| 77 | UNKNOWN | 0.018309 | 0.001268 | 44:03  |
| 78 | UNKNOWN | 0.000097 | 0.000095 | 44:04  |
| 79 | UNKNOWN | 0.014106 | 0.001175 | 44:05  |
| 80 | UNKNOWN | 0.000195 | 0.000140 | 44:06  |

|                                                 |         |          |          |        |
|-------------------------------------------------|---------|----------|----------|--------|
| 81                                              | UNKNOWN | 0.000097 | 0.000102 | 44:17  |
| 82                                              | UNKNOWN | 0.001072 | 0.000336 | 44:182 |
| 83                                              | UNKNOWN | 0.000585 | 0.000237 | 44:25  |
| 84                                              | UNKNOWN | 0.000877 | 0.000284 | 44:29  |
| 85                                              | UNKNOWN | 0.002241 | 0.000460 | 45:01  |
| 86                                              | UNKNOWN | 0.005067 | 0.000693 | 47:01  |
| 87                                              | UNKNOWN | 0.000292 | 0.000175 | 47:03  |
| 88                                              | UNKNOWN | 0.000974 | 0.000313 | 48:01  |
| 89                                              | UNKNOWN | 0.028118 | 0.001677 | 49:01  |
| 90                                              | UNKNOWN | 0.000097 | 0.000097 | 49:03  |
| 91                                              | UNKNOWN | 0.000097 | 0.000092 | 49:11  |
| 92                                              | UNKNOWN | 0.000195 | 0.000139 | 49:XX  |
| 93                                              | UNKNOWN | 0.013386 | 0.001139 | 50:01  |
| 94                                              | UNKNOWN | 0.000097 | 0.000099 | 50:02  |
| 95                                              | UNKNOWN | 0.000195 | 0.000139 | 50:34  |
| 96                                              | UNKNOWN | 0.124418 | 0.003305 | 51:01  |
| 97                                              | UNKNOWN | 0.000097 | 0.000097 | 51:04  |
| 98                                              | UNKNOWN | 0.000585 | 0.000239 | 51:05  |
| 99                                              | UNKNOWN | 0.001559 | 0.000402 | 51:07  |
| 100                                             | UNKNOWN | 0.006474 | 0.000824 | 51:08  |
| 101                                             | UNKNOWN | 0.000097 | 0.000096 | 51:65  |
| 102                                             | UNKNOWN | 0.002144 | 0.000469 | 51:XX  |
| 103                                             | UNKNOWN | 0.026918 | 0.001562 | 52:01  |
| 104                                             | UNKNOWN | 0.000487 | 0.000223 | 52:32  |
| 105                                             | UNKNOWN | 0.004970 | 0.000716 | 53:01  |
| 106                                             | UNKNOWN | 0.000097 | 0.000098 | 53:05  |
| 107                                             | UNKNOWN | 0.000097 | 0.000101 | 53:XX  |
| 108                                             | UNKNOWN | 0.000195 | 0.000139 | 54:01  |
| 109                                             | UNKNOWN | 0.029508 | 0.001646 | 55:01  |
| 110                                             | UNKNOWN | 0.000195 | 0.000137 | 55:02  |
| 111                                             | UNKNOWN | 0.000682 | 0.000256 | 55:60  |
| 112                                             | UNKNOWN | 0.004093 | 0.000645 | 56:01  |
| 113                                             | UNKNOWN | 0.000097 | 0.000096 | 56:05  |
| 114                                             | UNKNOWN | 0.000097 | 0.000096 | 56:40  |
| 115                                             | UNKNOWN | 0.012766 | 0.001090 | 57:01  |
| 116                                             | UNKNOWN | 0.000877 | 0.000290 | 57:02  |
| 117                                             | UNKNOWN | 0.001559 | 0.000388 | 57:03  |
| 118                                             | UNKNOWN | 0.000390 | 0.000194 | 57:66  |
| 119                                             | UNKNOWN | 0.014791 | 0.001177 | 58:01  |
| 120                                             | UNKNOWN | 0.000390 | 0.000192 | 58:47  |
| 121                                             | UNKNOWN | 0.000097 | 0.000103 | 59:XX  |
| 122                                             | UNKNOWN | 0.003521 | 0.000571 | 73:01  |
| 123                                             | UNKNOWN | 0.000195 | 0.000138 | 78:XX  |
| 124                                             | UNKNOWN | 0.000097 | 0.000102 | 81:01  |
| 125                                             | UNKNOWN | 0.011523 | 0.001628 | blank  |
| -----                                           |         |          |          |        |
| Sum of all 125 haplotype frequencies : 1.000000 |         |          |          |        |
| Sum of 125 listed frequencies : 1.000000        |         |          |          |        |

# Allele frequencies for the locus 3

```

-----
No. of gene copies in sample           : 10262
No. of random initial conditions for EM : 50
No. of bootstrap replicates            : 1000
No. of different maximum likelihoods found : 1
Epsilon value for stopping iterations   : 1.00000e-07
Logarithm of the sample maximum-likelihood : -27518.7

```

## Maximum-likelihood haplotype frequencies :

```

-----
Total number of possible haplotypes      :          73
Minimum frequency to reach for output    :      1.00e-05

```

| #  | Haplotype | Freq.    | s.d.     |        |
|----|-----------|----------|----------|--------|
| 1  | UNKNOWN   | 0.047887 | 0.002153 | 01:01  |
| 2  | UNKNOWN   | 0.017240 | 0.001290 | 01:02  |
| 3  | UNKNOWN   | 0.000487 | 0.000213 | 01:03  |
| 4  | UNKNOWN   | 0.000780 | 0.000277 | 01:60  |
| 5  | UNKNOWN   | 0.075288 | 0.002587 | 03:01  |
| 6  | UNKNOWN   | 0.001754 | 0.000429 | 03:85  |
| 7  | UNKNOWN   | 0.007601 | 0.000843 | 04:01  |
| 8  | UNKNOWN   | 0.018269 | 0.001305 | 04:02  |
| 9  | UNKNOWN   | 0.026753 | 0.001599 | 04:03  |
| 10 | UNKNOWN   | 0.007432 | 0.000801 | 04:04  |
| 11 | UNKNOWN   | 0.022074 | 0.001459 | 04:05  |
| 12 | UNKNOWN   | 0.000292 | 0.000175 | 04:06  |
| 13 | UNKNOWN   | 0.005554 | 0.000736 | 04:07  |
| 14 | UNKNOWN   | 0.001267 | 0.000354 | 04:08  |
| 15 | UNKNOWN   | 0.000097 | 0.000100 | 04:10  |
| 16 | UNKNOWN   | 0.000195 | 0.000141 | 04:11  |
| 17 | UNKNOWN   | 0.000097 | 0.000101 | 04:13  |
| 18 | UNKNOWN   | 0.001559 | 0.000395 | 04:164 |
| 19 | UNKNOWN   | 0.000195 | 0.000139 | 04:52  |
| 20 | UNKNOWN   | 0.000195 | 0.000146 | 04:96  |
| 21 | UNKNOWN   | 0.063745 | 0.002305 | 07:01  |
| 22 | UNKNOWN   | 0.000195 | 0.000137 | 07:11  |
| 23 | UNKNOWN   | 0.000974 | 0.000312 | 07:27  |
| 24 | UNKNOWN   | 0.007893 | 0.000858 | 08:01  |
| 25 | UNKNOWN   | 0.000780 | 0.000276 | 08:02  |
| 26 | UNKNOWN   | 0.002144 | 0.000450 | 08:03  |
| 27 | UNKNOWN   | 0.004483 | 0.000644 | 08:04  |
| 28 | UNKNOWN   | 0.000195 | 0.000137 | 08:10  |
| 29 | UNKNOWN   | 0.000390 | 0.000197 | 08:XX  |
| 30 | UNKNOWN   | 0.003131 | 0.000547 | 09:01  |
| 31 | UNKNOWN   | 0.018999 | 0.001286 | 10:01  |
| 32 | UNKNOWN   | 0.084996 | 0.002663 | 11:01  |
| 33 | UNKNOWN   | 0.000877 | 0.000284 | 11:02  |

|                                                        |         |          |          |        |
|--------------------------------------------------------|---------|----------|----------|--------|
| 34                                                     | UNKNOWN | 0.010850 | 0.001029 | 11:03  |
| 35                                                     | UNKNOWN | 0.156757 | 0.003640 | 11:04  |
| 36                                                     | UNKNOWN | 0.000097 | 0.000097 | 11:06  |
| 37                                                     | UNKNOWN | 0.000097 | 0.000096 | 11:10  |
| 38                                                     | UNKNOWN | 0.000097 | 0.000102 | 11:12  |
| 39                                                     | UNKNOWN | 0.000195 | 0.000137 | 11:13  |
| 40                                                     | UNKNOWN | 0.004288 | 0.000625 | 11:148 |
| 41                                                     | UNKNOWN | 0.000292 | 0.000181 | 11:15  |
| 42                                                     | UNKNOWN | 0.000097 | 0.000096 | 11:29  |
| 43                                                     | UNKNOWN | 0.000195 | 0.000139 | 11:39  |
| 44                                                     | UNKNOWN | 0.014951 | 0.001227 | 12:01  |
| 45                                                     | UNKNOWN | 0.000097 | 0.000093 | 12:02  |
| 46                                                     | UNKNOWN | 0.000195 | 0.000141 | 12:41  |
| 47                                                     | UNKNOWN | 0.042007 | 0.001945 | 13:01  |
| 48                                                     | UNKNOWN | 0.030735 | 0.001692 | 13:02  |
| 49                                                     | UNKNOWN | 0.012645 | 0.001122 | 13:03  |
| 50                                                     | UNKNOWN | 0.003216 | 0.000566 | 13:05  |
| 51                                                     | UNKNOWN | 0.000097 | 0.000098 | 13:104 |
| 52                                                     | UNKNOWN | 0.000195 | 0.000141 | 13:11  |
| 53                                                     | UNKNOWN | 0.000097 | 0.000097 | 13:14  |
| 54                                                     | UNKNOWN | 0.000682 | 0.000262 | 13:21  |
| 55                                                     | UNKNOWN | 0.000097 | 0.000101 | 13:88  |
| 56                                                     | UNKNOWN | 0.002144 | 0.000446 | 13:XX  |
| 57                                                     | UNKNOWN | 0.053322 | 0.002212 | 14:01  |
| 58                                                     | UNKNOWN | 0.005184 | 0.000702 | 14:04  |
| 59                                                     | UNKNOWN | 0.000097 | 0.000098 | 14:05  |
| 60                                                     | UNKNOWN | 0.000195 | 0.000141 | 14:07  |
| 61                                                     | UNKNOWN | 0.000487 | 0.000211 | 14:143 |
| 62                                                     | UNKNOWN | 0.000097 | 0.000101 | 14:16  |
| 63                                                     | UNKNOWN | 0.000195 | 0.000133 | 14:17  |
| 64                                                     | UNKNOWN | 0.047133 | 0.001988 | 15:01  |
| 65                                                     | UNKNOWN | 0.022716 | 0.001474 | 15:02  |
| 66                                                     | UNKNOWN | 0.001462 | 0.000370 | 15:03  |
| 67                                                     | UNKNOWN | 0.000195 | 0.000142 | 15:06  |
| 68                                                     | UNKNOWN | 0.002046 | 0.000452 | 15:99  |
| 69                                                     | UNKNOWN | 0.128755 | 0.003209 | 16:01  |
| 70                                                     | UNKNOWN | 0.021050 | 0.001390 | 16:02  |
| 71                                                     | UNKNOWN | 0.000292 | 0.000163 | 16:05  |
| 72                                                     | UNKNOWN | 0.002631 | 0.000508 | 16:XX  |
| 73                                                     | UNKNOWN | 0.010197 | 0.001663 | blank  |
| -----                                                  |         |          |          |        |
| Sum of all 73 haplotype frequencies : 1.000000         |         |          |          |        |
| Sum of 73 listed frequencies : 1.000000                |         |          |          |        |
| =====                                                  |         |          |          |        |
| == Summary of computations done within populations     |         |          |          |        |
| =====                                                  |         |          |          |        |
| ////////////////////////////////////                   |         |          |          |        |
| END OF RUN NUMBER 1 (08/10/14 at 19:02:55))            |         |          |          |        |
| Total computing time for this run : 10h 26m 23s 355 ms |         |          |          |        |
| ////////////////////////////////////                   |         |          |          |        |

#### **5.2.5.2 Πλήθος & Συχνότητες Απλοτύπων**

Στα αποτελέσματα της στατιστικής ανάλυσης των δεδομένων εισόδου, παρατηρούνται και οι τιμές του πλήθους των διαφορετικών απλοτύπων που προκύπτουν από τον συνδυασμό των αλληλόμορφων των γονιδιακών τόπων HLA-A, HLA-B και HLA-DRB1, καθώς επίσης και οι εκτιμώμενες τιμές των συχνοτήτων κάθε απλοτύπου, σύμφωνα με τον αλγόριθμο EM.

Η τιμή του πλήθους των διαφορετικών (unique) παρατηρούμενων (observed) απλοτύπων (συμπεριλαμβανομένων των κενών - blank) είναι ίση με 9.962.

Οι εκτιμώμενες τιμές των συχνοτήτων των 100 συχνότερων απλοτύπων που προκύπτουν από τον συνδυασμό των αλληλόμορφων των τριών γονιδιακών τόπων HLA-A, HLA-B και HLA-DRB1, όπως αυτές προέκυψαν από την ανάλυση του δείγματος φαίνονται στον πίνακα I.

Όπως παρατηρούμε οι εκτιμώμενες συχνότητες των 100 συχνότερων απλοτύπων του δείγματος αποτελούν το 36,937% του συνόλου των απλοτύπων. Αυτό σημαίνει πως μεγάλο πλήθος απλοτύπων παρατηρείται σε έναν μόνο δότη και για μια μόνο φορά εντός του δείγματος, γεγονός που απεικονίζει την μεγάλη ποικιλομορφία και τις χαμηλές τιμές των εκτιμώμενων συχνοτήτων των απλοτύπων.

**ΠΙΝΑΚΑΣ Ι.** Αποτελέσματα τιμών εκτιμώμενων συχνοτήτων των 100 συχνότερων απλοτύπων

| FREQUENCY RANK | HLA-A | HLA-B | HLA-DRB1      | FREQUENCY (%)  |
|----------------|-------|-------|---------------|----------------|
| 1              | 02:01 | 18:01 | 11:04         | 2,634%         |
| 2              | 01:01 | 08:01 | 03:01         | 2,155%         |
| 3              | 02:01 | 51:01 | 16:01         | 1,302%         |
| 4              | 24:02 | 18:01 | 11:04         | 1,123%         |
| 5              | 02:01 | 51:01 | 11:01         | 1,035%         |
| 6              | 24:02 | 35:02 | 11:04         | 0,942%         |
| 7              | 33:01 | 14:02 | 01:02         | 0,750%         |
| 8              | 32:01 | 40:02 | 16:02         | 0,712%         |
| 9              | 02:01 | 51:01 | 11:04         | 0,636%         |
| 10             | 11:01 | 35:01 | 01:01         | 0,630%         |
| 11             | 02:01 | 51:01 | 14:01         | 0,616%         |
| 12             | 24:02 | 51:01 | 11:04         | 0,609%         |
| 13             | 01:01 | 52:01 | 15:02         | 0,588%         |
| 14             | 02:01 | 35:01 | 14:01         | 0,514%         |
| 15             | 02:01 | 39:01 | 16:01         | 0,512%         |
| 16             | 11:01 | 52:01 | 15:02         | 0,499%         |
| 17             | 03:01 | 35:01 | 01:01         | 0,491%         |
| 18             | 02:01 | 13:02 | 07:01         | 0,490%         |
| 19             | 02:01 | 44:05 | 16:01         | 0,474%         |
| 20             | 30:01 | 13:02 | 07:01         | 0,451%         |
| 21             | 02:01 | 51:01 | 13:01         | 0,439%         |
| 22             | 24:02 | 08:01 | 03:01         | 0,433%         |
| 23             | 02:01 | 27:02 | 16:01         | 0,407%         |
| 24             | 02:01 | 44:02 | 11:04         | 0,401%         |
| 25             | 02:01 | 44:02 | 16:01         | 0,392%         |
| 26             | 03:01 | 07:02 | 15:01         | 0,383%         |
| 27             | 32:01 | 44:02 | 11:04         | 0,381%         |
| 28             | 02:01 | 35:03 | 14:01         | 0,365%         |
| 29             | 01:01 | 57:01 | 07:01         | 0,356%         |
| 30             | 03:01 | 18:01 | 11:04         | 0,354%         |
| 31             | 01:01 | 37:01 | 16:01         | 0,344%         |
| 32             | 03:01 | 18:05 | 16:01         | 0,337%         |
| 33             | 02:01 | 18:01 | 16:01         | 0,314%         |
| 34             | 11:01 | 18:01 | 11:04         | 0,308%         |
| 35             | 23:01 | 49:01 | 11:01         | 0,306%         |
| 36             | 02:01 | 35:02 | 11:04         | 0,305%         |
| 37             | 32:01 | 51:01 | 11:01         | 0,303%         |
| 38             | 23:01 | 44:03 | 07:01         | 0,300%         |
| 39             | 24:02 | 39:01 | 16:01         | 0,300%         |
| 40             | 02:01 | 35:03 | 11:04         | 0,299%         |
| 41             | 32:01 | 18:01 | 11:04         | 0,289%         |
| 42             | 11:01 | 35:01 | 14:01         | 0,289%         |
| 43             | 03:01 | 07:02 | 16:01         | 0,285%         |
| 44             | 02:01 | 18:01 | 14:01         | 0,285%         |
| 45             | 24:02 | 18:01 | 16:01         | 0,280%         |
| 46             | 02:01 | 08:01 | 03:01         | 0,279%         |
| 47             | 03:01 | 41:01 | 04:05         | 0,278%         |
| 48             | 24:02 | 51:01 | 04:03         | 0,276%         |
| 49             | 11:01 | 51:01 | 11:04         | 0,273%         |
| 50             | 01:01 | 18:01 | 11:04         | 0,265%         |
| 51             | 02:01 | 73:01 | 04:05         | 0,265%         |
| 52             | 26:01 | 51:01 | 16:01         | 0,262%         |
| 53             | 11:01 | 35:01 | 16:01         | 0,261%         |
| 54             | 03:01 | 13:02 | 07:01         | 0,255%         |
| 55             | 68:01 | 18:01 | 11:04         | 0,250%         |
| 56             | 03:01 | 51:01 | 11:01         | 0,245%         |
| 57             | 25:01 | 18:01 | 15:01         | 0,242%         |
| 58             | 02:01 | 51:01 | 01:01         | 0,241%         |
| 59             | 01:01 | 51:01 | 11:04         | 0,240%         |
| 60             | 24:02 | 35:03 | 11:04         | 0,233%         |
| 61             | 03:01 | 38:01 | 13:01         | 0,233%         |
| 62             | 02:01 | 51:01 | 07:01         | 0,232%         |
| 63             | 02:01 | 51:01 | 04:03         | 0,229%         |
| 64             | 26:01 | 55:01 | 16:01         | 0,226%         |
| 65             | 24:02 | 35:03 | 14:01         | 0,225%         |
| 66             | 02:01 | 18:01 | 11:01         | 0,225%         |
| 67             | 24:02 | 55:01 | 16:01         | 0,224%         |
| 68             | 02:05 | 50:01 | 07:01         | 0,210%         |
| 69             | 01:01 | 35:01 | 01:01         | 0,209%         |
| 70             | 11:01 | 18:03 | 11:04         | 0,208%         |
| 71             | 24:02 | 13:02 | 07:01         | 0,207%         |
| 72             | 02:01 | 27:05 | 01:01         | 0,205%         |
| 73             | 24:02 | 51:01 | 11:01         | 0,204%         |
| 74             | 23:01 | 49:01 | 11:04         | 0,200%         |
| 75             | 11:01 | 35:01 | 11:01         | 0,194%         |
| 76             | 24:02 | 38:01 | 13:01         | 0,193%         |
| 77             | 02:01 | 35:08 | 16:01         | 0,189%         |
| 78             | 02:01 | 51:01 | 04:07         | 0,186%         |
| 79             | 68:01 | 18:01 | 16:01         | 0,184%         |
| 80             | 24:02 | 35:03 | 12:01         | 0,183%         |
| 81             | 24:02 | 07:02 | 15:01         | 0,175%         |
| 82             | 01:01 | 40:06 | 14:04         | 0,174%         |
| 83             | 26:01 | 51:01 | 11:01         | 0,173%         |
| 84             | 26:01 | 52:01 | 15:02         | 0,171%         |
| 85             | 26:01 | 39:01 | 16:01         | 0,171%         |
| 86             | 02:01 | 07:02 | 15:01         | 0,171%         |
| 87             | 03:01 | 35:01 | 11:01         | 0,170%         |
| 88             | 26:01 | 38:01 | 04:02         | 0,166%         |
| 89             | 24:02 | 44:02 | 11:04         | 0,166%         |
| 90             | 24:02 | 35:01 | 01:01         | 0,165%         |
| 91             | 11:01 | 55:01 | 16:01         | 0,164%         |
| 92             | 33:03 | 58:01 | 13:02         | 0,163%         |
| 93             | 24:02 | 18:01 | 15:01         | 0,162%         |
| 94             | 33:01 | 14:02 | 11:04         | 0,160%         |
| 95             | 30:01 | 39:01 | 16:01         | 0,159%         |
| 96             | 24:02 | 52:01 | 15:02         | 0,158%         |
| 97             | 03:01 | 07:02 | 11:01         | 0,157%         |
| 98             | 01:01 | 44:02 | 11:04         | 0,156%         |
| 99             | 33:03 | 58:01 | 03:01         | 0,155%         |
| 100            | 68:01 | 51:01 | 13:02         | 0,155%         |
|                |       |       | <b>ΣΥΝΟΛΟ</b> | <b>36,937%</b> |

Επίσης, στα αποτελέσματα των υπολογισμών παρατηρούνται και οι εκτιμώμενες σχετικές συχνότητες των φαινοτύπων των 5.131 εθελοντών δοτών του μητρώου, καθώς και των γονοτύπων, από τα ζεύγη των απλοτύπων που συμμετέχουν στον σχηματισμό του φαινοτύπου κάθε ατόμου.

### ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΑΠΛΟΤΥΠΩΝ ARLEQUIN 3.5

```
=====
== Haplotypes frequency estimation : (Greeks)
=====

    Dempster, A., N. Laird and D. Rubin, 1977.
    Excoffier, L. and M. Slatkin. 1995.
    Lange, K., 1997.
    Weir, B. S., 1996.
    Excoffier, L., Laval, G., and Balding, 2004.

-----
Conventional EM algorithm (one pass)
-----

No. of gene copies in sample           : 10262
No. of random initial conditions for EM : 50
No. of different maximum likelihoods found : 50
Epsilon value for stopping iterations   : 1.000000e-07
LogLHood : -66935.60925

Reference: Excoffier, L. and M. Slatkin. 1995.
No. of bootstraps for generating s.d.'s : 1000

-----
Maximum-likelihood haplotype frequencies :
-----

Total number of possible haplotypes      :          9962
Minimum frequency to reach for output    :          1.00e-05

-----
#   Haplotype      Freq.      s.d.
-----

1    UNKNOWN  0.000195  0.000150  01:01 07:02 03:01
2    UNKNOWN  0.000627  0.000294  01:01 07:02 11:04
3    UNKNOWN  0.000104  0.000114  01:01 07:02 14:01
4    UNKNOWN  0.000491  0.000276  01:01 07:02 15:01
5    UNKNOWN  0.000285  0.000172  01:01 07:05 03:01
6    UNKNOWN  0.000158  0.000129  01:01 07:05 04:05
7    UNKNOWN  0.000073  0.000088  01:01 07:05 16:01
8    UNKNOWN  0.000487  0.000246  01:01 08:01 01:01
9    UNKNOWN  0.021545  0.001469  01:01 08:01 03:01
10   UNKNOWN  0.000199  0.000146  01:01 08:01 08:01
. . . . .
2738 UNKNOWN  0.000195  0.000142  blank blank 13:XX
2739 UNKNOWN  0.003068  0.000663  blank blank blank
```

-----  
Sum of all 9962 haplotype frequencies : 1.000000

Sum of 2739 listed frequencies : 0.999939  
-----

Maximum-likelihood frequencies of genotypes  
-----

(Haplotypes that are not listed above have a negative id)

=====

Phenotype # 1  
=====

| Name          | Abs. Freq. | Rel. Freq. |
|---------------|------------|------------|
| CBMDP-0000501 | 1          | 0.000727   |

List of genotypes ( 6)  
-----

| Gen. # | Rel. Freq. | Exp. Freq. | Hapl. IDs | Haplotypes        |
|--------|------------|------------|-----------|-------------------|
| 1      | 0.942886   | 0.000686   | 491       | 02:01 51:01 16:01 |
|        |            |            | 308       | 02:01 18:01 11:04 |
| 2      | 0.054866   | 0.000040   | 316       | 02:01 18:01 16:01 |
|        |            |            | 484       | 02:01 51:01 11:04 |
| 3      | 0.002248   | 0.000002   | 491       | 02:01 51:01 16:01 |
|        |            |            | 2685      | blank 18:01 11:04 |

=====

Phenotype # 2  
=====

| Name          | Abs. Freq. | Rel. Freq. |
|---------------|------------|------------|
| CBMDP-0000502 | 1          | 0.000037   |

List of genotypes ( 4)  
-----

| Gen. # | Rel. Freq. | Exp. Freq. | Hapl. IDs | Haplotypes        |
|--------|------------|------------|-----------|-------------------|
| 1      | 0.972850   | 0.000036   | 9         | 01:01 08:01 03:01 |
|        |            |            | 1356      | 24:02 27:05 04:04 |
| 2      | 0.027150   | 0.000001   | 49        | 01:01 27:05 04:04 |
|        |            |            | 1303      | 24:02 08:01 03:01 |

. . . . .

=====

Phenotype # 5131  
=====

| Name          | Abs. Freq. | Rel. Freq. |
|---------------|------------|------------|
| CBMDP-0005632 | 1          | 0.000000   |

List of genotypes ( 6)  
-----

| Gen. # | Rel. Freq. | Exp. Freq. | Hapl. IDs | Haplotypes        |
|--------|------------|------------|-----------|-------------------|
| 1      | 1.000000   | 0.000000   | 84        | 01:01 35:08 14:01 |
|        |            |            | 2730      | blank 55:01 07:01 |

\*\*\*\*\*

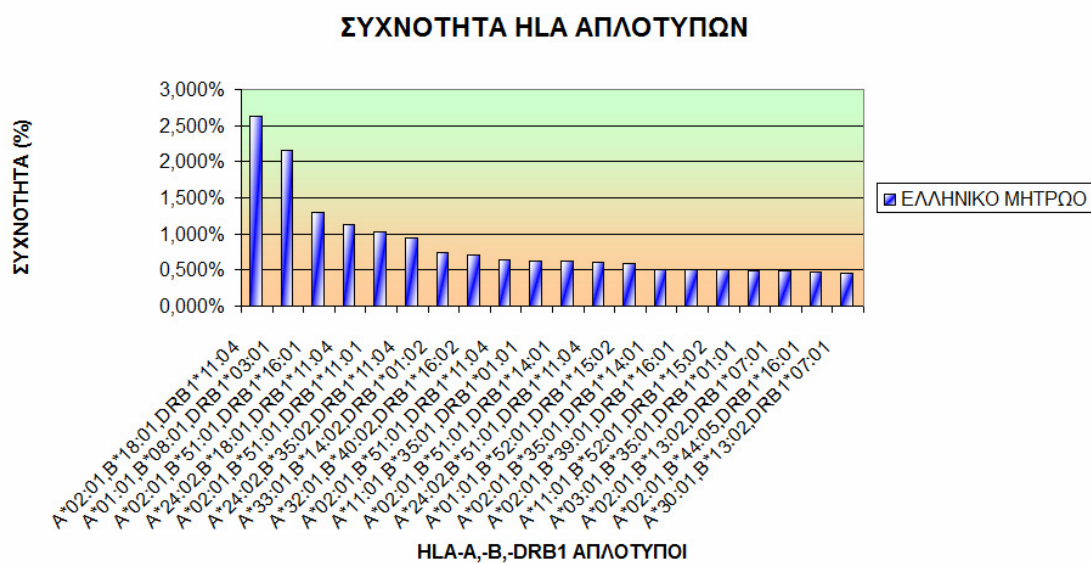
### Σύγκριση με Γερμανικό Μητρώο

Στο σημείο αυτό κρίνεται σκόπιμο να παρουσιαστούν ορισμένα συγκριτικά στοιχεία όσον αφορά στις εκτιμώμενες συχνότητες των απλοτύπων του Ελληνικού μητρώου ΚΕΔΜΟΠ των 5.131 εγγεγραμμένων εθελοντών δοτών σε σχέση με τις αντίστοιχες συχνότητες απλοτύπων ενός Γερμανικού μητρώου 8.862 εγγεγραμμένων εθελοντών δοτών [1].

Στον πίνακα Κ φαίνονται οι 20 συχνότεροι HLA απλότυποι του Ελληνικού Μητρώου κατά φθίνουσα σειρά ως προς τις εκτιμώμενες τιμές των συχνοτήτων τους.

**ΠΙΝΑΚΑΣ Κ.** Συχνότητες 20 συχνότερων απλοτύπων του Ελληνικού Μητρώου

| ΣΥΧΝΟΤΗΤΑ (%)<br>ΕΛΛΗΝΙΚΟ ΜΗΤΡΩΟ | HLA ΑΠΛΟΤΥΠΟΣ              |
|----------------------------------|----------------------------|
| 2,634%                           | A*02:01,B*18:01,DRB1*11:04 |
| 2,155%                           | A*01:01,B*08:01,DRB1*03:01 |
| 1,302%                           | A*02:01,B*51:01,DRB1*16:01 |
| 1,123%                           | A*24:02,B*18:01,DRB1*11:04 |
| 1,035%                           | A*02:01,B*51:01,DRB1*11:01 |
| 0,942%                           | A*24:02,B*35:02,DRB1*11:04 |
| 0,750%                           | A*33:01,B*14:02,DRB1*01:02 |
| 0,712%                           | A*32:01,B*40:02,DRB1*16:02 |
| 0,636%                           | A*02:01,B*51:01,DRB1*11:04 |
| 0,630%                           | A*11:01,B*35:01,DRB1*01:01 |
| 0,616%                           | A*02:01,B*51:01,DRB1*14:01 |
| 0,609%                           | A*24:02,B*51:01,DRB1*11:04 |
| 0,588%                           | A*01:01,B*52:01,DRB1*15:02 |
| 0,514%                           | A*02:01,B*35:01,DRB1*14:01 |
| 0,512%                           | A*02:01,B*39:01,DRB1*16:01 |
| 0,499%                           | A*11:01,B*52:01,DRB1*15:02 |
| 0,491%                           | A*03:01,B*35:01,DRB1*01:01 |
| 0,490%                           | A*02:01,B*13:02,DRB1*07:01 |
| 0,474%                           | A*02:01,B*44:05,DRB1*16:01 |

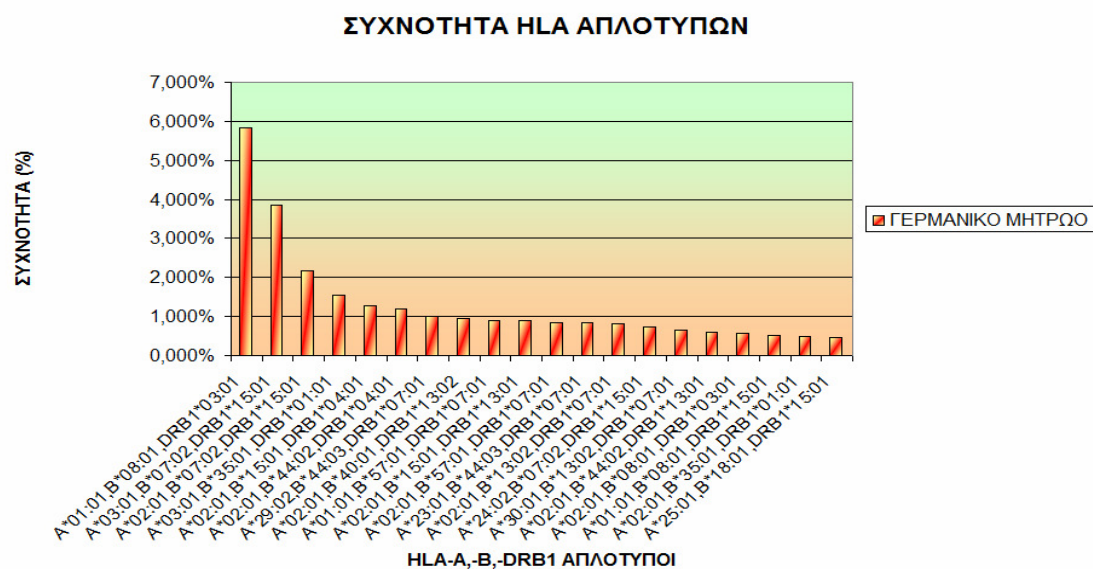


**ΕΙΚΟΝΑ 5.12** Γραφική παράσταση συχνοτήτων των 20 συχνότερων απλοτύπων του Ελληνικού Μητρώου

Στον πίνακα Λ φαίνονται οι 20 συχνότεροι HLA απλότυποι του Γερμανικού Μητρώου κατά φθίνουσα σειρά ως προς τις εκτιμώμενες τιμές των συχνοτήτων τους.

**ΠΙΝΑΚΑΣ Λ.** Συχνότητες 20 συχνότερων απλοτύπων του Γερμανικού Μητρώου

| ΣΥΧΝΟΤΗΤΑ (%)<br>ΓΕΡΜΑΝΙΚΟ ΜΗΤΡΩΟ | HLA ΑΠΛΟΤΥΠΟΣ              |
|-----------------------------------|----------------------------|
| 5,826%                            | A*01:01,B*08:01,DRB1*03:01 |
| 3,843%                            | A*03:01,B*07:02,DRB1*15:01 |
| 2,181%                            | A*02:01,B*07:02,DRB1*15:01 |
| 1,540%                            | A*03:01,B*35:01,DRB1*01:01 |
| 1,286%                            | A*02:01,B*15:01,DRB1*04:01 |
| 1,207%                            | A*02:01,B*44:02,DRB1*04:01 |
| 1,005%                            | A*29:02,B*44:03,DRB1*07:01 |
| 0,951%                            | A*02:01,B*40:01,DRB1*13:02 |
| 0,901%                            | A*01:01,B*57:01,DRB1*07:01 |
| 0,899%                            | A*02:01,B*15:01,DRB1*13:01 |
| 0,831%                            | A*02:01,B*57:01,DRB1*07:01 |
| 0,829%                            | A*23:01,B*44:03,DRB1*07:01 |
| 0,822%                            | A*02:01,B*13:02,DRB1*07:01 |
| 0,740%                            | A*24:02,B*07:02,DRB1*15:01 |
| 0,648%                            | A*30:01,B*13:02,DRB1*07:01 |
| 0,597%                            | A*02:01,B*44:02,DRB1*13:01 |
| 0,580%                            | A*02:01,B*08:01,DRB1*03:01 |
| 0,519%                            | A*01:01,B*08:01,DRB1*15:01 |
| 0,488%                            | A*02:01,B*35:01,DRB1*01:01 |



**ΕΙΚΟΝΑ 5.13** Γραφική παράσταση συχνοτήτων των 20 συχνότερων απλοτύπων του Γερμανικού Μητρώου

Στον πίνακα Μ φαίνονται οι 15 κοινός συχνότεροι HLA απλότυποι του Ελληνικού και του αντίστοιχου Γερμανικού Μητρώου.

**ΠΙΝΑΚΑΣ Μ.** Συχνότητες 15 κοινών συχνότερων απλοτύπων Ελληνικού & Γερμανικού Μητρώου

| ΣΥΧΝΟΤΗΤΑ (%)<br>ΕΛΛΗΝΙΚΟ ΜΗΤΡΩΟ | ΗΛΑ ΑΠΛΟΤΥΠΟΣ              | ΣΥΧΝΟΤΗΤΑ (%)<br>ΓΕΡΜΑΝΙΚΟ ΜΗΤΡΩΟ |
|----------------------------------|----------------------------|-----------------------------------|
| 2,634%                           | A*02:01,B*18:01,DRB1*11:04 | 0,344%                            |
| 2,155%                           | A*01:01,B*08:01,DRB1*03:01 | 5,826%                            |
| 0,750%                           | A*33:01,B*14:02,DRB1*01:02 | 0,429%                            |
| 0,630%                           | A*11:01,B*35:01,DRB1*01:01 | 0,423%                            |
| 0,491%                           | A*03:01,B*35:01,DRB1*01:01 | 1,540%                            |
| 0,490%                           | A*02:01,B*13:02,DRB1*07:01 | 0,822%                            |
| 0,451%                           | A*30:01,B*13:02,DRB1*07:01 | 0,648%                            |
| 0,407%                           | A*02:01,B*27:02,DRB1*16:01 | 0,244%                            |
| 0,383%                           | A*03:01,B*07:02,DRB1*15:01 | 3,843%                            |
| 0,356%                           | A*01:01,B*57:01,DRB1*07:01 | 0,901%                            |
| 0,300%                           | A*23:01,B*44:03,DRB1*07:01 | 0,829%                            |
| 0,279%                           | A*02:01,B*08:01,DRB1*03:01 | 0,580%                            |
| 0,242%                           | A*25:01,B*18:01,DRB1*15:01 | 0,472%                            |
| 0,210%                           | A*02:05,B*50:01,DRB1*07:01 | 0,271%                            |
| 0,207%                           | A*24:02,B*13:02,DRB1*07:01 | 0,282%                            |



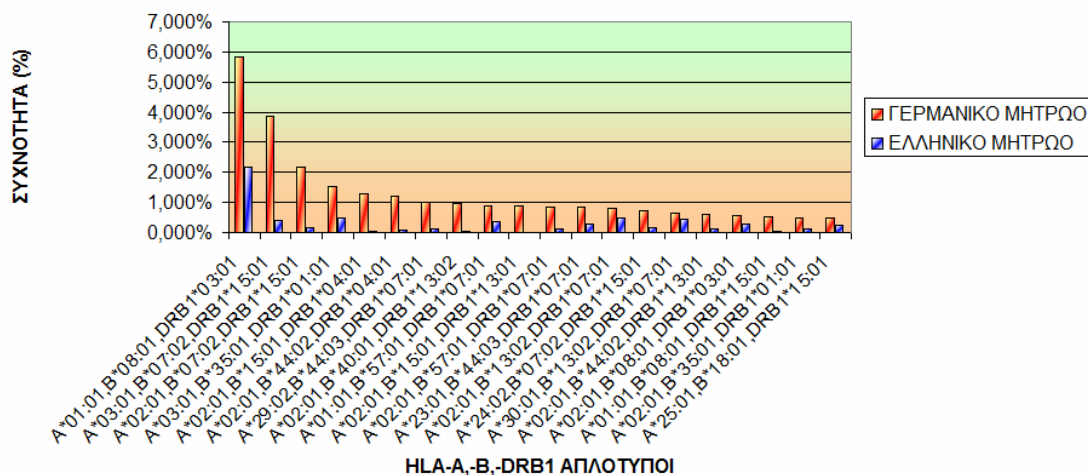
**ΕΙΚΟΝΑ 5.14** Γραφική παράσταση συχνοτήτων των 15 κοινών συχνότερων απλοτύπων του Ελληνικού & του Γερμανικού Μητρώου

Στον πίνακα N φαίνονται οι 20 κοινοί συχνότεροι ΗΛΑ απλότυποι του Γερμανικού και του αντίστοιχου Ελληνικού Μητρώου, ως προς φθίνουσα σειρά ταξινόμησης των εκτιμώμενων τιμών συχνοτήτων απλοτύπων του Γερμανικού Μητρώου.

**ΠΙΝΑΚΑΣ Ν.** Συχνότητες 20 κοινών συχνότερων απλοτύπων Ελληνικού & Γερμανικού Μητρώου ως προς φθίνουσα σειρά ταξινόμησης των συχνοτήτων του Γερμανικού Μητρώου

| ΣΥΧΝΟΤΗΤΑ (%)<br>ΓΕΡΜΑΝΙΚΟ ΜΗΤΡΩΟ | ΗΛΑ ΑΠΛΟΤΥΠΟΣ              | ΣΥΧΝΟΤΗΤΑ (%)<br>ΕΛΛΗΝΙΚΟ ΜΗΤΡΩΟ |
|-----------------------------------|----------------------------|----------------------------------|
| 5,826%                            | A*01:01,B*08:01,DRB1*03:01 | 2,1545%                          |
| 3,843%                            | A*03:01,B*07:02,DRB1*15:01 | 0,3833%                          |
| 2,181%                            | A*02:01,B*07:02,DRB1*15:01 | 0,1708%                          |
| 1,540%                            | A*03:01,B*35:01,DRB1*01:01 | 0,4905%                          |
| 1,286%                            | A*02:01,B*15:01,DRB1*04:01 | 0,0427%                          |
| 1,207%                            | A*02:01,B*44:02,DRB1*04:01 | 0,0975%                          |
| 1,005%                            | A*29:02,B*44:03,DRB1*07:01 | 0,1105%                          |
| 0,951%                            | A*02:01,B*40:01,DRB1*13:02 | 0,0471%                          |
| 0,901%                            | A*01:01,B*57:01,DRB1*07:01 | 0,3559%                          |
| 0,899%                            | A*02:01,B*15:01,DRB1*13:01 | 0,0000%                          |
| 0,831%                            | A*02:01,B*57:01,DRB1*07:01 | 0,1364%                          |
| 0,829%                            | A*23:01,B*44:03,DRB1*07:01 | 0,3000%                          |
| 0,822%                            | A*02:01,B*13:02,DRB1*07:01 | 0,4895%                          |
| 0,740%                            | A*24:02,B*07:02,DRB1*15:01 | 0,1750%                          |
| 0,648%                            | A*30:01,B*13:02,DRB1*07:01 | 0,4513%                          |
| 0,597%                            | A*02:01,B*44:02,DRB1*13:01 | 0,1227%                          |
| 0,580%                            | A*02:01,B*08:01,DRB1*03:01 | 0,2789%                          |
| 0,519%                            | A*01:01,B*08:01,DRB1*15:01 | 0,0356%                          |
| 0,488%                            | A*02:01,B*35:01,DRB1*01:01 | 0,1011%                          |
| 0,472%                            | A*25:01,B*18:01,DRB1*15:01 | 0,2423%                          |

**ΣΥΧΝΟΤΗΤΑ ΗΛΑ ΑΠΛΟΤΥΠΩΝ**  
**Συχνότεροι κοινοί ΗΛΑ απλότυποι**  
**Ελληνικού-Γερμανικού Μητρώου**  
**ως προς το Γερμανικό Μητρώο**



**ΕΙΚΟΝΑ 5.15** Γραφική παράσταση συχνοτήτων των 20 κοινών συχνότερων απλοτύπων του Ελληνικού & του Γερμανικού Μητρώου ως προς το Γερμανικό Μητρώο

### 5.2.5.3 Πλήθος & Συχνότητες Αντιγόνων

#### ΣΕΝΑΡΙΟ Β : Low Resolution HLA-A, HLA-B & HLA-DRB1

Για την υλοποίηση αναλύσεων και την εκτίμηση συχνοτήτων σε επίπεδο αντιγόνων, εισάγουμε τα προς ανάλυση δεδομένα του μητρώου ΚΕΔΜΟΠ των 5.131 Ελλήνων εθελοντών δοτών σε μορφή 2-digit.

```
[Data]

[[Samples]]
SampleName="Greeks"
SampleSize=5131
SampleData={

CBMDP-0000501  1      02      18      16
                02      51      11
CBMDP-0000502  1      01      08      03
                24      27      04
CBMDP-0000503  1      02      35      16
                30      51      04
CBMDP-0000504  1      01      44      15
                32      52      11
CBMDP-0000505  1      24      35      15
                32      44      16
. . . . .
}
```

Στα αποτελέσματα της στατιστικής ανάλυσης των δεδομένων εισόδου, παρατηρούνται οι τιμές του πλήθους των διαφορετικών αντιγόνων που προκύπτουν για τους γονιδιακούς τόπους HLA-A, HLA-B και HLA-DRB1 ξεχωριστά, καθώς επίσης και οι τιμές των συχνοτήτων κάθε αντιγόνου.

Οι τιμές του πλήθους των παρατηρούμενων (observed) αντιγόνων για κάθε γονιδιακό τόπο (συμπεριλαμβανομένων των κενών - blank) φαίνονται στον πίνακα Ξ.

**ΠΙΝΑΚΑΣ Ξ.** Αποτελέσματα πλήθους παρατηρούμενων αντιγόνων

| ΓΟΝΙΔΙΑΚΟΣ ΤΟΠΟΣ<br>(LOCUS) | ΠΛΗΘΟΣ<br>UNIQUE ANTIGONΩΝ |
|-----------------------------|----------------------------|
| HLA-A                       | 20                         |
| HLA-B                       | 35                         |
| HLA-DRB1                    | 14                         |

Όπως είναι φανερό, η μεγαλύτερη ποικιλομορφία αντιγόνων εμφανίζεται για το γονίδιο HLA-B στο σύνολο των δεδομένων του μητρώου, και η μικρότερη εμφανίζεται για το γονίδιο HLA-DRB1.

Οι τιμές των συχνοτήτων όλων των αντιγόνων των τριών γονιδιακών τόπων HLA-A, HLA-B και HLA-DRB1, όπως αυτές προέκυψαν από την ανάλυση του δείγματος φαίνονται στον πίνακα Ο.

**ΠΙΝΑΚΑΣ Ο.** Αποτελέσματα συχνοτήτων αντιγόνων

| HLA-A    |               | HLA-B    |               | HLA-DRB1 |               |
|----------|---------------|----------|---------------|----------|---------------|
| ΑΝΤΙΓΟΝΟ | ΣΥΧΝΟΤΗΤΑ (%) | ΑΝΤΙΓΟΝΟ | ΣΥΧΝΟΤΗΤΑ (%) | ΑΝΤΙΓΟΝΟ | ΣΥΧΝΟΤΗΤΑ (%) |
| 02       | 26,092%       | 35       | 17,355%       | 11       | 25,764%       |
| 24       | 14,958%       | 51       | 13,504%       | 16       | 15,232%       |
| 01       | 10,391%       | 18       | 11,635%       | 13       | 9,145%        |
| 03       | 9,575%        | 44       | 7,525%        | 04       | 9,094%        |
| 11       | 7,955%        | 08       | 4,573%        | 03       | 7,688%        |
| 32       | 7,257%        | 07       | 3,916%        | 15       | 7,336%        |
| 26       | 5,118%        | 40       | 3,769%        | 01       | 6,621%        |
| 68       | 3,666%        | 39       | 3,413%        | 07       | 6,480%        |
| 30       | 3,236%        | 15       | 3,340%        | 14       | 5,938%        |
| 23       | 2,486%        | 55       | 3,036%        | 10       | 1,898%        |
| 33       | 2,317%        | 13       | 2,967%        | 08       | 1,563%        |
| 29       | 1,892%        | 49       | 2,850%        | 12       | 1,524%        |
| 31       | 1,628%        | 38       | 2,786%        | blank    | 1,405%        |
| blank    | 1,539%        | 52       | 2,737%        | 09       | 0,313%        |
| 25       | 0,750%        | 27       | 2,572%        |          |               |
| 69       | 0,507%        | 14       | 2,326%        |          |               |
| 66       | 0,497%        | 41       | 1,937%        |          |               |
| 74       | 0,117%        | 57       | 1,559%        |          |               |
| 34       | 0,010%        | 58       | 1,517%        |          |               |
| 80       | 0,010%        | blank    | 1,478%        |          |               |
|          |               | 37       | 1,419%        |          |               |
|          |               | 50       | 1,367%        |          |               |
|          |               | 47       | 0,536%        |          |               |
|          |               | 53       | 0,517%        |          |               |
|          |               | 56       | 0,429%        |          |               |
|          |               | 73       | 0,352%        |          |               |
|          |               | 45       | 0,224%        |          |               |
|          |               | 34       | 0,166%        |          |               |
|          |               | 48       | 0,097%        |          |               |
|          |               | 42       | 0,029%        |          |               |
|          |               | 54       | 0,020%        |          |               |
|          |               | 78       | 0,020%        |          |               |
|          |               | 31       | 0,010%        |          |               |
|          |               | 59       | 0,010%        |          |               |
|          |               | 81       | 0,010%        |          |               |

### ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΑΝΤΙΓΟΝΩΝ ARLEQUIN 3.5

\*\*\*\*\*

Allele frequencies :

(1000 bootstrap replicates)

Allele frequencies for the locus 1

-----

No. of gene copies in sample : 10262

No. of random initial conditions for EM : 50

No. of bootstrap replicates : 1000

No. of different maximum likelihoods found : 1

Epsilon value for stopping iterations : 1.00000e-07

Logarithm of the sample maximum-likelihood : -20712.5

-----

Maximum-likelihood haplotype frequencies :

-----

Total number of possible haplotypes : 20  
 Minimum frequency to reach for output : 1.00e-05

| #  | Haplotype | Freq.    | s.d.     |       |
|----|-----------|----------|----------|-------|
| 1  | UNKNOWN   | 0.103909 | 0.002896 | 01    |
| 2  | UNKNOWN   | 0.260924 | 0.004107 | 02    |
| 3  | UNKNOWN   | 0.095751 | 0.002845 | 03    |
| 4  | UNKNOWN   | 0.079551 | 0.002612 | 11    |
| 5  | UNKNOWN   | 0.024861 | 0.001506 | 23    |
| 6  | UNKNOWN   | 0.149582 | 0.003464 | 24    |
| 7  | UNKNOWN   | 0.007503 | 0.000850 | 25    |
| 8  | UNKNOWN   | 0.051183 | 0.002161 | 26    |
| 9  | UNKNOWN   | 0.018919 | 0.001347 | 29    |
| 10 | UNKNOWN   | 0.032362 | 0.001737 | 30    |
| 11 | UNKNOWN   | 0.016277 | 0.001245 | 31    |
| 12 | UNKNOWN   | 0.072565 | 0.002426 | 32    |
| 13 | UNKNOWN   | 0.023165 | 0.001431 | 33    |
| 14 | UNKNOWN   | 0.000097 | 0.000094 | 34    |
| 15 | UNKNOWN   | 0.004970 | 0.000693 | 66    |
| 16 | UNKNOWN   | 0.036657 | 0.001779 | 68    |
| 17 | UNKNOWN   | 0.005067 | 0.000681 | 69    |
| 18 | UNKNOWN   | 0.001169 | 0.000349 | 74    |
| 19 | UNKNOWN   | 0.000097 | 0.000096 | 80    |
| 20 | UNKNOWN   | 0.015390 | 0.002283 | blank |

Sum of all 20 haplotype frequencies : 1.000000

Sum of 20 listed frequencies : 1.000000

Allele frequencies for the locus 2

No. of gene copies in sample : 10262  
 No. of random initial conditions for EM : 50  
 No. of bootstrap replicates : 1000  
 No. of different maximum likelihoods found : 1  
 Epsilon value for stopping iterations : 1.00000e-07  
 Logarithm of the sample maximum-likelihood : -25768.8

Maximum-likelihood haplotype frequencies :

Total number of possible haplotypes : 35  
 Minimum frequency to reach for output : 1.00e-05

| #                                                   | Haplotype | Freq.    | s.d.     |       |
|-----------------------------------------------------|-----------|----------|----------|-------|
| 1                                                   | UNKNOWN   | 0.039158 | 0.001870 | 07    |
| 2                                                   | UNKNOWN   | 0.045731 | 0.002029 | 08    |
| 3                                                   | UNKNOWN   | 0.029673 | 0.001616 | 13    |
| 4                                                   | UNKNOWN   | 0.023255 | 0.001443 | 14    |
| 5                                                   | UNKNOWN   | 0.033396 | 0.001713 | 15    |
| 6                                                   | UNKNOWN   | 0.116350 | 0.003067 | 18    |
| 7                                                   | UNKNOWN   | 0.025719 | 0.001469 | 27    |
| 8                                                   | UNKNOWN   | 0.000097 | 0.000100 | 31    |
| 9                                                   | UNKNOWN   | 0.001657 | 0.000390 | 34    |
| 10                                                  | UNKNOWN   | 0.173549 | 0.003708 | 35    |
| 11                                                  | UNKNOWN   | 0.014193 | 0.001154 | 37    |
| 12                                                  | UNKNOWN   | 0.027864 | 0.001642 | 38    |
| 13                                                  | UNKNOWN   | 0.034134 | 0.001779 | 39    |
| 14                                                  | UNKNOWN   | 0.037685 | 0.001897 | 40    |
| 15                                                  | UNKNOWN   | 0.019372 | 0.001368 | 41    |
| 16                                                  | UNKNOWN   | 0.000292 | 0.000168 | 42    |
| 17                                                  | UNKNOWN   | 0.075254 | 0.002565 | 44    |
| 18                                                  | UNKNOWN   | 0.002241 | 0.000479 | 45    |
| 19                                                  | UNKNOWN   | 0.005360 | 0.000731 | 47    |
| 20                                                  | UNKNOWN   | 0.000974 | 0.000300 | 48    |
| 21                                                  | UNKNOWN   | 0.028502 | 0.001659 | 49    |
| 22                                                  | UNKNOWN   | 0.013673 | 0.001097 | 50    |
| 23                                                  | UNKNOWN   | 0.135040 | 0.003285 | 51    |
| 24                                                  | UNKNOWN   | 0.027371 | 0.001505 | 52    |
| 25                                                  | UNKNOWN   | 0.005165 | 0.000711 | 53    |
| 26                                                  | UNKNOWN   | 0.000195 | 0.000133 | 54    |
| 27                                                  | UNKNOWN   | 0.030358 | 0.001719 | 55    |
| 28                                                  | UNKNOWN   | 0.004288 | 0.000634 | 56    |
| 29                                                  | UNKNOWN   | 0.015592 | 0.001206 | 57    |
| 30                                                  | UNKNOWN   | 0.015170 | 0.001184 | 58    |
| 31                                                  | UNKNOWN   | 0.000097 | 0.000097 | 59    |
| 32                                                  | UNKNOWN   | 0.003518 | 0.000590 | 73    |
| 33                                                  | UNKNOWN   | 0.000195 | 0.000139 | 78    |
| 34                                                  | UNKNOWN   | 0.000097 | 0.000092 | 81    |
| 35                                                  | UNKNOWN   | 0.014783 | 0.001941 | blank |
| Sum of all 35 haplotype frequencies : 1.000000      |           |          |          |       |
| Sum of 35 listed frequencies : 1.000000             |           |          |          |       |
| Allele frequencies for the locus 3                  |           |          |          |       |
| -----                                               |           |          |          |       |
| No. of gene copies in sample : 10262                |           |          |          |       |
| No. of random initial conditions for EM : 50        |           |          |          |       |
| No. of bootstrap replicates : 1000                  |           |          |          |       |
| No. of different maximum likelihoods found : 1      |           |          |          |       |
| Epsilon value for stopping iterations : 1.00000e-07 |           |          |          |       |

Logarithm of the sample maximum-likelihood : -19830.8

Maximum-likelihood haplotype frequencies :

Total number of possible haplotypes : 14

Minimum frequency to reach for output : 1.00e-05

| #  | Haplotype | Freq.    | s.d.     |       |
|----|-----------|----------|----------|-------|
| 1  | UNKNOWN   | 0.066210 | 0.002354 | 01    |
| 2  | UNKNOWN   | 0.076883 | 0.002469 | 03    |
| 3  | UNKNOWN   | 0.090937 | 0.002683 | 04    |
| 4  | UNKNOWN   | 0.064797 | 0.002410 | 07    |
| 5  | UNKNOWN   | 0.015633 | 0.001250 | 08    |
| 6  | UNKNOWN   | 0.003128 | 0.000536 | 09    |
| 7  | UNKNOWN   | 0.018983 | 0.001366 | 10    |
| 8  | UNKNOWN   | 0.257635 | 0.004160 | 11    |
| 9  | UNKNOWN   | 0.015236 | 0.001176 | 12    |
| 10 | UNKNOWN   | 0.091452 | 0.002831 | 13    |
| 11 | UNKNOWN   | 0.059377 | 0.002291 | 14    |
| 12 | UNKNOWN   | 0.073358 | 0.002521 | 15    |
| 13 | UNKNOWN   | 0.152318 | 0.003501 | 16    |
| 14 | UNKNOWN   | 0.014053 | 0.002265 | blank |

Sum of all 14 haplotype frequencies : 1.000000

#### 5.2.5.4 Έλεγχος Ισορροπίας Hardy-Weinberg

Ο στατιστικός έλεγχος υποθέσεων (hypothesis testing) είναι μια συμπερασματική διαδικασία/μέθοδος που προσφέρει η Στατιστική Συμπερασματολογία και βρίσκει εφαρμογή σε στοχαστικά προβλήματα απόφασης μεταξύ δύο εναλλακτικών υποθέσεων. Η μία υπόθεση έχει επικρατήσει να συμβολίζεται ως  $H_0$  και ονομάζεται μηδενική υπόθεση (null hypothesis) και η άλλη ως  $H_1$  και ονομάζεται εναλλακτική υπόθεση (alternative hypothesis). Όταν το αποτέλεσμα του ελέγχου δεν έχει προέλθει από τυχαίους παράγοντες, τότε θεωρείται στατιστικά σημαντικό, ενώ στην αντίθετη περίπτωση που έχει προέλθει από τυχαίους παράγοντες θεωρείται στατιστικά μη σημαντικό. Το κατά πόσο μικρή πρέπει να είναι η πιθανότητα ώστε να μην υπάρχουν τυχαία αποτελέσματα έχει καθοριστεί κατά κύριο λόγο από τις κοινωνικές επιστήμες ως πιθανότητα σφάλματος **p-value** το πολύ 5%, ώστε να προκύπτουν στατιστικά σημαντικά αποτελέσματα. Ωστόσο υπάρχουν και μικρότερα επίπεδα σημαντικότητας όπως 1% . Ο έλεγχος ισορροπίας HW διεξήχθη διότι, τόσο ο αλγόριθμος EM, όσο και

ο υπολογισμός των συχνοτήτων των δεδομένων του μητρώου των δοτών, απαιτούν από τον υπό ανάλυση πληθυσμό να βρίσκεται θεωρητικά σε ισορροπία Hardy-Weinberg. Ωστόσο, η παρουσία ελλιπών (recessive-blank) αλληλόμορφων απαγορεύει την διενέργεια επακριβούς ελέγχου HWE επί των δεδομένων. Για το λόγο αυτό αγνοήθηκε η παρουσία ελλιπών αλληλόμορφων κατά τη διενέργεια του ελέγχου HWE. Η τιμή p-value απεικονίζει την πιθανότητα εύρεσης της διαφοράς μεταξύ παρατηρούμενων (observed) και αναμενόμενων (expected) τιμών συχνοτήτων υπό την ισχύ της ισορροπίας HW στα δεδομένα του υπό ανάλυση πληθυσμού.

### ΕΛΕΓΧΟΣ ΥΠΟΘΕΣΕΩΝ

**H<sub>0</sub>** : Ο υπό ανάλυση πληθυσμός βρίσκεται σε ισορροπία HW.

**H<sub>1</sub>** : Ο υπό ανάλυση πληθυσμός δεν βρίσκεται σε ισορροπία HW.

### P-VALUE

❖ Αν **p-value > 0,05**

1. Δεν απορρίπτεται η H<sub>0</sub>
2. Η διαφοροποίηση μεταξύ των παρατηρούμενων (observed) συχνοτήτων γονοτύπων και των αντίστοιχων αναμενόμενων (expected) υπό την ισχύ της ισορροπίας HW **δεν είναι στατιστικά σημαντική.**

❖ Αν **p-value < 0,05**

1. **Απορρίπτεται η H<sub>0</sub>**
2. Η διαφοροποίηση μεταξύ των παρατηρούμενων (observed) συχνοτήτων γονοτύπων και των αντίστοιχων αναμενόμενων (expected) υπό την ισχύ της ισορροπίας HW **είναι στατιστικά σημαντική.**

**ΠΙΝΑΚΑΣ Π.** Αποτελέσματα ελέγχου ισορροπίας Hardy-Weinberg (HWE)

| ΓΟΝΙΔΙΑΚΟΣ ΤΟΠΟΣ (LOCUS) | ΠΑΡΑΤΗΡΟΥΜΕΝΗ ΕΤΕΡΟΖΥΓΩΤΙΑ (Observed Heterozygosity) | ΑΝΑΜΕΝΟΜΕΝΗ ΕΤΕΡΟΖΥΓΩΤΙΑ (Expected Heterozygosity) | P-value |
|--------------------------|------------------------------------------------------|----------------------------------------------------|---------|
| HLA-A                    | 0,87702                                              | 0,88831                                            | 0,00000 |
| HLA-B                    | 0,94095                                              | 0,95044                                            | 0,00002 |
| HLA-DRB1                 | 0,92360                                              | 0,92671                                            | 0,00000 |

Από τα αποτελέσματα της διενέργειας ελέγχου HWE γίνεται σαφές ότι η διαφοροποίηση μεταξύ παρατηρούμενων και αναμενόμενων τιμών συχνοτήτων γονοτύπων του δείγματος του υπό ανάλυση πληθυσμού είναι στατιστικά σημαντική για όλους τους γονιδιακούς τόπους του HLA που εξετάζονται και απορρίπτεται η μηδενική υπόθεση H<sub>0</sub>. Αυτό σημαίνει πως ο υπό ανάλυση πληθυσμός δεν βρίσκεται σε ισορροπία HWE στην πραγματικότητα και η πιθανότητα εύρεσης των πραγματικών αποτελεσμάτων όταν θεωρείται ότι ο πληθυσμός βρίσκεται σε ισορροπία HW είναι μικρότερη του 5%.

## ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΛΕΓΧΟΥ ΗWE ARLEQUIN 3.5

### Settings used for Calculations

#### General settings:

```

Deletion Weight           = 1
Transition Weight Weight  = 1
Tranversion Weight Weight = 1
Epsilon Value             = 1e-07
Significant digits for output = 5
Use original haplotype definition
Allowed level of missing data = 0.05

```

#### Active Tasks:

#### Hardy-Weinberg equilibrium test:

```

No. of steps in Markov chain      = 1000000
No. of Dememorisation Steps      = 100000
Required precision on Probability = 0
Test association at the Locus level

```

### == ANALYSES AT THE INTRA-POPULATION LEVEL

== Sample :   Greeks

== Hardy-Weinberg equilibrium : (Greeks)

Guo, S. and Thompson, E. 1992.

Levene H. (1949).

Exact test using a Markov chain (for all Loci):

Forecasted chain length       :1000000

Dememorization steps         :100000

| Locus | #Genot | Obs.Het. | Exp.Het. | P-value | s.d.    | Steps done |
|-------|--------|----------|----------|---------|---------|------------|
| 1     | 5131   | 0.87702  | 0.88831  | 0.00000 | 0.00000 | 1001000    |
| 2     | 5131   | 0.94095  | 0.95044  | 0.00081 | 0.00002 | 1001000    |
| 3     | 5131   | 0.92360  | 0.92671  | 0.00000 | 0.00000 | 1001000    |

== Summary of computations done within populations

### 5.2.5.5 Ανισορροπία Σύνδεσης

Για τον προσδιορισμό της τιμής της ανισορροπίας σύνδεσης μεταξύ ζευγών αλληλόμορφων των γονιδιακών τόπων HLA-A, HLA-B και HLA-DRB1 μέσω του πακέτου Arlequin 3.5 θεωρούμε ότι η γαμετική φάση είναι γνωστή και αγνοούμε την ύπαρξη ελλειπών αλληλόμορφων εντός του δείγματος του υπό ανάλυση πληθυσμού. Με τον τρόπο αυτό υπολογίζονται οι τιμές των παραμέτρων  $D$ ,  $D'$  και  $r^2$  που σχετίζονται με τον παράγοντα της ανισορροπίας γενετικής σύνδεσης.

Στον πίνακα P φαίνονται οι 20 μεγαλύτερες τιμές ανισορροπίας σύνδεσης για τα 20 πιο συχνά παρατηρούμενα ζεύγη αλληλόμορφων των τριών εξεταζόμενων γονιδιακών τόπων HLA-A, HLA-B και HLA-DRB1, ως διαφορές μεταξύ της τιμής της παρατηρούμενης (observed) και της αναμενόμενης (expected) συχνότητας του συνδυασμού των αλληλόμορφων του κάθε ζεύγους. Όπως είναι φανερό, οι μεγαλύτερες τιμές για την παράμετρο D προκύπτουν για τα ζεύγη αλληλόμορφων των γονιδιακών τόπων HLA-B και HLA-DRB1, γεγονός που επιβεβαιώνει και την μικρή φυσική απόσταση των συγκεκριμένων HLA γονιδίων επάνω στο χρωμόσωμα 6, σε σχέση με την απόσταση των υπόλοιπων ζευγών των HLA γονιδίων.

**ΠΙΝΑΚΑΣ Ρ.** Αποτελέσματα 20 μεγαλύτερων τιμών ανισορροπίας σύνδεσης για τα 20 συχνότερα παρατηρούμενα (observed) ζεύγη αλληλόμορφων των εξεταζόμενων γονιδιακών τόπων

| HLA-A-B |       |       |                       |                       |         | HLA-B-DRB1 |       |                       |                       |         | HLA-A-DRB1 |       |                       |                       |         |
|---------|-------|-------|-----------------------|-----------------------|---------|------------|-------|-----------------------|-----------------------|---------|------------|-------|-----------------------|-----------------------|---------|
| RANK    | A     | B     | OBSERVED<br>FREQUENCY | EXPECTED<br>FREQUENCY | D       | B          | DRB1  | OBSERVED<br>FREQUENCY | EXPECTED<br>FREQUENCY | D       | A          | DRB1  | OBSERVED<br>FREQUENCY | EXPECTED<br>FREQUENCY | D       |
| 1       | 02:01 | 51:01 | 0,0526                | 0,0298                | 0,0228  | 18:01      | 11:04 | 0,0580                | 0,0170                | 0,0410  | 01:01      | 03:01 | 0,0236                | 0,0077                | 0,0159  |
| 2       | 01:01 | 08:01 | 0,0263                | 0,0046                | 0,0217  | 08:01      | 03:01 | 0,0380                | 0,0035                | 0,0345  | 02:01      | 16:01 | 0,0427                | 0,0308                | 0,0119  |
| 3       | 11:01 | 35:01 | 0,0223                | 0,0062                | 0,0161  | 52:01      | 15:02 | 0,0185                | 0,0006                | 0,0179  | 24:02      | 11:04 | 0,0340                | 0,0227                | 0,0113  |
| 4       | 02:01 | 18:01 | 0,0413                | 0,0255                | 0,0158  | 13:02      | 07:01 | 0,0181                | 0,0019                | 0,0162  | 02:01      | 14:01 | 0,0193                | 0,0128                | 0,0065  |
| 5       | 33:01 | 14:02 | 0,0119                | 0,0004                | 0,0115  | 35:01      | 01:01 | 0,0192                | 0,0038                | 0,0154  | 11:01      | 01:01 | 0,0100                | 0,0038                | 0,0062  |
| 6       | 32:01 | 40:02 | 0,0120                | 0,0015                | 0,0105  | 35:02      | 11:04 | 0,0183                | 0,0042                | 0,0141  | 02:01      | 11:04 | 0,0428                | 0,0375                | 0,0053  |
| 7       | 03:01 | 07:02 | 0,0123                | 0,0027                | 0,0096  | 14:02      | 01:02 | 0,0141                | 0,0004                | 0,0137  | 26:01      | 16:01 | 0,0102                | 0,0064                | 0,0038  |
| 8       | 30:01 | 13:02 | 0,0096                | 0,0006                | 0,0090  | 51:01      | 11:01 | 0,0240                | 0,0108                | 0,0132  | 03:01      | 11:01 | 0,0106                | 0,0074                | 0,0032  |
| 9       | 24:02 | 35:02 | 0,0128                | 0,0038                | 0,0090  | 39:01      | 16:01 | 0,0158                | 0,0034                | 0,0124  | 02:01      | 13:01 | 0,0128                | 0,0101                | 0,0027  |
| 10      | 01:01 | 37:01 | 0,0088                | 0,0014                | 0,0074  | 07:02      | 15:01 | 0,0132                | 0,0015                | 0,0117  | 03:01      | 16:01 | 0,0136                | 0,0113                | 0,0023  |
| 11      | 26:01 | 38:01 | 0,0083                | 0,0014                | 0,0069  | 44:05      | 16:01 | 0,0095                | 0,0018                | 0,0077  | 11:01      | 11:04 | 0,0140                | 0,0125                | 0,0015  |
| 12      | 02:01 | 44:02 | 0,0155                | 0,0096                | 0,0059  | 35:01      | 14:01 | 0,0116                | 0,0043                | 0,0073  | 24:02      | 11:01 | 0,0137                | 0,0123                | 0,0014  |
| 13      | 24:02 | 35:03 | 0,0118                | 0,0062                | 0,0056  | 44:03      | 07:01 | 0,0083                | 0,0012                | 0,0071  | 02:01      | 07:01 | 0,0150                | 0,0152                | -0,0002 |
| 14      | 03:01 | 35:01 | 0,0123                | 0,0068                | 0,0055  | 55:01      | 16:01 | 0,0102                | 0,0039                | 0,0063  | 32:01      | 11:04 | 0,0109                | 0,0114                | -0,0005 |
| 15      | 24:02 | 18:01 | 0,0193                | 0,0154                | 0,0039  | 44:02      | 11:04 | 0,0116                | 0,0064                | 0,0052  | 01:01      | 11:04 | 0,0151                | 0,0160                | -0,0009 |
| 16      | 32:01 | 35:01 | 0,0095                | 0,0057                | 0,0038  | 51:01      | 16:01 | 0,0204                | 0,0164                | 0,0040  | 02:01      | 11:01 | 0,0192                | 0,0203                | -0,0011 |
| 17      | 02:01 | 35:03 | 0,0095                | 0,0102                | -0,0007 | 18:01      | 15:01 | 0,0081                | 0,0051                | 0,0030  | 01:01      | 16:01 | 0,0110                | 0,0132                | -0,0022 |
| 18      | 24:02 | 51:01 | 0,0168                | 0,0180                | -0,0012 | 51:01      | 14:01 | 0,0085                | 0,0068                | 0,0017  | 02:01      | 01:01 | 0,0092                | 0,0114                | -0,0022 |
| 19      | 24:02 | 35:01 | 0,0094                | 0,0113                | -0,0019 | 51:01      | 11:04 | 0,0193                | 0,0199                | -0,0006 | 24:02      | 16:01 | 0,0153                | 0,0186                | -0,0033 |
| 20      | 02:01 | 35:01 | 0,0108                | 0,0187                | -0,0079 | 18:01      | 16:01 | 0,0092                | 0,0140                | -0,0048 | 02:01      | 03:01 | 0,0133                | 0,0180                | -0,0047 |

Στον πίνακα Σ φαίνονται οι 10 υψηλότερες θετικές και οι 10 υψηλότερες αρνητικές τιμές ανισορροπίας σύνδεσης για τα παρατηρούμενα ζεύγη αλληλόμορφων των τριών εξεταζόμενων γονιδιακών τόπων HLA-A, HLA-B και HLA-DRB1, ως διαφορές μεταξύ της τιμής της παρατηρούμενης (observed) και της αναμενόμενης (expected) συχνότητας του συνδυασμού των αλληλόμορφων του κάθε ζεύγους.

**ΠΙΝΑΚΑΣ Σ.** Αποτελέσματα 10 υψηλότερων θετικών και αρνητικών τιμών ανισορροπίας σύνδεσης για τα παρατηρούμενα (observed) ζεύγη αλληλόμορφων των εξεταζόμενων γονιδιακών τόπων

| HLA-A-B |       |        |                       |                       |         | HLA-B-DRB1 |       |                       |                       |         | HLA-A-DRB1 |       |                       |                       |         |
|---------|-------|--------|-----------------------|-----------------------|---------|------------|-------|-----------------------|-----------------------|---------|------------|-------|-----------------------|-----------------------|---------|
| RANK    | A     | B      | OBSERVED<br>FREQUENCY | EXPECTED<br>FREQUENCY | D>0     | B          | DRB1  | OBSERVED<br>FREQUENCY | EXPECTED<br>FREQUENCY | D>0     | A          | DRB1  | OBSERVED<br>FREQUENCY | EXPECTED<br>FREQUENCY | D>0     |
| 1       | 02:01 | 51:01  | 0,0526                | 0,0298                | 0,0228  | 18:01      | 11:04 | 0,0580                | 0,0170                | 0,0410  | 01:01      | 03:01 | 0,0236                | 0,0077                | 0,0159  |
| 2       | 01:01 | 08:01  | 0,0263                | 0,0046                | 0,0217  | 08:01      | 03:01 | 0,0380                | 0,0035                | 0,0345  | 02:01      | 16:01 | 0,0427                | 0,0308                | 0,0119  |
| 3       | 11:01 | 35:01  | 0,0223                | 0,0062                | 0,0161  | 52:01      | 15:02 | 0,0185                | 0,0006                | 0,0179  | 24:02      | 11:04 | 0,0340                | 0,0227                | 0,0113  |
| 4       | 02:01 | 18:01  | 0,0413                | 0,0255                | 0,0158  | 13:02      | 07:01 | 0,0181                | 0,0019                | 0,0162  | 33:01      | 01:02 | 0,0078                | 0,0003                | 0,0075  |
| 5       | 33:01 | 14:02  | 0,0119                | 0,0004                | 0,0115  | 35:01      | 01:01 | 0,0192                | 0,0038                | 0,0154  | 32:01      | 16:02 | 0,0090                | 0,0015                | 0,0075  |
| 6       | 32:01 | 40:02  | 0,0120                | 0,0015                | 0,0105  | 35:02      | 11:04 | 0,0183                | 0,0042                | 0,0141  | 02:01      | 14:01 | 0,0193                | 0,0128                | 0,0065  |
| 7       | 03:01 | 07:02  | 0,0123                | 0,0027                | 0,0096  | 14:02      | 01:02 | 0,0141                | 0,0004                | 0,0137  | 11:01      | 01:01 | 0,0100                | 0,0038                | 0,0062  |
| 8       | 30:01 | 13:02  | 0,0096                | 0,0006                | 0,0090  | 51:01      | 11:01 | 0,0240                | 0,0108                | 0,0132  | 02:01      | 11:04 | 0,0428                | 0,0375                | 0,0053  |
| 9       | 24:02 | 35:02  | 0,0128                | 0,0038                | 0,0090  | 39:01      | 16:01 | 0,0158                | 0,0034                | 0,0124  | 24:02      | 04:03 | 0,0078                | 0,0039                | 0,0039  |
| 10      | 23:01 | 49:01  | 0,0082                | 0,0007                | 0,0075  | 07:02      | 15:01 | 0,0132                | 0,0015                | 0,0117  | 26:01      | 16:01 | 0,0102                | 0,0064                | 0,0038  |
| RANK    | A     | B      | OBSERVED<br>FREQUENCY | EXPECTED<br>FREQUENCY | D<0     | B          | DRB1  | OBSERVED<br>FREQUENCY | EXPECTED<br>FREQUENCY | D<0     | A          | DRB1  | OBSERVED<br>FREQUENCY | EXPECTED<br>FREQUENCY | D<0     |
| 1       | 29:02 | 15:01  | 0,0000                | 0,1000                | -0,1000 | 15:03      | 15:01 | 0,0002                | 0,0998                | -0,0996 | 69:01      | 10:01 | 0,0001                | 0,0974                | -0,0973 |
| 2       | 25:01 | 41:01  | 0,0000                | 0,0986                | -0,0986 | 14:02      | 08:04 | 0,0001                | 0,0986                | -0,0985 | 26:08      | 14:01 | 0,0005                | 0,0970                | -0,0965 |
| 3       | 02:02 | 07:02  | 0,0000                | 0,0973                | -0,0973 | 73:01      | 04:03 | 0,0000                | 0,0976                | -0,0976 | 68:02      | 15:02 | 0,0001                | 0,0962                | -0,0961 |
| 4       | 74:01 | 18:01  | 0,0003                | 0,0968                | -0,0965 | 44:03      | 14:04 | 0,0000                | 0,0975                | -0,0975 | 02:02      | 13:02 | 0,0001                | 0,0958                | -0,0957 |
| 5       | 30:01 | 40:06  | 0,0001                | 0,0965                | -0,0964 | 57:01      | 04:01 | 0,0000                | 0,0973                | -0,0973 | 66:01      | 10:01 | 0,0001                | 0,0936                | -0,0935 |
| 6       | 29:02 | 44:05  | 0,0002                | 0,0966                | -0,0964 | 27:05      | 04:01 | 0,0003                | 0,0973                | -0,0970 | 24:03      | 10:01 | 0,0002                | 0,0936                | -0,0934 |
| 7       | 33:03 | 41:01  | 0,0001                | 0,0959                | -0,0958 | 40:01      | 08:01 | 0,0005                | 0,0964                | -0,0959 | 25:01      | 13:03 | 0,0001                | 0,0934                | -0,0933 |
| 8       | 03:01 | 15:24  | 0,0000                | 0,0954                | -0,0954 | 44:25      | 11:04 | 0,0000                | 0,0949                | -0,0949 | 69:01      | 04:02 | 0,0000                | 0,0933                | -0,0933 |
| 9       | 03:01 | 18:18  | 0,0000                | 0,0954                | -0,0954 | 51:05      | 11:04 | 0,0000                | 0,0949                | -0,0949 | 32:01      | 04:08 | 0,0001                | 0,0933                | -0,0932 |
| 10      | 03:01 | 44:182 | 0,0001                | 0,0954                | -0,0953 | 55:01      | 09:01 | 0,0002                | 0,0950                | -0,0948 | 01:02      | 13:01 | 0,0012                | 0,0933                | -0,0921 |

### ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΙΣΟΡΡΟΠΙΑΣ ΣΥΝΔΕΣΗΣ ARLEQUIN 3.5

#### Project information:

```

NbSamples      = 1
DataType       = STANDARD
GenotypicData  = 1
GameticPhase   = 1
RecessiveData  = 0

```

#### Settings used for Calculations

#### General settings:

```

Deletion Weight      = 1
Transition Weight Weight = 1
Tranversion Weight Weight = 1
Epsilon Value        = 1e-07
Significant digits for output = 5
Use original haplotype definition
Allowed level of missing data = 0.05

```

#### Active Tasks:

#### Linkage disequilibrium test:

```

No. of steps in Markov chain      = 10000
No. of Dememorisation Steps       = 1000
Required precision on Probability = 0

```

D and D' coefficients for all pairs of alleles : Yes

Print Histogram and table : No

=====

== ANALYSES AT THE INTRA-POPULATION LEVEL

=====

=====

== Sample :      Greeks

=====

=====

== Pairwise linkage disequilibrium : (Greeks)

=====

Slatkin, M. 1994a.

Slatkin, M. and Excoffier, L. 1996.

Lewontin, R. C., and K. Kojima 1960.

Test of linkage disequilibrium for all pairs of loci:

-----

Exact test using a Markov chain :

Chain length       :10000

Dememorization    :1000

Note: The test is only performed for polymorphic pairs of loci

Pair (0, 1) Exact P=0.00425743 +- 0.000492643 (10100 Steps done)

Pair (0, 2) Exact P=0 +- 0 (10100 Steps done)

Pair (1, 2) Exact P=0 +- 0 (10100 Steps done)

Table of significant linkage disequilibrium (significance level=0.05):

|         |  |   |   |   |
|---------|--|---|---|---|
| -----   |  |   |   |   |
| Locus # |  | 0 | 1 | 2 |
| -----   |  |   |   |   |
| 0       |  | * | + | + |
| 1       |  | + | * | + |
| 2       |  | + | + | * |

Number of linked loci per polymorphic locus (significance level=0.05):

-----

y\Locus:       0     1     2    No. of Loci

|       |   |   |   |   |
|-------|---|---|---|---|
| ----- |   |   |   |   |
| 1     | 2 | 2 | 2 | 3 |
| ----- |   |   |   |   |
| 2     | 2 | 2 | 2 | 3 |
| ----- |   |   |   |   |
| 3     | 2 | 2 | 2 | 3 |
| ----- |   |   |   |   |
| 4     | 2 | 2 | 2 | 3 |
| ----- |   |   |   |   |
| 5     | 2 | 2 | 2 | 3 |
| ----- |   |   |   |   |

Percentage of linked loci per locus (significance level=0.05):

| y\Locus: | 0     | 1     | 2     | No. of Loci |
|----------|-------|-------|-------|-------------|
| 1        | 100.0 | 100.0 | 100.0 | 3           |
| 2        | 100.0 | 100.0 | 100.0 | 3           |
| 3        | 100.0 | 100.0 | 100.0 | 3           |
| 4        | 100.0 | 100.0 | 100.0 | 3           |
| 5        | 100.0 | 100.0 | 100.0 | 3           |

-----  
Loci 0 and 1  
-----

1: Observed contingency table

. . . . .

2: Expected table

. . . . .

3: Table of disequilibrium values ( $D = p_{ab} - p_a p_b$ ) for all two-locus haplotypes

. . . . .

4: Table of standardized disequilibrium values ( $D' = D/D_{max}$ )

. . . . .

5: Table of standardized disequilibrium values ( $r^2$ )

. . . . .

6: Table of Chi-square values (  $\chi^2 = n \cdot (p_a \cdot (1-p_a) \cdot p_b \cdot (1-p_b))$  )

. . . . .

7: Table of Chi-square P values (1 d.f.)

-----  
Loci 0 and 2  
-----

. . . . .

-----  
Loci 1 and 2  
-----

. . . . .

### 5.2.5.6 Ανάλυση Πιθανοτήτων Ιστοσυμβατότητας & Βελτιστοποίηση Μητρώου

Κάθε άτομο του υπό ανάλυση πληθυσμού του μητρώου Ελλήνων εθελοντών δοτών του ΚΕΔΜΟΠ διαθέτει ένα σύνολο χαρακτηριστικών, έτσι ώστε ένα μόσχευμα να είναι δυνατόν να μεταμοσχευθεί μεταξύ ατόμων που έχουν τα ίδια χαρακτηριστικά. Συγκεκριμένα, ο τύπος των χαρακτηριστικών αυτών θεωρούμε ότι είναι ο HLA φαινότυπος κάθε ατόμου, ο οποίος προσδιορίζεται από τα αλληλόμορφα που διαθέτει κάθε άτομο για κάθε έναν από τους τρεις τυποποιημένους σε υψηλή ανάλυση γονιδιακούς τόπους HLA-A, HLA-B και HLA-DRB1.

Για τον υπολογισμό των πιθανοτήτων ιστοσυμβατότητας και τη βελτιστοποίηση του μητρώου θεωρούμε ότι ισχύουν τα εξής :

- Όλοι οι HLA φαινότυποι είναι διαφορετικοί μεταξύ τους, ώστε κάθε δότης διαθέτει έναν ξεχωριστό τύπο χαρακτηριστικών.
- Κάθε συμβατός δότης με έναν ασθενή που θα χρειαστεί μεταμόσχευση είναι πλήρως συμβατός.
- Δότες και λήπτες είναι τυχαία κατανεμημένοι εντός του ίδιου πληθυσμού.

Με βάση τα παραπάνω, ορίζεται ο παρακάτω μαθηματικός υπολογισμός ο οποίος αποτελεί την **πιθανότητα εύρεσης τουλάχιστον ενός πλήρως συμβατού δότη για τον επόμενο ασθενή του μητρώου που θα χρειαστεί μόσχευμα.**

$$p(n) = \sum_{i=1}^P f_i \left[ 1 - (1 - f_i)^n \right] \quad (24)$$

όπου :

$P$  είναι το αρχικό μέγεθος του πλήθους των 5.131 εθελοντών δοτών του μητρώου.

$n$  είναι το εξεταζόμενο μέγεθος του μητρώου των εθελοντών δοτών.

$f_i$  είναι η σχετική συχνότητα του  $i$ -οστού HLA φαινοτύπου του μητρώου.

Εκτός από την πιθανότητα ιστοσυμβατότητας  $p(n)$ , επίσης υπολογίζουμε και την παράγωγό της :

$$p'(n) = - \sum_{i=1}^P f_i \cdot \ln(1 - f_i) (1 - f_i)^n \quad (25)$$

Κατόπιν, ορίζεται η παρακάτω ποσότητα, η οποία αποτελεί *το πλήθος των νέων δοτών που απαιτείται να εισαχθούν στο υπό ανάλυση μητρώο μεγέθους  $n$  , έτσι ώστε η πιθανότητα ισοσυμβατότητας  $p(n)$  για τον επόμενο ασθενή του μητρώου να αυξηθεί κατά 1%.*

$$d(n) = \frac{1}{100 \cdot p'(n)} \quad (26)$$

Παρακάτω παρουσιάζεται ο κώδικας υλοποίησης και το αποτέλεσμα της κλήσης συνάρτησης σε Matlab που υπολογίζει την πιθανότητα  $p(n)$  για  $n = 5.131$

#### **ΚΩΔΙΚΑΣ ΣΥΝΑΡΤΗΣΗΣ ΥΠΟΛΟΓΙΣΜΟΥ ΤΗΣ ΠΟΣΟΤΗΤΑΣ $p(n)$**

```
function x = f1(f,n)
% Η συνάρτηση f1 δέχεται ως ορίσματα το διάνυσμα των συχνοτήτων των
% φαινοτύπων του μητρώου και το μέγεθος του μητρώου.

p=0;
for (i=1:n)
    p=sum((f(i))*(1-(1-f(i))^n))+p;
end
p
end
```

#### **ΕΙΣΑΓΩΓΗ ΔΕΔΟΜΕΝΩΝ ΣΤΟ COMMAND WINDOW ΤΟΥ MATLAB**

```
>> f=importdata('f.txt');
>> n=5131;
```

#### **ΚΛΗΣΗ ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ f1 ΜΕ ΟΡΙΣΜΑΤΑ ΤΑ ΔΕΔΟΜΕΝΑ & ΑΠΟΤΕΛΕΣΜΑΤΑ**

```
>> f1(f,n)

p =

    0.0517
```

Παρακάτω παρουσιάζεται ο κώδικας υλοποίησης και το αποτέλεσμα της κλήσης συνάρτησης σε Matlab που υπολογίζει την ποσότητα  $d(n)$  για  $n = 5.131$

#### **ΚΩΔΙΚΑΣ ΣΥΝΑΡΤΗΣΗΣ ΥΠΟΛΟΓΙΣΜΟΥ ΤΗΣ ΠΟΣΟΤΗΤΑΣ $d(n)$**

```
function x = f2(f,n)
% Η συνάρτηση f2 δέχεται ως ορίσματα το διάνυσμα των συχνοτήτων των
% φαινοτύπων του μητρώου και το μέγεθος του μητρώου.

p2=0;
for (i=1:n)
    p2=-sum((f(i))*(log(1-f(i)))*(1-f(i))^n)+p2;
end

d=1/(100*p2)
end
```

## ΚΛΗΣΗ ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ f2 ΜΕ ΟΡΙΣΜΑΤΑ ΤΑ ΔΕΔΟΜΕΝΑ & ΑΠΟΤΕΛΕΣΜΑΤΑ

```
>> f2(f,n)
```

```
d =
```

```
2.7960e+003
```

Άρα, με βάση τα διαθέσιμα δεδομένα του μητρώου τα αποτελέσματα των υπολογισμών έχουν ως εξής :

$$p(5.131) = 0,0517$$

$$d(5.131) = 2.796$$

Σε μια προσπάθεια γενίκευσης των αποτελεσμάτων της έως τώρα ανάλυσης και βελτιστοποίησης των μεθοδολογιών σχεδιασμού μητρώων εθελοντών δοτών μυελού των οστών, οι τιμές των πιθανοτήτων ιστοσυμβατότητας  $p(n)$  μπορούν να εκτιμηθούν και για διαφορετικά μεγέθη μητρώων, όπου το εξεταζόμενο μέγεθος μητρώου  $n$  είναι μεγαλύτερο από το διαθέσιμο μέγεθος του μητρώου των 5.131 εθελοντών δοτών. Ωστόσο, η αξιοπιστία των εκτιμήσεων εξαρτάται σε κάθε περίπτωση από τα χαρακτηριστικά των HLA δεδομένων του κάθε μητρώου, το επίπεδο της ανάλυσης της HLA τυποποίησης των γονιδίων και από το μέγεθος του μητρώου.

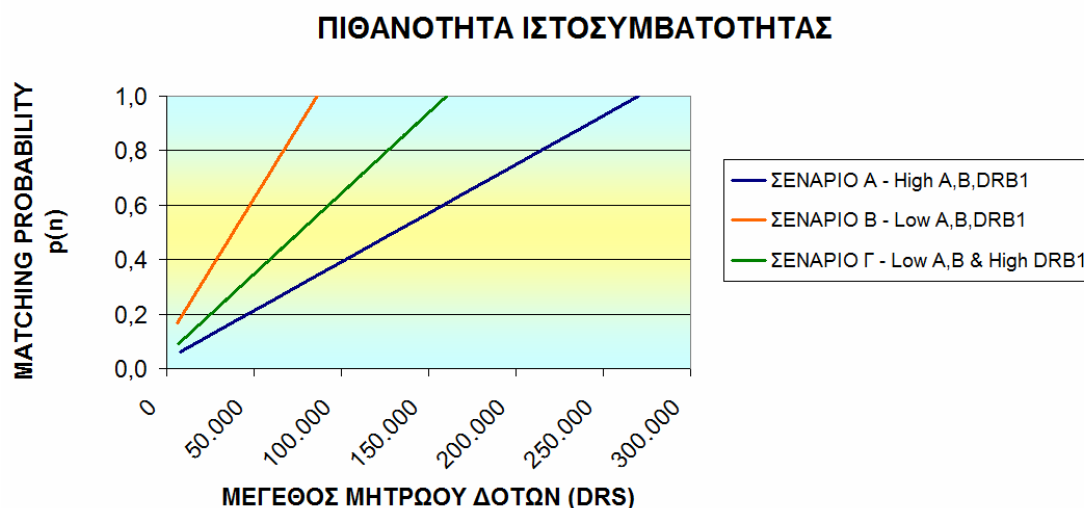
Αρχικά, ξεκινώντας από τα ως άνω αποτελέσματα των υπολογισμών των μεγεθών  $p(n)$  και  $d(n)$  για  $n = 5.131$ , υπολογίσαμε τις απαιτήσεις του αρχικού μητρώου για προσθήκη νέων δοτών σε αυτό, έτσι ώστε η πιθανότητα ιστοσυμβατότητας  $p(n)$  να αυξηθεί σταδιακά με μια επιλογή βήματος 0,001 (0,1%), από την τιμή  $p(5.131)$  έως την ιδανική τιμή 1 (100%). Οι υπολογισμοί έγιναν και για τα τρία διαφορετικά σενάρια Α, Β και Γ του διαφορετικού επιπέδου HLA τυποποίησης των γονιδίων HLA-A, HLA-B και HLA-DRB1, όπως αυτά έχουν ήδη οριστεί.

Ξεκινώντας από το αρχικό μέγεθος του μητρώου των 5.131 δοτών, μέσω των υπολογισμών προκύπτουν οι εκτιμήσεις τιμών για τα τρία σενάρια όπως παρουσιάζονται στον πίνακα Γ.

**ΠΙΝΑΚΑΣ Τ.** Εκτιμήσεις τιμών των μεγεθών  $p(n)$  και  $d(n)$  για τα τρία σενάρια της HLA ανάλυσης για το αρχικό μητρώο των  $n = 5.131$  εθελοντών δοτών

| $n = 5.131$                                                |        |        |
|------------------------------------------------------------|--------|--------|
| ΣΕΝΑΡΙΟ                                                    | $p(n)$ | $d(n)$ |
| (Α) High Resolution HLA-A, HLA-B, HLA-DRB1                 | 0,0517 | 2.796  |
| (Β) Low Resolution HLA-A, HLA-B, HLA-DRB1                  | 0,1569 | 960    |
| (Γ) Low Resolution HLA-A, HLA-B & High Resolution HLA-DRB1 | 0,0809 | 1.689  |

Στην εικόνα 5.16 φαίνεται η γραφική παράσταση της πιθανότητας ιστοσυμβατότητας  $p(n)$  ως προς το μέγεθος του μητρώου των εθελοντών δοτών (DRS - Donor Registry Size) για τα τρία διαφορετικά σενάρια.



**ΕΙΚΟΝΑ 5.16** Γραφική παράσταση της πιθανότητας ιστοσυμβατότητας  $p(n)$  ως προς το μέγεθος του μητρώου των εθελοντών δοτών (DRS)

Από τα αποτελέσματα των υπολογισμών, είναι φανερό πως οι απαιτήσεις για προσθήκη νέων δοτών στο μητρώο, έτσι ώστε να αυξηθεί το μέγεθός του και κατ' επέκταση η πιθανότητα ιστοσυμβατότητας  $p(n)$ , είναι μεγαλύτερες για το αυστηρότερο, από πλευράς HLA τυποποίησης των γονιδίων, σενάριο Α. Αντιθέτως, η τιμή της πιθανότητας  $p(n)$  προσεγγίζει το 100% σε μικρότερα μεγέθη μητρώου, εφόσον οι απαιτήσεις για προσθήκη νέων δοτών είναι μικρότερες.

Η σχέση της πιθανότητας  $p(n)$  και του μεγέθους του μητρώου είναι, όπως φαίνεται, σχεδόν γραμμική και για κάθε αύξηση της τιμής της πιθανότητας  $p(n)$  κατά 0,1% , οι

απαιτήσεις προσθήκης νέων δοτών κυμαίνονται στην τιμή 280 για το σενάριο Α, στην τιμή 97 για το σενάριο Β και στην τιμή 169 για το σενάριο Γ. Για την προσέγγιση της τιμής της πιθανότητας  $p(n)$  στο 100%, το μέγεθος του μητρώου από τις απαιτήσεις προσθήκης νέων δοτών κυμαίνεται στην τιμή 270.191 για το σενάριο Α, στην τιμή 86.096 για το σενάριο Β και στην τιμή 160.331 για το σενάριο Γ.

## **ΕΝΑΛΛΑΚΤΙΚΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ – ΜΗ ΓΡΑΜΜΙΚΑ ΜΟΝΤΕΛΑ ΑΝΑΛΥΣΗΣ**

### **1. Διακριτό μη γραμμικό μοντέλο**

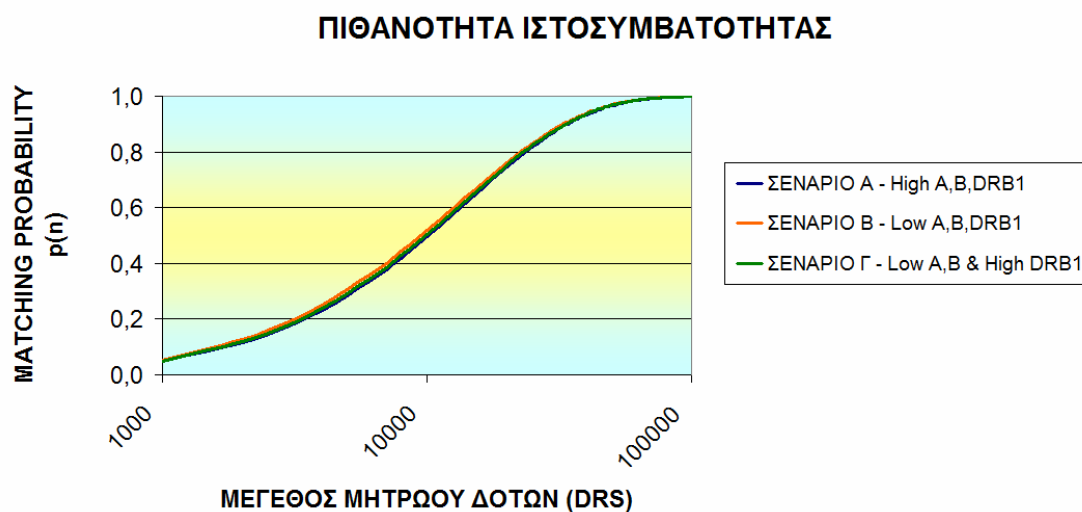
Στο σημείο αυτό, θα γίνει μια προσπάθεια υπολογισμού των πιθανοτήτων ιστοσυμβατότητας  $p(n)$  ως προς το μέγεθος του μητρώου των δοτών (DRS), με βάση μια εναλλακτική προσέγγιση όπως παρουσιάζεται στο [19]. Η εναλλακτική αυτή προσέγγιση επικεντρώνεται σε μια διαφορετική ανάλυση των συχνοτήτων των φαινοτύπων των ατόμων του μητρώου. Ενώ στην προηγούμενη προσέγγιση οι τιμές των συχνοτήτων των φαινοτύπων που χρησιμοποιήθηκαν ήταν οι σχετικές συχνότητες των 5.131 φαινοτύπων, όπως αυτές προέκυψαν από τις εκτιμήσεις της ανάλυσης του πακέτου Arlequin, στην εναλλακτική προσέγγιση υπολογίζουμε αρχικά το πλήθος των διαφορετικών (unique) φαινοτύπων με ακρίβεια 8/8 για το υπό ανάλυση μητρώο. Έτσι, για κάθε μελετούμενο σενάριο της HLA ανάλυσης, το μητρώο καθορίζεται από τους φαινοτύπους  $1, \dots, J$  όπου ο καθένας από αυτούς έχει μια συχνότητα  $f_i > 0$  και ισχύει η σχέση  $\sum_{i=1}^J f_i = 1$ . Όσον αφορά το ποσοστό της συμβατότητας δότη-λήπτη φαινοτύπου  $i$ , αυτό αντιστοιχεί σε μια πιθανότητα  $a$ .

Αν θεωρήσουμε ότι δεν υπάρχει αρχικό stock δοτών, η πιθανότητα ιστοσυμβατότητας ορίζεται ως εξής :

$$p(n) = 1 - \sum_{i=1}^J f_i \cdot e^{-anf_i} \quad (27)$$

$$\text{με } \sum_{i=1}^J f_i = 1 \quad \text{και} \quad f_i \geq 0 \quad \forall i = 1, \dots, J$$

Μια ρεαλιστική τιμή για την μεταβλητή  $a$  είναι  $a = \frac{1}{3}$ , δηλαδή το  $\frac{1}{3}$  των δοτών που είναι συμβατοί σε επίπεδο μητρώου, να είναι πλήρως συμβατοί και διαθέσιμοι για μεταμόσχευση (France Greffe de Moelle Registry). Με βάση τα παραπάνω, υπολογίζουμε την τιμή της πιθανότητας  $p(n)$  για αυξανόμενες τιμές του μεγέθους του μητρώου των δοτών  $n$ . Το πλήθος των διαφορετικών φαινοτύπων με ακρίβεια 8/8 για τα τρία διαφορετικά σενάρια της HLA ανάλυσης είναι 4.925 φαινότυποι για το σενάριο Α, 4.714 φαινότυποι για το σενάριο Β και 4.847 φαινότυποι για το σενάριο Γ. Οι υπολογιζόμενες τιμές για την πιθανότητα  $p(n)$  ως προς το μέγεθος του μητρώου κυμαίνονται στα ίδια επίπεδα και για τα τρία σενάρια. Σε σχέση με το γραμμικό μοντέλο, παρατηρείται ότι για την προσέγγιση της τιμής της πιθανότητας  $p(n)$  στο 100% το μέγεθος του μητρώου των δοτών κυμαίνεται γύρω από την τιμή 148.000, η οποία βρίσκεται στα επίπεδα του μέσου όρου των αντίστοιχων τιμών των πιθανοτήτων  $p(n)$  του γραμμικού μοντέλου και οι τιμές των πιθανοτήτων  $p(n)$  για το μέγεθος του μητρώου των 5.131 είναι αυξημένες σε σχέση με το γραμμικό μοντέλο, αφού κυμαίνονται στα επίπεδα του 30%.



**ΕΙΚΟΝΑ 5.17** Γραφική παράσταση της πιθανότητας ιστοσυμβατότητας  $p(n)$  ως προς το μέγεθος του μητρώου των εθελοντών δοτών (DRS) σε λογαριθμική κλίμακα του άξονα x, σύμφωνα με το διακριτό μη γραμμικό μοντέλο.

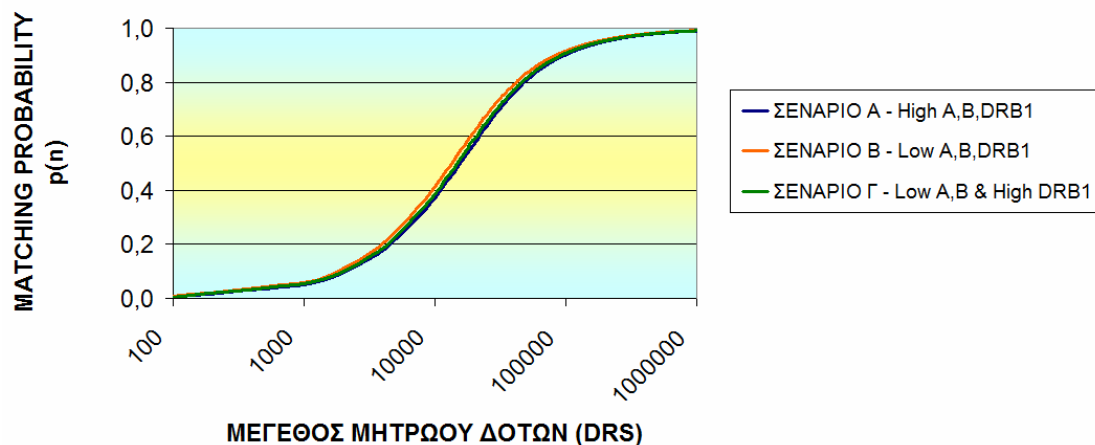
## 2. Συνεχές μη γραμμικό μοντέλο

Τέλος, ένα πιο γενικευμένο μοντέλο ανάλυσης των πιθανοτήτων ιστοσυμβατότητας  $p(n)$  ως προς το μέγεθος του μητρώου των δοτών, όπως παρουσιάζεται στο [19], θεωρεί ότι παρόλο που το πλήθος των διαφορετικών φαινοτύπων που υπάρχουν σε ένα μητρώο δοτών και κατά συνέπεια και των συχνοτήτων τους είναι πεπερασμένο (διακριτό), μπορούμε να διαχειριστούμε τους φαινοτύπους και τις συχνότητές τους ως μια συνεχή κατανομή θετικών τιμών. Ακριβέστερα, θεωρούμε ότι οι φαινότυποι προκύπτουν τυχαία από μια εκθετική κατανομή πυκνότητας πιθανότητας  $\lambda \cdot e^{-\lambda i}$  ( $i \in [0, +\infty]$ ), ώστε η τιμή της πιθανότητας ιστοσυμβατότητας  $p(n)$  να υπολογίζεται ως εξής :

$$p(n) = 1 - \frac{1 - e^{-an\lambda}}{an\lambda} \quad (28)$$

όπου η τιμή της μεταβλητής  $\lambda$  προκύπτει από τον ενδιάμεσο της κατανομής και ορίζεται ως  $\lambda = -\frac{1}{X} \ln \frac{1}{2}$ , όπου το πλήθος  $X$  των συχνότερων φαινοτύπων αναπαριστά αθροιστικά το 50% του συνόλου των φαινοτύπων του πληθυσμού. Η τιμή της μεταβλητής  $X$  υπολογίστηκε ως 2.265 φαινότυποι (από τους 4.925) για το σενάριο Α, 1.949 φαινότυποι (από τους 4.714) για το σενάριο Β και 2.148 φαινότυποι (από τους 4.847) για το σενάριο Γ.

### ΠΙΘΑΝΟΤΗΤΑ ΙΣΤΟΣΥΜΒΑΤΟΤΗΤΑΣ



**ΕΙΚΟΝΑ 5.18** Γραφική παράσταση της πιθανότητας ιστοσυμβατότητας  $p(n)$  ως προς το μέγεθος του μητρώου των εθελοντών δοτών (DRS) σε λογαριθμική κλίμακα του άξονα x, σύμφωνα με το συνεχές μη γραμμικό μοντέλο.

Σε σχέση με το διακριτό μη γραμμικό μοντέλο, παρατηρείται ότι για την προσέγγιση της τιμής της πιθανότητας  $p(n)$  στο 100%, όταν το μέγεθος του μητρώου των δοτών κυμαίνεται γύρω από την τιμή 148.000, η τιμή της πιθανότητας  $p(n)$  βρίσκεται στα επίπεδα του 94%, αφήνοντας περιθώρια μικρής αύξησης για ακόμη μεγαλύτερα μεγέθη μητρώου. Όσον αφορά, το μέγεθος του αρχικού μας μητρώου, για τις τιμές του μεγέθους του μητρώου γύρω από την τιμή 5.131 του αρχικού μητρώου, η αντίστοιχη τιμή της πιθανότητας  $p(n)$  είναι ελαφρώς μειωμένη και για τα τρία σενάρια στα επίπεδα του 23%.

# **Κ Ε Φ Α Λ Α Ι Ο 6<sup>ο</sup>**

## **ΣΥΜΠΕΡΑΣΜΑΤΑ & ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ**

Στο κεφάλαιο αυτό παρουσιάζονται τα βασικότερα συμπεράσματα της παρούσας διπλωματικής εργασίας και οι προκλήσεις που προέκυψαν κατά τη διερεύνηση και τη μελέτη του αντικειμένου της εργασίας, καθώς επίσης προτείνονται και ορισμένες επιπρόσθετες μελλοντικές επεκτάσεις του έργου που πραγματοποιήθηκε στα πλαίσια της εργασίας.

### **6.1 Σύνοψη & Συμπεράσματα**

Στην παρούσα εργασία, εφαρμόστηκαν καθιερωμένες μέθοδοι για τον υπολογισμό των κατανομών των συχνοτήτων των HLA αντιγόνων και αλληλόμορφων και για την εκτίμηση των συχνοτήτων των HLA αplotύπων και των πιθανοτήτων ιστοσυμβατότητας μεταξύ των δοτών που είναι εγγεγραμμένοι στο μητρώο των Ελλήνων εθελοντών δοτών μυελού των οστών του ΚΕΔΜΟΠ, συμπεριλαμβανομένων υπολογισμών που υλοποιήθηκαν για τον προσδιορισμό συγκεκριμένων μετρικών βιοστατιστικής ανάλυσης HLA δεδομένων. Ως γενικό αποτέλεσμα, μπορούμε να συμπεράνουμε τα σημαντικά οφέλη της ενδοπληθυσμιακής προσέλευσης εθελοντών δοτών αιμοποιητικών βλαστικών κυττάρων. Το μοντέλο που χρησιμοποιήθηκε στις αναλύσεις βασίστηκε στην υπόθεση ότι τα HLA δεδομένα των δοτών οι οποίοι είναι πλήρως συμβατοί σε ασθενείς που έχουν ανάγκη από μεταμόσχευση θα μπορούν να προσδιοριστούν κατά τη διαδικασία αναζήτησης συμβατού δότη. Ωστόσο, υπάρχουν αποδείξεις ότι πλήρως συμβατοί δότες μπορούν να παραμένουν άγνωστοι κατά τη διαδικασία, λόγω ελλειπών HLA στοιχείων. Όσον αφορά τα ποιοτικά και ποσοτικά χαρακτηριστικά του δείγματος του μητρώου που χρησιμοποιήθηκε κατά την διενέργεια των αναλύσεων, καθώς ο προσδιορισμός του βαθμού ιστοσυμβατότητας κατά τη μεταμόσχευση μυελού των οστών βασίζεται στην υψηλής ανάλυσης HLA τυποποίηση των γονιδίων HLA-A, HLA-B και HLA-DRB1, το δείγμα του μητρώου των 5.131 Ελλήνων

εθελοντών δοτών του ΚΕΔΜΟΠ διαθέτει επαρκή συνάφεια ως προς την αναζήτηση μεμονωμένων συμβατών δοτών αλλά και ως προς τις εφαρμογές στρατηγικού σχεδιασμού μητρώων δοτών. Σχετικά με την ποικιλομορφία των δεδομένων του μητρώου, το γονίδιο HLA-B είναι το πιο πολυμορφικό ως προς το πλήθος των διαφορετικών αντιγόνων και αλληλόμορφων που διαθέτει με συχνότερο αντιγόνο το HLA-B\*35 με συχνότητα 17,355% και με συχνότερο αλληλόμορφο το HLA-B\*51:01 με συχνότητα 12,442%. Αντίστοιχα, για το γονίδιο HLA-A, συχνότερο είναι το αντιγόνο HLA-A\*02 με συχνότητα 26,092% και το αλληλόμορφο HLA-A\*02:01 με συχνότητα 23,4% και για το γονίδιο HLA-DRB1, συχνότερο είναι το αντιγόνο HLA-DRB1\*11 με συχνότητα 25,764% και το αλληλόμορφο HLA-DRB1\*11:04 με συχνότητα 15,676%. Το πλήθος των διαφορετικών απλοτύπων που προέκυψαν είναι 9.962 με αθροιστική συχνότητα των 100 συχνότερων απλοτύπων ίση με 36,937%. Συχνότερος είναι ο απλότυπος A\*02:01, B\*18:01, DRB1\*11:04 με συχνότητα 2,634%. Από τα αποτελέσματα είναι φανερό πως υπάρχει μεγάλο πλήθος απλοτύπων, οι οποίοι έχουν πολύ μικρή συχνότητα ή παρατηρούνται κατά μοναδικότητα στα άτομα του δείγματος. Η ανθρωπογεωγραφία του δείγματος μπορεί να ερμηνεύσει τα αποτελέσματα των αναλύσεων, καθώς τα περισσότερα άτομα που είναι εγγεγραμμένα στο μητρώο του ΚΕΔΜΟΠ προέρχονται από την Δυτική Ελλάδα. Από την αύξηση του μεγέθους του μητρώου, με παράλληλη επέκταση των δράσεων προσέλκυσης δοτών στο σύνολο των περιοχών της ηπειρωτικής αλλά και της νησιωτικής Ελλάδας, τα ποιοτικά χαρακτηριστικά του μητρώου μπορούν να γίνουν ακόμη ισχυρότερα. Επίσης, η αθροιστική συχνότητα των συχνότερων απλοτύπων μπορεί να ενισχυθεί. Το μεγάλο πλήθος των μοναδικών απλοτύπων, φανερώνει ότι στην πορεία του χρόνου υπήρξε έντονη η μετακίνηση του Ελληνικού πληθυσμού και η εξέλιξή του ήταν πολύπλοκη. Από τον έλεγχο ισορροπίας Hardy-Weinberg που διενεργήθηκε επί των δεδομένων του μητρώου είναι φανερό πως το συγκεκριμένο δείγμα του Ελληνικού πληθυσμού διαθέτει υψηλό βαθμό ετερωζυγωτίας 87,7% για το γονίδιο HLA-A, 94,1% για το γονίδιο HLA-B και 92,4% για το γονίδιο HLA-DRB1 και ότι το συγκεκριμένο πληθυσμιακό δείγμα δεν βρίσκεται σε ισορροπία Hardy-Weinberg για κανέναν HLA γονιδιακό τόπο. Αυτό σημαίνει πως ο πληθυσμός του δείγματος δεν χαρακτηρίζεται από τυχαίο ζευγάρωμα και οι υπό μελέτη γονιδιακοί τόποι εμφανίζουν μεταβαλλόμενο ρυθμό μετάλλαξης. Επίσης, όπως προαναφέρθηκε, είναι πιθανόν να έχουν προκύψει αλλαγές στη δομή του πληθυσμού λόγω μετανάστευσης, γεγονός που προκάλεσε μεταβολές και στις γονιδιακές συχνότητες από τη μια γενεά στην επόμενη. Έχει αποδειχθεί ότι οι αποκλίσεις από την ισορροπία Hardy-Weinberg

δεν επηρεάζουν αρνητικά την ακρίβεια των εκτιμήσεων των συχνοτήτων των απλοτύπων με τον αλγόριθμο EM. Σχετικά τις τιμές της ανισορροπίας γενετικής σύνδεσης (LD), οι μεγαλύτερες τιμές για την παράμετρο D προκύπτουν για τα ζεύγη αλληλόμορφων των γονιδιακών τόπων HLA-B και HLA-DRB1, γεγονός που επιβεβαιώνει και την μικρή φυσική απόσταση των συγκεκριμένων HLA γονιδίων επάνω στο χρωμόσωμα 6, σε σχέση με την απόσταση των υπόλοιπων ζευγών των HLA γονιδίων. Συγκεκριμένα, το ζεύγος B\*18:01, DRB1\*11:04 έχει την υψηλότερη θετική τιμή παραμέτρου D ίση με 0,0410 και το ζεύγος B\*15:03, DRB1\*15:01 έχει την υψηλότερη αρνητική τιμή παραμέτρου D ίση με -0,0996. Αυτό σημαίνει πως το πρώτο ζεύγος εμφανίζεται ως συνδυασμός πιο συχνά από όλα τα άλλα ζεύγη στον πληθυσμό απ' ό,τι θα αναμενόταν από μια τυχαία διαμόρφωση των απλοτύπων από τα αλληλόμορφα, με βάση τις συχνότητές τους, ενώ αντίθετα το δεύτερο ζεύγος εμφανίζεται λιγότερο συχνά από όλα τα άλλα ζεύγη. Τα συγκεκριμένα αλληλόμορφα γονίδια έχουν την τάση να κληρονομούνται μαζί, ως ζεύγη και να παραμένουν γενετικά συνδεδεμένα, χωρίς να διαχωρίζονται κατά την μείωση. Αντίστοιχα ζεύγη αλληλόμορφων για τα γονίδια HLA-A και HLA-B αποτελούν το ζεύγος A\*02:01, B\*51:01 με την υψηλότερη θετική τιμή παραμέτρου D ίση με 0,0228 και το ζεύγος A\*29:02, B\*15:01 με την υψηλότερη αρνητική τιμή παραμέτρου D ίση με -0,1000 και για τα γονίδια HLA-A και HLA-DRB1 το ζεύγος A\*01:01, DRB1\*03:01 με την υψηλότερη θετική τιμή παραμέτρου D ίση με 0,0159 και το ζεύγος A\*69:01, DRB1\*10:01 με την υψηλότερη αρνητική τιμή παραμέτρου D ίση με -0,0973. Για το αρχικό διαθέσιμο μητρώο των 5.131 Ελλήνων εθελοντών δοτών, η πιθανότητα ιστοσυμβατότητας μεταξύ ατόμων εντός του συγκεκριμένου μητρώου εκτιμήθηκε στην τιμή 5,17% με παράλληλες απαιτήσεις προσθήκης 2.796 νέων δοτών έτσι ώστε η πιθανότητα ιστοσυμβατότητας να αυξηθεί κατά 1%. Στο γραμμικό μοντέλο, όπου θεωρούμε ότι οι απαιτήσεις προσθήκης νέων δοτών παραμένουν σταθερές σε πλήθος, το μέγεθος του μητρώου όταν η πιθανότητα ιστοσυμβατότητας προσεγγίζει το 80% κυμαίνεται στην τιμή 214.000 για το σενάριο Α. Για το σενάριο Β το αντίστοιχο μέγεθος του μητρώου για πιθανότητα ιστοσυμβατότητας 80% κυμαίνεται στην τιμή 66.900 και για το σενάριο Γ στην τιμή 126.700. Τα υπόλοιπα δυο μοντέλα είναι πιο γενικευμένα μοντέλα που αφήνουν περιθώρια αύξησης της τιμής της πιθανότητας ιστοσυμβατότητας για μεγαλύτερα μεγέθη μητρώου. Αξίζει να σημειωθεί πως σε αυτά μοντέλα, η προσέγγιση των τιμών των πιθανοτήτων ιστοσυμβατότητας πραγματοποιείται σε μικρότερα μεγέθη μητρώου δοτών και οι τιμές των πιθανοτήτων συγκλίνουν και για τα τρία σενάρια.

Τα βασικά αποτελέσματα των αναλύσεων που υλοποιήθηκαν είναι αρκετά ισχυρά για να θεωρηθούν ποιοτικά τα δεδομένα του αρχικού μητρώου και οι περιορισμοί των αρχικών συνθηκών της βιοστατιστικής ανάλυσής τους. Επιβεβαιώνονται τα ισχυρά πλεονεκτήματα της προσέλευσης και εγγραφής εθελοντών δοτών εντός του ίδιου πληθυσμού με στόχο την αύξηση της πιθανότητας ιστοσυμβατότητας μεταξύ Ελλήνων εθελοντών δοτών που ανήκουν στο ίδιο μητρώο. Επίσης, μέσω της εκτεταμένης διασυνοριακής ανταλλαγής αιμοποιητικών βλαστικών κυττάρων εθελοντών δοτών, η εγχώρια αύξηση των διαθέσιμων HLA δεδομένων θα οδηγήσει στην διεθνή συνεισφορά του ενοποιημένου ελληνικού μητρώου στην μεταμόσχευση μυελού των οστών σε παγκόσμιο επίπεδο. Η διαθεσιμότητα μεγαλύτερων δειγμάτων, πλήρως τυποποιημένων δοτών επιτρέπει σημαντικές βελτιώσεις των υπολογιστικών μοντέλων βιοστατιστικής ανάλυσης των HLA δεδομένων. Ωστόσο, όσον αφορά στις εφαρμογές για το στρατηγικό σχεδιασμό μητρώων δοτών μυελού των οστών, η οργάνωση και η ιεράρχηση των προσπαθειών προσέλευσης νέων δοτών ίσως είναι αποδοτικότερη εάν οι υπολογισμοί βασιστούν περισσότερο σε ημι-ποσοτικά δεδομένα που προέρχονται από την ανάλυση βασικών παραμέτρων όπως, όπως για παράδειγμα, το τρέχον πλήθος εθελοντών δοτών ή η ποικιλομορφία εντός του πληθυσμού. Ως εκ τούτου, τα αποτελέσματα που παρουσιάζονται στην παρούσα εργασία θα πρέπει να χρησιμοποιηθούν ως ένα βήμα προς την ενεργοποίηση του μιας ορθολογικής στρατηγικής προσέλευσης νέων δοτών σε εθνικό και διεθνές επίπεδο παρά ως τελικό αποτέλεσμα. Παρόλο που οι συχνότητες των απλοτύπων μπορούν επίσης να εκτιμηθούν από ελλιπή HLA δεδομένα θα ήταν πιο χρήσιμο για την συνεχιζόμενη βελτίωση των δεδομένων των μητρώων στο μέλλον να καταχωρούνται μόνο τα πλήρη δεδομένα της HLA τυποποίησης σε υψηλή ανάλυση, τουλάχιστον για τις γονιδιακές θέσεις HLA-A, HLA-B και HLA-DRB1. Τα εμπόδια που σχετίζονται με το κόστος αυτής της προσέγγισης θα πρέπει όλο και περισσότερο να μειωθούν λόγω της χρήσης των νέας γενιάς τεχνολογιών αλληλούχησης κατά την διενέργεια της υψηλής ανάλυσης HLA τυποποίησης.

## 6.2 Μελλοντική Εργασία

Η παρούσα διπλωματική εργασία μπορεί να αποτελέσει τη βάση για περαιτέρω ανάπτυξη πάνω σε θέματα μελέτης ζητημάτων βιοστατιστικής ανάλυσης HLA δεδομένων δειγμάτων πληθυσμών, προσδιορισμού πιθανοτήτων ιστοσυμβατότητας μεταξύ δότη και λήπτη για την διενέργεια μεταμοσχεύσεων αιμοποιητικών βλαστικών κυττάρων και στρατηγικού σχεδιασμού και βελτιστοποίησης μητρώων εθελοντών δοτών μυελού των οστών, αλλά και σε ζητήματα φυλογενετικής ανάλυσης πληθυσμών και ανθρωπολογικών μελετών.

Επεκτείνοντας καθιερωμένες μεθόδους εκτίμησης τιμών συχνοτήτων HLA απλοτύπων και πιθανοτήτων ιστοσυμβατότητας, μπορεί να κατασκευαστεί ένα εξελιγμένο μοντέλο εκτίμησης της ανάγκης και των επιπτώσεων υλοποίησης προσπαθειών προσέλκυσης και εγγραφής νέων εθελοντών δοτών σε διάφορους πληθυσμούς, στα πλαίσια ενός διεθνούς περιβάλλοντος ανταλλαγής HLA δεδομένων εθελοντών δοτών μυελού των οστών. Παρά τους περιορισμούς λόγω της φύσης των δεδομένων εισόδου, η διαθεσιμότητα μεγαλύτερων δειγμάτων, πλήρως HLA τυποποιημένων σε υψηλή ανάλυση εθελοντών δοτών, θα οδηγήσει σε σημαντικές βελτιώσεις των υπολογιστικών μοντέλων βιοστατιστικής ανάλυσης των HLA δεδομένων και των αντίστοιχων εκτιμήσεων συχνοτήτων αλληλόμορφων, αντιγόνων και απλοτύπων, πιθανοτήτων ιστοσυμβατότητας, αλλά και σε ολοκληρωμένο και αποτελεσματικότερο σχεδιασμό βελτιστοποιημένων μητρώων HLA δεδομένων εθελοντών δοτών μυελού των οστών.

---

## ΑΝΑΦΟΡΕΣ & ΔΙΑΔΙΚΤΥΑΚΕΣ ΠΗΓΕΣ

---

- [1] Alexander H. Schmidt, Daniel Baier, Ute V. Solloch, Andrea Stahr, Nezih Cereb, Ralf Wassmuth, Gerhrd Ehninger and Claudia Rutt, «Estimation of high-resolution HLA-A, -B, -C, -DRB1 allele and haplotype frequencies based on 8.862 German stem cell donors and Implications for strategic donor registry planning», *Human Immunology* 70, (2009) 895-902.
- [2] J.S. Dhaliwal, M. Shahnaz, C.L. Too, A. Azrena, L. Maiselamah, Y.Y. Lee, Y.A. Irda and M. Salawati, «HLA-A, -B and -DR Allele and Haplotype Frequencies in Malays», *Asian Pacific Journal Of Allergy and Immunology* (2007) 25:47-51.
- [3] L. Excoffier, P. Smouse, and J. Quattro, «Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data», (1992), *Genetics* 131:479-491.
- [4] Ines Alvarez, Milka Bengochea, Roberto Toledo, Elena Carretto and Pedro C. Hidalgo, «HLA Class I Antigen and HLA-A, -B and -C Haplotype Frequencies in Uruguayans», *Human Biology*, (2006), v. 78, no. 4, pp. 513-525.
- [5] Richard M. Single, Pierre-Antoine Gourraud, Alex K. Lancaster, Farren Briggs, Lisa F. Barcellos, Jill A. Hollenbach, Steven J. Mack and Glenys Thomson, «Estimating Haplotype Frequencies and Linkage Disequilibrium Parameters in the HLA and KIR Regions», *Haplotype Estimation and Linkage Disequilibrium Manual Version 0.1.8*, (2011).
- [6] L. Matevosyan, S. Chattopadhyay, V. Madelian, S. Avagyan, M. Nazaretyan, A. Hyussian, E. Vardapetyan, R. Arutunyan and F. Jordan, «HLA-A, HLA-B and HLA-DRB1 allele distribution in a large Armenian population sample», *Tissue Antigens ISSN 0001-2815*, (2010).
- [7] Laurent Excoffier and Heidi Lischer, «Arlequin Ver. 3.5 : An Integrated Software Package for Population Genetics Data Analysis», *Computational and Molecular Population Genetics Lab (CMPG), Institute of Ecology and Evolution, University of Bern, Switzerland* (2011).
- [8] G. Mertens , Z. Berneman, B. Vandekerckhove and MP. Emonds, «Valid Belgian population frequency data on HLA A-B-DR haplotypes by means of family studies» , *Arc Public Health*, (2003), 61:191-200.
- [9] Alicia Sanchez-Mazas and Jose Manuel Nunes, «HLA data analysis in anthropology : basic theory and practice», *Laboratory of Anthropology, Genetics and Peopling History (AGP), Department of Anthropology and Ecology, University of Geneva*, (2002).
- [10] Alicia Sanchez-Mazas, «HLA data analysis in anthropology : some practical applications» , *Laboratory of Anthropology, Genetics and Peopling History (AGP), Department of Anthropology and Ecology, University of Geneva*, (2002).

- [11] Knud Christensen , «Population Genetics» , *Animal Genetics*, (2003).
- [12] K. Rohde and R. Fuerst , «Haplotyping and Estimation of Haplotype Frequencies for Closely Linked Biallelic Multilocus Genetic Phenotypes Including Nuclear Family Information», *Human Mutation*, (2001), 17:289-295.
- [13] Fang Yuan and Yongzhi Xi, «Statistic and Analytical Strategies for HLA Data» , *INTECH, Chapter 4 : HLA and Associated Important Diseases*, (2014).
- [14] S.G.E Marsh, E.D. Albert, W.F. Bodmer, R.E. Bontrop, B. Dupont, H.A. Erlich, M. Fernandez-Vina, D.E. Geraghty, R. Holdsworth, C.K. Hurley, M. Lau, K.W. Lee, B. Mach, M. Maier, W.R. Mayr, C.R. Muller, P. Parham, E.W. Petersdorf, T. Sasazuki, J.L. Strominger, A. Svejgaard, P.I. Terasaki, J.M. Tiercy and J. Trowsdale, «Nomenclature for factors of the HLA system, 2010» , *Tissue Antigens ISSN 0001-2815*, (2010).
- [15] Dario Gasbarra, Sangita Kulathinal, Matti Pirinen and Mikko J. Sillanpaa , «Estimating haplotype frequencies by combining data from large DNA pools with database information» , *IEEE Transactions on Computational Biology and Bioinformatics*, (2010).
- [16] Bruno A. Lima and Helena Alves , «HLA-A, -C, -B and -DRB1 Allelic and Haplotypic Diversity in Bone Marrow Volunteer Donors from Northern Portugal» , *Organs Tissues & Cells* (16:19-26),(2013).
- [17] G. Sulcebe, A. Sanchez-Mazas, J.M. Tiercy, E. Shyti, I. Mone, Z. Ylli and V. Kardhashi , «HLA allele and haplotype frequencies in the Albanian population and their relationship with the other European populations» , *International Journal of Immunogenetics – Blackwell Publishing Ltd.*, (2009).
- [18] Gokay Bozkurt , «HLA-A, -B and –DR allele and Haplotype Frequencies in the European American, African American and Hispanic Populations in Dallas, Texas : Relatedness to the North American Population», *Balkan Med J* (28:366-374), (2011).
- [19] Frederique Feve & Jean-Pierre Florens , «A Mathematical Model for Bone Marrow Donors' Registries and Cord Blood Banks» , *Toulouse School of Economics Working Paper Series* 10-177, (2010).
- [20] Laurent Excoffier and Montgomery Slatkin, «Maximum-Likelihood Estimation of Molecular Haplotype Frequencies in a diploid population», *Mol. Biol. Evol.* 12(5:921-927), (1995).
- [21] Matthew B. Hamilton , «Population Genetics» , *Wiley-Blackwell*, (2009).
- [22] Hans-Peter Eberhard and Carlheinz R. Muller , «The impact of HLA-C matching on donor identification rates in a European-Caucasian population», *Frontiers in Immunology*, (2014).
- [23] U. Shankarkumar, «The Human Leukocyte Antigen (HLA) System», *Kamla-Raj – Int J Hum Genet*, 4(2:91-103), (2004).

- [24] David Steiner , «Probabilistic Matching in search for unrelated hematopoietic stem cell donors» , *Doctoral Thesis, Czech Technical University in Prague – Faculty of Electrical Engineering, (2013).*
- [26] Alexander H. Schmidt, Jurgen Sauter, Julia Pingel and Gerhard Ehninger, «Toward an Optimal Global Stem Cell Donor Recruitment Strategy», *PloS One 9(1), (2014).*
- [27] Κουνιάκη Αδ. , ΔΙΑΛΕΞΗ : «Αντιγόνα Ιστοσυμβατότητας (HLA), Κυτταροτοξικά Αντισώματα έναντι των HLA, Δοκιμασίες Διασταύρωσης (Crossmatch-XM) μεταξύ Δότη-Λήπτη», 18<sup>ο</sup> Σεμινάριο Συνεχιζόμενης Εκπαίδευσης : Η συμβολή του Εργαστηρίου στις Μεταμοσχεύσεις, Γ.Ν.Α. «Ο ΕΥΑΓΓΕΛΙΣΜΟΣ», Τμ. Ανοσολογίας-Ιστοσυμβατότητας, (2014).
- [28] ΣΗΜΕΙΩΣΕΙΣ ΜΑΘΗΜΑΤΟΣ : «Ανοσολογία» , Τμήμα Νοσηλευτικής ΕΚΠΑ.
- [29] Wikipedia: The Free Encyclopedia. <http://en.wikipedia.org/>
- [30] NMDP. <https://bioinformatics.bethematchclinical.org/>
- [31] DNA Type Lookup Tool. <http://www.marrow-donor.org/cgi-bin/DNA/dnatyp.pl>
- [32] HLA Nomenclature. <http://hla.alleles.org/>
- [33] Be The Match. <http://bethematch.org/>
- [34] WMDA. <https://www.wmda.info/>
- [35] EBI. <http://www.ebi.ac.uk/>
- [36] Anthony Nolan. <http://www.anthonynolan.org/>
- [37] HLA Explorer. <http://www.hlaexplorer.net/>
- [38] HLA Net. [www.hla-net.eu](http://www.hla-net.eu)
- [39] GENDX. <http://www.gendx.com/>
- [40] NCBI. <http://www.ncbi.nlm.nih.gov/>
- [41] CIBMTR. <http://www.cibmtr.org/>
- [42] IMGT. <http://www.imgt.org/>
- [43] IHWG. <http://www.ihwg.org/>
- [44] International HLA & Immunogenetics Workshop. <http://www.ihw.org/>
- [45] Allele Frequencies in Worldwide Populations. <http://www.allelefrequencies.net>
- [46] ARLEQUIN : An Integrated Software for Population Genetics Data Analysis.  
<http://cmpg.unibe.ch/software/arlequin35/>
- [47] Center For Statistical Genetics (CSG), University of Michigan  
<http://csg.sph.umich.edu/> , <http://genome.sph.umich.edu>
- [48] Section on Statistical Genetics (SSG), University of Alabama  
<http://www.soph.uab.edu/ssg/>
- [49] ΣΥΛΛΟΓΟΣ ΦΛΟΓΑ. <http://www.floga.org.gr/>
- [50] ΣΥΛΛΟΓΟΣ ΕΛΠΙΔΑ. <http://www.elpida.org/>
- [51] ΣΥΛΛΟΓΟΣ ΟΡΑΜΑ ΕΛΠΙΔΑΣ. <http://www.oramaelpidas.gr/>
- [52] ΕΘΝΙΚΟΣ ΟΡΓΑΝΙΣΜΟΣ ΜΕΤΑΜΟΣΧΕΥΣΕΩΝ. <http://www.eom.gr/>
- [53] ΚΕΔΜΟΠ. <http://www.xarisezoi.gr/>

# **Estimation of high-resolution HLA-A, -B, -DRB1 allele and haplotype frequencies based on 5131 Greek stem cell donors and implications for strategic donor registry planning**

P. Kormpakis<sup>1</sup>, A. Sarma<sup>1</sup>, G. Oikonomopoulou<sup>1</sup>, E. Kefala<sup>1</sup>, D. Malliori<sup>1</sup>, A. Mouzaki<sup>1</sup>, G. Androulakis<sup>2</sup>, A. Spyridonidis<sup>1</sup>.

<sup>1</sup> Center to advance public awareness & recruitment of Bone Marrow Donor Volunteers of the University of Patras (CBMDP), School of Medicine, University of Patras, Rio, Patras, Greece  
email: [cbmdpatras@upatras.gr](mailto:cbmdpatras@upatras.gr),

<sup>2</sup>Department of Business Administration, University of Patras, Rio, Patras, Greece  
email: [gandroul@upatras.gr](mailto:gandroul@upatras.gr)

---

## ABSTRACT

*The frequency distributions of HLA alleles and the corresponding haplotypes in certain populations can be used to optimize the search of individual hematopoietic stem cells of donors and efficient management of donor registries. As part of this study, we determined the frequency distribution of HLA alleles, antigens and haplotypes in a population of 5131 Greek bone marrow donor volunteers for HLA-A, HLA-B and HLA-DRB1 genes, typed at high resolution, which is then used to estimate the probability of finding a compatible donor within the registry of Greek donors. The calculated haplotype frequencies used to estimate the probability (p) of finding a compatible donor within the set, according to the size (n) of the available donor registry. It was found that the number of different haplotypes that obtained is 2511, the most common has frequency 2.63% and the sum of the frequencies of the 100 most frequent being 36.937%. The histocompatibility possibility (p) is 5.17% by requiring the addition of 2796 new donors in the registry, so that this probability would be increased by 1%.*

**KEY WORDS :** Hematopoietic Stem Cell Transplantation, Bone Marrow, Histocompatibility, MHC, HLA, NMDP, Linkage Disequilibrium, Genetic Equilibrium, Hardy-Weinberg, Arlequin, EM Algorithm, Allele, Gene, Antigen, Haplotype, Genotype, Phenotype, Donor Volunteer Registry, Frequency, Biostatistics.

## 1. Introduction

The HLA system of human leukocyte antigens, located on the short arm of chromosome 6, consisting of the most polymorphic loci of the human genome [1]. The HLA system is comprised of class I, II and III non-overlapping portion of the genome encoding heterodimeric cell surface glycoproteins. The genetic variation at these loci plays an important role in immune and hematopoietic stem cells. The class I and II HLA genes play an essential role in the tissue compatibility and the susceptibility of disease, in the presentation of peptides to T cells and their recognition as foreign or not. Population studies of HLA system are of particular interest not only for phylogenetic analyzes, anthropological studies and correlation studies with diseases, but also for the extension of the single networking of HLA registry data of registered unrelated donors of hematopoietic stem cells. The compatibility analysis of HLA-A, HLA-B and HLA-DRB1 genes between donor and recipient improves survival after transplantation of hematopoietic stem cells. Accordingly, the donor search aims to find a donor who is compatible with their respective patients at least for these three HLA genes, typed at high resolution. In this context, the availability of high resolution HLA haplotype frequency distributions is helpful for two reasons. First, it increases the efficiency of individual donor searches, as many registered donors are listed only with incomplete HLA information. One or more relevant HLA genes may be typed only at low / intermediate resolution or not at all. Search coordinators can use high-resolution haplotype frequency distributions of the donor population to identify donors who have, despite their incomplete HLA typing information, high probabilities to be optimal matches for specific patients. Advanced search algorithms like HapLogic or OptiMatch use frequency distributions of HLA haplotypes for the automatic classification of potential donors with incomplete HLA data by matching probabilities. Second, population specific HLA haplotype frequency distributions determine how many patients find matching donors in a registry with  $n$  donors from the analyzed population [2].

The frequency distributions of HLA alleles and haplotypes have implications for strategic donor registry planning, as the information and donor attracting centers must work with limited resources and have to decide to what extent these resources will be allocated to the donor recruitment efforts. Donor registry HLA data are generally heterogeneous with respect to typed HLA genes, typing resolution and typing method. Also, the discovery of new allelic variants compromises high-resolution typing data over time. Furthermore, HLA data of registered stem cell donors often accumulate over several years. Additional HLA typing after registration may be carried out within the scope of specific donor searches (patient-directed typing) or initiated by the donor center or registry (prospective typing). In both cases, the selection for additional typing is based on the donor's HLA phenotype as it is known until then. Recently it was shown that such data characteristics do not generally prevent the estimation of haplotype frequencies through the expectation-maximization (EM) algorithm.

The Center to advance public awareness & recruitment of Bone Marrow Donor Volunteers of the University of Patras (CBMDP) has began to create an data registry of bone marrow donor volunteers through awareness actions and attractions that develops and by October 2009 donors are registered by filling in special bone marrow donor registration form, and then high resolution typing is implemented for HLA-A, HLA-B and HLA-DRB1 genes in the sample obtained upon registration (swab or blood). This approach may be advantageous for the operations of the center and in this way the recorded data can be useful as a bias-free data set for estimation of the HLA haplotype frequencies without relevant problems.

We estimated the HLA gene, antigen and haplotype frequencies of the Greek stem cell donor pool based on 5131 bone marrow volunteer donors of CBMDP who were typed at high resolution for HLA-A, HLA-B and HLA-DRB1 at recruitment [2-9]. Our results are used to estimate the probability of finding fully matched donors by donor registry size. In particular, we analyze how matching probabilities depend on various matching requirements.

## 2. Subjects and Methods

### 2.1. Donor Registry

Analyses were carried out with the Arlequin 3.5 [10-12], a biostatistical and population genetics software package. The Arlequin software package is a free computer program made by the Computational and Molecular Population Genetics Laboratory of the Institute of Ecology and Evolution of the University of Bern. Data was collected from donors who were recruited at the registry of Greek bone marrow volunteer donors of CBMDP during the period between October 2009 and April 2014, through donation departments of various hospitals in Greece or through awareness of bone marrow donor volunteers events that organized by the CBMDP ([www.xarisezoi.gr](http://www.xarisezoi.gr)).

The CBMDP donor registry that was analyzed, consists of 5131 Greek bone marrow donors and includes information on identity (gender, age, phone numbers, addresses, place of origin, relatives), recording data (date, place of registration, information source), biological factors (sample type, height, weight, number of pregnancies) and medical history data. The data is in anonymous form in accordance with the confidentially agreement signed in the form of a contract between the donor volunteer and the CBMDP, governed by Greek law. Each donor has signed a printed consent for receiving a sample (swab or blood) for HLA typing. For each donor, the HLA-A, HLA-B and HLA-DRB1 genes were typed at high resolution, according to the course of CBMDP.

### 2.2. HLA typing and processing of HLA data

High-resolution HLA typing was carried out at the ASHI-accredited laboratory of Histogenetics, Inc. (Ossining, NY). Exons 2 and 3 were sequenced for Class I genes. For HLA-DRB1, exon 2 was sequenced. Cis-trans ambiguities within the relevant exons were resolved. Synonymous mutations were not taken into account. Alleles with synonymous mutations inside or outside the relevant exons are merged. The form of nomenclature of alleles [13], as resulting from the high-resolution HLA typing is as follows :  
e.g. A\*01:DCEJ, A\*24:DRZV, B\*08:BYXR, B\*27:BHXW, DRB1\*03:01:01, DRB1\*04:04.

The nomenclature codes of alleles are an elementary notation used by the NMDP to represent combinations of allele database registries of donors at the molecular HLA typing. Once identified a new combination of alleles, a code is assigned to the combination, which is then activated (ACTIVATED) and is available for use by the representation of a particular gene locus and allele specific family to which it belongs. The “DNA Type Lookup Tool” is used to convert the codes of each allele in the usual form of four digits (4-digit). The conversion of the codes is implemented via the access in data tables of allele codes, which are updated on a daily basis. For each code, the first code is selected from the conversion result, which is the most frequent recorded code up to that time. Less code number represents greater matching probability. So, for the donor above, the result of conversion is as follows : A\*01:01, A\*24:02, B\*08:01, B\*27:05, DRB1\*03:01, DRB1\*04:04.

#### *Missing Data*

The Arlequin manages missing data (NULL) for any allele as any other recorded allele. This, in the analysis of HLA haplotypes creates data that do not correspond to reality. However, the frequency of missing values is taken into account when calculating the allele and haplotype frequencies.

### 2.3. Test of Hardy-Weinberg Equilibrium (HWE)

The Arlequin 3.5 software package was used to estimate the frequency of HLA haplotypes via the EM algorithm and to check for Hardy-Weinberg equilibrium [14-15]. The goal is to find the best estimation of the frequencies of HLA haplotypes within the population, given certain information about the phenotypes that contained in the data set. The EM algorithm is based on the pursuit of maximizing the expectation and is a recursive frequency estimation method of the haplotypes of a population, taking into account the frequencies of the alleles and processing missing data of the sample. The Hardy-Weinberg equilibrium is a fundamental concept of population genetics. It is a mathematical equation that describes the distribution and expression of the alleles in a population and defines the circumstances in which the frequency of an allele is expected to change. Deviation from any of the conditions under which the principle of Hardy-Weinberg equilibrium can change the frequency of alleles of a population from one generation to the other. The test for HWE was carried out because both the EM algorithm and the calculation of matching probabilities by donor registry size require the analyzed population to be in HWE. The p-values of less than 5% were considered as statistically significant. However, the principle of HWE is a simplified approach, because it applies fully only under certain special situations that are rarely encountered in human populations [7]. The HWE law applies clearly only for large populations, where there is random mating and no selection, mutation and migration. Any deviation from Hardy-Weinberg equilibrium may be a major source of errors, because the EM algorithm is based on equilibrium values in the step of expectation. However, any deviation from HWE values will not lead to significant variation in the estimation of HLA haplotype frequencies [10].

### 2.4. Linkage Disequilibrium (LD)

To determine the value of linkage disequilibrium between pairs of HLA-A, HLA-B and HLA-DRB1 alleles, we calculated the value of parameter  $D$  via Arlequin 3.5 [3, 15]. In population genetics, linkage disequilibrium is the non-random association (correlation) of alleles at two or more loci from individual ancestral chromosomes. The linkage disequilibrium is an indispensable tool for implementing genetic association studies. In other words, linkage disequilibrium is the occurrence of certain combinations of alleles or genetic markers in a population more frequently or less frequently than would be expected from a random configuration of haplotypes from alleles, based on their frequencies. We calculated the greater linkage disequilibrium values for the most commonly observed allele pairs of the three loci. Also, the value of linkage disequilibrium does not greatly affect the frequencies of the common HLA haplotypes [10]. To determine the value of linkage disequilibrium between pairs of alleles via Arlequin 3.5, we presume that the gametic phase is unknown and ignore the existence of defective alleles in the sample of the population.

### 2.5. Matching Probability

Each individual of the population of the Greek donors in the CBMDP registry has a set of features, so a graft can be transplanted between individuals having the same characteristics. Specifically, the type of characteristics we presume is the HLA phenotype of each individual, which is determined by the alleles of each person for each of the three loci HLA-A, HLA-B and HLA-DRB1, typed at high resolution [16].

#### *Analyzed HLA Matching Scenarios*

By the calculations performed by the Arlequin 3.5 software, we examined three different scenarios in the levels of HLA analysis of the HLA data of the registered donors of CBMDP registry to determine the degree of HLA histocompatibility.

**SCENARIO A :** *High-resolution matching for HLA-A, HLA-B & HLA-DRB1.*

The basic scenario refers to data management, when is represented a high resolution HLA typing for HLA-A, HLA-B and HLA-DRB1 genes. The high-resolution HLA typing of alleles corresponds to the 4-digit representation (allelic level).

**SCENARIO B :** *Low-resolution matching for HLA-A, HLA-B & HLA-DRB1.*

The low-resolution HLA typing of alleles corresponds to the 2-digit representation (antigen level).

**SCENARIO C :** *Low-resolution matching for HLA-A, HLA-B & High-resolution matching for HLA-DRB1.*

### ***Matching Probabilities Estimation Models***

#### ***Linear Model for the initial donor registry***

Based on the above, is defined the following mathematical calculation [2, 16-19] which is the probability of finding at least one fully matched donor for the next patient of the registry in need of a transplant

$$p(n) = \sum_{i=1}^P f_i \left[ 1 - (1 - f_i)^n \right]$$

where :

$P$  is the initial registry size of the 5131 donors.

$n$  is the test donor registry size.

$f_i$  is the relative frequency of the  $i$ -th phenotype.

Apart from the matching probability  $p(n)$ , we also calculate and its derivative

$$p'(n) = - \sum_{i=1}^P f_i \cdot \ln(1 - f_i) (1 - f_i)^n$$

Then, the following quantity is established, which indicates the number of new donors who must be recruited at donor registry size  $n$  in order to increase the matching probability for the next patient by 1%.

$$d(n) = \frac{1}{100 \cdot p'(n)}$$

Given the above, we will calculate the quantities  $p(n)$  and  $d(n)$  for  $n=5131$  and we will consider that this model has linear behavior, namely that the requirements of adding new donors to increase the value of the matching probability by a fixed percentage are unchanged.

#### ***Discrete Model***

At this point, it will become an attempt to calculate matching probabilities by donor registry size (DRS), based on an alternative approach as presented in [16]. This alternative approach focuses on a different analysis of the phenotype frequencies of registry individuals. While the previous approach use the relative phenotype frequencies of the 5131 phenotypes of the donors, such as these derived from the estimations of the analysis of the Arlequin software package, with this approach we initially count the number of single (unique) phenotypes with accuracy 8/8 for the analyzed registry. Thus, for each analyzed matching scenario, the registry is determined by the phenotypes  $1, \dots, J$ , where each of them has a frequency  $f_i > 0$  with

$\sum_{i=1}^J f_i = 1$ . The percentage of matching probability between donor and recipient phenotypes corresponds to a probability  $\alpha$ .

Assuming that there is no initial donor stock, the matching probability defined as follows :

$$p(n) = 1 - \sum_{i=1}^J f_i \cdot e^{-anf_i}$$

$$\text{with } \sum_{i=1}^J f_i = 1 \quad \text{and} \quad f_i \geq 0 \quad \forall i = 1, \dots, J$$

A realistic value for the variable  $\alpha$  is  $\alpha = \frac{1}{3}$ , that the  $\frac{1}{3}$  of donors, compatible at the registry level are fully compatible and available for a graft (France Greffe de Moelle Registry). Based on the above, we calculate the value of the probability  $p(n)$  for increasing values of the donor registry size  $n$ .

### Continuous Model

Finally, a more generalized matching probability analysis model by the donor registry size, as presented in [16], assume that although the number of different phenotypes that presented in a donor registry and consequently of their frequencies, is finite (discrete), we can manage the phenotypes and their frequencies as a continuous distribution of positive values. More specifically, we assume that the phenotypes are randomly generated from an exponential distribution of density  $\lambda \cdot e^{-\lambda i}$  ( $i \in [0, +\infty]$ ), that the value of matching probability  $p(n)$  is calculated as follows :

$$p(n) = 1 - \frac{1 - e^{-an\lambda}}{an\lambda}$$

where the value of the variable  $\lambda$  resulting from the median of the distribution and is defined as  $\lambda = -\frac{1}{X} \ln \frac{1}{2}$ , where the number  $X$  of the most frequent phenotypes represent 50% of all phenotypes of the population.

## **3. Results**

### **3.1. Donor Population Sample Attributes**

The characteristics of the sample of donor population about gender, age and region of origin are shown in Table 1. All donors declared their Greek origin and a fraction of 1% foreigners residing in Greece. The places of origin and residence of all donors are distributed in all administrative regions of the Greek territory (continental and insular country). Because of the seat of CBMDP in Patras and actions were making at times in various places of Greece, the number of donors registered divided by 17 in Epirus to 1882 in Western Greece.

### **3.2. Number & Frequencies of Antigens, Alleles and Haplotypes**

The results of biostatistics analysis of the HLA data, were observed the values of the number of different antigens and alleles derived for HLA-A, HLA-B and HLA-DRB1 loci separately, as well as the values of frequencies of each allele and antigen. The values of the number of observed antigens and alleles for each locus are shown in Table 2.

### ***Cumulative Allele Frequencies***

Figure 1 shows the graph of the cumulative frequency of alleles. As is obvious, the greatest diversity of alleles displayed for the HLA-B gene in data registry and the smallest is displayed for the HLA-DRB1 gene. The values of the frequencies of all antigens and alleles of the three HLA loci, as resulted from the analysis of the sample, are shown in Table 3 and 4 respectively.

The results of the analysis of the HLA data include values of the number of different HLA haplotypes that result from the combination of the alleles of the three HLA loci, as well as the estimated values of the frequency of each haplotype, according to the EM algorithm. The number of unique haplotypes is equal to 2511 (with a frequency threshold value 0.0001). The estimated values of the frequencies of the 100 most frequent haplotypes that result from the combination of the HLA alleles of the three HLA loci, as resulted from the analysis of the sample are shown in Table 5. As is obvious, the estimated frequencies of the 100 most frequent haplotypes in the sample account for 36.937% of all haplotypes.

### **3.3. HWE Test & Linkage Disequilibrium (LD)**

From the results of the HWE test, it becomes clear that the differentiation between the observed and expected genotype frequencies of the analyzed population sample is statistically significant for all examined HLA loci. This means that the analyzed population is not in HWE and the possibility of finding the actual results when it is considered that the population is in HWE is less than 5%.

Table 6 shows the 10 higher positive and 10 negative linkage disequilibrium values for the observed allelic pairs of the three HLA loci.

### ***Comparison with German Registry***

At this point, it is appropriate to present some comparative data on the estimated haplotype frequencies of the CBMDP donor registry of 5131 Greek registered donors with respect to the corresponding haplotype frequencies of a donor registry of 8862 German registered donors [2]. Regarding the number of alleles of HLA loci calculated in these two registries, there were found, for the HLA-A gene 61 different alleles in the German registry and 64 in the Greek, for the HLA-B gene 101 different alleles in the German registry and 107 in the Greek, and for the HLA-DRB1 gene 68 different alleles in the German registry and 65 in the Greek. Table 7 shows the 15 most frequent common HLA haplotypes of the Greek and German registry. It is worth noting that the number of the different haplotypes of the Greek registry of 5131 donor volunteers was estimated at 2511, with a cumulative frequency of the 50 most frequent equal to 26.721%, while the number of different haplotypes of the German registry of 8862 donor volunteers was estimated at 3810, with a cumulative frequency of 50 most frequent equal to 35.4%.

Assuming that the haplotype frequencies of the 5131 CBMDP Greek donor registry are normally distributed, we calculate the mean  $\mu$  of  $3.9 \times 10^{-4}$  and standard deviation  $\sigma$  equal to 0.0010. This means that most of the total 2511 haplotype frequencies are between 0.039% and 0.139%. The values for the mean and standard deviation of the normal distribution of the haplotype frequencies of the German donor registry are  $2.83 \times 10^{-4}$  and 0.0014 respectively. This means that most of the total 3810 haplotypes have a frequency between 0.0283% and 0.1683%. The graph of the normal curve of the frequency distribution of the two registries is shown in Figure 5.

### 3.4. Matching Probabilities & Registry Optimization

Based on the size of the available CBMDP donor volunteer registry of 5131 Greek donors, the calculated results indicate that the probability  $p(n)$  of finding at least one fully compatible donor for the next patient of the registry in need of a transplant is 5.17% and the number  $d(n)$  of new donors who must be recruited at donor registry size  $n$  in order to increase the matching probability for the next patient by 1% is 2796 donors.

In an attempt to generalize the results of the analysis and to optimize the bone marrow registries design methodologies, the matching probability  $p(n)$  values and the number of new donors  $d(n)$  can be estimated for different registry sizes that are much larger than the original registry size of 5131 donors. However, the reliability of the estimations depends, in each case, on the characteristics of the HLA data of each registry, the level of the HLA gene typing and the registry size. Initially, starting from the above calculation results for the quantities  $p(n)$  and  $d(n)$ , we calculated the requirements of the original registry to add new donors to it, so that the matching probability  $p(n)$  to be increased gradually. The calculations were made for three different scenarios A, B and C as they have already been defined. Starting from the initial size of 5131 donors, through the calculations we estimated the corresponding values for the three scenarios as presented in Table 8.

#### *Linear Model for the available initial registry*

Figure 2 shows the graph of matching probability by the donor registry size for the three different scenarios. From the calculated results, it is obvious that the requirements for new donors addition in the registry, so as to increase its size and thus the matching probability  $p(n)$ , are much greater for the more strict in terms of HLA gene typing scenario A. Conversely, the value of the probability  $p(n)$  approaches the value of 100% in smaller registry sizes, when the requirements for addition of new donors are lower. The relationship between the matching probability and the registry size, it seems, almost linear. In order to approximate the value of 100% for the matching probability, the registry size requirements ranging in a value equal to 270191 for scenario A, 86096 for scenario B and 160331 for scenario C.

#### *Discrete Model*

The number of different phenotypes with 8/8 accuracy for the three different scenarios is equal to 4925 for scenario A. Compared with the linear model, it is observed that for the approximation of a matching probability  $p(n)$  value near to 100%, the donor registry size is around a value of 148000, which is near to the average value of the respective matching probability  $p(n)$  values of the linear model and the matching probability values for the initial donor registry size of 5131 donors are increased compared with the linear model and range at around 30%. Figure 3 shows the graph of matching probability  $p(n)$  by the donor registry size in logarithmic x axis, according to the discrete non-linear model.

#### *Continuous Model*

The value of the variable  $X$  was calculated at 2265 phenotypes for scenario A. Compared with the discrete model, it is observed that for the approximation of a matching probability  $p(n)$  value near to 100%, when the donor registry size is around the value of 148000 donors, the matching probability  $p(n)$  value is at around 93%, leaving small room for growth for even larger donor registry sizes. For donor registry size values around the value of 5131 donors of the initial donor registry, the corresponding matching probability value is slightly reduced at the level of 22%. Figure 4 shows the graph of matching probability  $p(n)$  by the donor registry size in logarithmic x axis, according to the continuous non-linear model.

#### 4. Discussion

In this work, we applied standard methods for calculation of the frequency distribution of HLA antigens and alleles and for estimation of HLA haplotype frequencies and matching probabilities between registered donors of the CBMDP bone marrow registry of Greek donor volunteers, including calculations carried out to identify specific biostatistical analysis metrics of HLA data. As a general result, we can deduce the significant benefits of intrapopulation hematopoietic stem cell donor volunteers attracting. The model used in the analyzes was based on the hypothesis that the HLA data of donors who are fully compatible in patients in need of a transplant can be determined when a compatible donor search process is in progress. However, there is evidence that there are fully compatible donors may remain unknown when a donor search process is in progress, due to missing HLA data [20-22].

At regard of qualitative and quantitative characteristics of the donor registry sample that was used during the analyzing processes, as the determining of the histocompatibility degree in a bone marrow transplantation process is based on high-resolution HLA typing of HLA-A, HLA-B and HLA-DRB1 genes, the sample of 5131 donors of CBMDP bone marrow donor volunteer registry has sufficient consistency to find single compatible donors and for implications for strategic donor registry planning.

About the diversity of donor registry data, the HLA-B gene is the most polymorphic in the number of different antigens and alleles, with most frequent the HLA-B\*35 antigen at a frequency of 17.355% and most frequent allele the HLA-B\*51:01 at a frequency of 12.442%. Accordingly, about the HLA-A gene, the most frequent antigen is the HLA-A\*02 at frequency 26.092% and the allele HLA-A\*02:01 at a frequency of 23.4% and about the HLA-DRB1 gene, the most frequent antigen is the HLA-DRB1\*11 with a frequency of 25.764% and the allele HLA-DRB1\*11:04 at a frequency of 15.676%. The number of the different obtained haplotypes is 2511 with a cumulative frequency of the 100 most frequent to be equal to 36.937%. The most frequent haplotype is the A\*02:01, B\*18:01, DRB1\*11:04 with a frequency of 2.634%.

From the results it is evident that there is a large number of haplotypes, which are of very low frequency or they are observed by unique individuals in the sample. The human geography of the sample may interpret the results of the analyzes, as the most donors who are registered in the CBMDP donor registry come from Western Greece. By increasing the registry size, while expanding donor recruitment activities in all regions of mainland but also of islands, the quality of the registry can become even stronger. Also, the cumulative frequency of the most frequent haplotypes can be enhanced. The large number of unique haplotypes indicates that over time there was a strong movement of the Greek population and its evolution was complicated.

The quality of a donor registry as to the number of the different haplotypes that include depends both on the donor registry size and by the values of the parameters of mean  $\mu$  and standard deviation  $\sigma$  of the normal distribution that the haplotype frequencies follow. The differences are apparent in the form of the density function curve in each case. Changes in the mean value are shown as shifts of the curve at a new location and changes in the standard deviation value cause changes in the shape of the curve (narrowness and height). The lower the value of standard deviation is, the higher and closer the normal curve is, i.e. the smaller the interval at which the distribution extends. The curve shows in position  $x=\mu$  a maximum

value  $y = \frac{1}{\sigma\sqrt{2\pi}}$  and in positions  $x=\mu-\sigma$  and  $x=\mu+\sigma$  has turning points. Therefore, opening

up the curve, the low height of the curve and the small registry size lead to a greater number of haplotypes more polymorphic, with a more evenly distributed set of frequencies.

From the HWE test that was conducted on the donor registry data, it is obvious that this sample of Greek population has a high degree of heterozygosity 87.7% for the HLA-A gene, 94.1% of the HLA-B gene and 92.4 for the HLA-DRB1 gene and that this sample is not in HWE for any HLA locus. This means that the population of the sample is not characterized by random mating and the three analyzed HLA loci show changing mutation rate. Also, as mentioned above, it is likely changes to have been occurred in the structure of the population due to immigration, which resulted in changes in gene frequency from one generation to the

next. It has been shown that deviations from HWE do not adversely affect the accuracy of the estimations of haplotype frequencies with the EM algorithm [23].

At regard of the values of linkage disequilibrium (LD), the higher values for parameter  $D$  were obtained for the allele pairs of the HLA-B and HLA-DRB1 genes, something that confirms the close physical distance of these HLA genes on the chromosome 6, in relation to the distance of the other pairs of the HLA genes. Specifically, the pair B\*18:01, DRB1\*11:04 has the highest positive value of parameter  $D$  of 0.0410 and the pair B\*15:03, DRB1\*15:01 has the highest negative value of parameter  $D$  of -0.0996. This means that the first pair is shown as a combination more frequently than all the other pairs in the population than would be expected from a random configuration of the haplotypes from the alleles, based on their frequencies, whereas the second pair is shown less frequently than all the other pairs. These alleles tend to be inherited together as a pair and can remain genetically linked, without being separated by the recombination. Corresponding allele pairs for HLA-A and HLA-B genes represent by the pair A\*02:01, B\*51:01 with the highest positive value of parameter  $D$  equal to 0.0228 and the pair A\*29:02, B\*15:01 with the highest negative value of parameter  $D$  equal to -0.1000 and for the HLA-A and HLA-DRB1 genes represent by the pair A\*01:01, DRB1\*03:01 with the highest positive value of parameter  $D$  equal to 0.0159 and the pair A\*69:01, DRB1\*10:01 with the highest negative value of parameter  $D$  equal to -0.0973.

For the available initial registry of 5131 Greek donors, the matching probability between individuals within that registry estimated as a value equal to 5.17% with adding requirements of 2796 new donors so that the matching probability to be increased by 1%. In the linear model, where we believe that the addition requirements of new donors remain constant in number, the donor registry size range around the value of 214000 for scenario A when the matching probability approaches the value of 80%, 66900 for scenario B and 126700 for scenario C. The other models leave margins in order to increase the value of the matching probability for larger registry sizes, achieving the maximum value when the registry size range around the value of 148000 donors. It is noteworthy that in these models, the approximation of the values of matching probability takes place in smaller donor registry sizes.

The main results of the analysis are strong enough to be considered as qualitative the data of the initial donor registry and the limitations of the initial conditions of the biostatistical analysis. There are confirmed the strong advantages of attracting and recording bone marrow donor volunteers within the same population in order to increase the matching probability between Greek donors who belong to the same registry. Also, through extensive cross-border exchange of hematopoietic stem cell donor volunteers, the domestic increase of the available HLA data will lead to the international contribution of the consolidated Greek donor registry in bone marrow transplantation at a global level [20, 21]. The availability of larger samples, of fully HLA typed donors allows significant improvements in computational models for HLA data biostatistical analysis.

However, with regard to applications for strategic bone marrow donor registry planning, the organization and prioritization of efforts to attract new donors might be more efficient if the calculations are based more on semi-quantitative data derived from the analysis of basic parameters such as, for example, the number of donor volunteers or the diversity within the population. Therefore, the results presented in this work should be used as a step in the activation of a rational strategy of attracting new donors at national and international level, than as a final result.

Although the haplotype frequencies can also be estimated from incomplete HLA data, it would be more useful for the ongoing improvement of donor registry data, in the future to be registered only the complete data of the HLA typing processes in high-resolution, at least for the HLA-A, HLA-B and HLA-DRB1 loci. The problems relate to the cost of this approach should increasingly be reduced due to the use of next generation sequencing technologies during the high resolution HLA typing methods.

## References

- [1] U. Shankarkumar, «The Human Leukocyte Antigen (HLA) System», *Kamla-Raj – Int J Hum Genet*, 4(2:91-103), (2004).
- [2] Alexander H. Schmidt, Daniel Baier, Ute V. Solloch, Andrea Stahr, Nezh Cereb, Ralf Wassmuth, Gerhrd Ehninger and Claudia Rutt, «Estimation of high-resolution HLA-A, -B, -C, -DRB1 allele and haplotype frequencies based on 8.862 German stem cell donors and Implications for strategic donor registry planning», *Human Immunology* 70, (2009) 895-902.
- [3] Richard M. Single, Pierre-Antoine Gourraud, Alex K. Lancaster, Farren Briggs, Lisa F. Barcellos, Jill A. Hollenbach, Steven J. Mack and Glenys Thomson, «Estimating Haplotype Frequencies and Linkage Disequilibrium Parameters in the HLA and KIR Regions», *Haplotype Estimation and Linkage Disequilibrium Manual Version 0.1.8*, (2011).
- [4] J.S. Dhaliwal, M. Shahnaz, C.L. Too, A. Azrena, L. Maiselamah, Y.Y. Lee, Y.A. Irida and M. Salawati, «HLA-A, -B and -DR Allele and Haplotype Frequencies in Malays», *Asian Pacific Journal Of Allergy and Immunology* (2007) 25:47-51.
- [5] Ines Alvarez, Milka Bengochea, Roberto Toledo, Elena Carretto and Pedro C. Hidalgo, «HLA Class I Antigen and HLA-A, -B and -C Haplotype Frequencies in Uruguayans», *Human Biology*, (2006), v. 78, no. 4, pp. 513-525.
- [6] L. Matevosyan, S. Chattopadhyay, V. Madelian, S. Avagyan, M. Nazaretyan, A. Hyussian, E. Vardapetyan, R. Arutunyan and F. Jordan, «HLA-A, HLA-B and HLA-DRB1 allele distribution in a large Armenian population sample», *Tissue Antigens ISSN 0001-2815*, (2010).
- [7] G. Mertens, Z. Berneman, B. Vandekerckhove and MP. Emonds, «Valid Belgian population frequency data on HLA A-B-DR haplotypes by means of family studies», *Arc Public Health*, (2003), 61:191-200.
- [8] G. Sulcebe, A. Sanchez-Mazas, J.M. Tiercy, E. Shyti, I. Mone, Z. Ylli and V. Kardhashi, «HLA allele and haplotype frequencies in the Albanian population and their relationship with the other European populations», *International Journal of Immunogenetics – Blackwell Publishing Ltd.*, (2009).
- [9] Gokay Bozkurt, «HLA-A, -B and -DR allele and Haplotype Frequencies in the European American, African American and Hispanic Populations in Dallas, Texas : Relatedness to the North American Population», *Balkan Med J* (28:366-374), (2011).
- [10] Laurent Excoffier and Heidi Lischer, «Arlequin Ver. 3.5 : An Integrated Software Package for Population Genetics Data Analysis», *Computational and Molecular Population Genetics Lab (CMPG), Institute of Ecology and Evolution, University of Bern, Switzerland* (2011).
- [11] Alicia Sanchez-Mazas and Jose Manuel Nunes, «HLA data analysis in anthropology : basic theory and practice», *Laboratory of Anthropology, Genetics and Peopling History (AGP), Department of Anthropology and Ecology, University of Geneva*, (2002).
- [12] Alicia Sanchez-Mazas, «HLA data analysis in anthropology : some practical applications», *Laboratory of Anthropology, Genetics and Peopling History (AGP), Department of Anthropology and Ecology, University of Geneva*, (2002).
- [13] S.G.E Marsh, E.D. Albert, W.F. Bodmer, R.E. Bontrop, B. Dupont, H.A. Erlich, M. Fernandez-Vina, D.E. Geraghty, R. Holdsworth, C.K. Hurley, M. Lau, K.W. Lee, B. Mach, M. Maiers, W.R. Mayr, C.R. Muller, P. Parham, E.W. Petersdorf, T. Sasazuki, J.L. Strominger, A. Svejgaard, P.I. Terasaki, J.M. Tiercy and J. Trowsdale, «Nomenclature for factors of the HLA system, 2010», *Tissue Antigens ISSN 0001-2815*, (2010).
- [14] K. Rohde and R. Fuerst, «Haplotyping and Estimation of Haplotype Frequencies for Closely Linked Biallelic Multilocus Genetic Phenotypes Including Nuclear Family Information», *Human Mutation*, (2001), 17:289-295.
- [15] Fang Yuan and Yongzhi Xi, «Statistic and Analytical Strategies for HLA Data», *INTECH, Chapter 4 : HLA and Associated Important Diseases*, (2014).
- [16] Frederique Feve & Jean-Pierre Florens, «A Mathematical Model for Bone Marrow Donors' Registries and Cord Blood Banks», *Toulouse School of Economics Working Paper Series 10-177*, (2010).
- [17] Laurent Excoffier and Montgomery Slatkin, «Maximum-Likelihood Estimation of Molecular Haplotype Frequencies in a diploid population», *Mol. Biol. Evol.* 12(5:921-927), (1995).
- [18] Hans-Peter Eberhard and Carlheinz R. Muller, «The impact of HLA-C matching on donor identification rates in a European-Caucasian population», *Frontiers in Immunology*, (2014).
- [19] David Steiner, «Probabilistic Matching in search for unrelated hematopoietic stem cell donors», *Doctoral Thesis, Czech Technical University in Prague – Faculty of Electrical Engineering*, (2013).
- [20] Alexander H. Schmidt, Jurgen Sauter, Julia Pingel and Gerhard Ehninger, «Toward an Optimal Global Stem Cell Donor Recruitment Strategy», *PloS One* 9(1), (2014).
- [21] Schipper RF, D'Amato J, Bakker JT, van Rood JJ, Oudshoorn M. «HLA gene and haplotype frequencies in Bone Marrow Donors Worldwide Registries», *Human Immunology*, (1997), 52: 54-71.

- [22] Schmidt AH, Solloch U, Sauter J, Hofmann JA, Pingel J, «Simulation shows that registered stem cell donors with full donor-patient HLA-match can remain unidentified», *Tissue Antigens*, (2013), 81: 270 (Abstract).
- [23] Fallin D, Schork NJ, «Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data», *Am J Hum Genet.* (2000), 67(4): 947–59.

**TABLE 1.** Population sample characteristics of 5131 CBMDP Greek donor volunteer registry.

| POPULATION SAMPLE CHARACTERISTICS OF<br>5131 CBMDP GREEK DONOR VOLUNTEER REGISTRY |                       |               |               |               |                |           |       |       |
|-----------------------------------------------------------------------------------|-----------------------|---------------|---------------|---------------|----------------|-----------|-------|-------|
|                                                                                   | G E N D E R           |               |               |               |                |           |       |       |
|                                                                                   | M E N                 |               |               |               | W O M E N      |           |       |       |
| PERCENTAGE (%)                                                                    | 47,63                 |               |               |               | 52,37          |           |       |       |
| NUMBER                                                                            | 2.444                 |               |               |               | 2.687          |           |       |       |
|                                                                                   |                       |               |               |               |                |           |       |       |
|                                                                                   | A G E                 |               |               |               |                |           |       |       |
|                                                                                   | < 20 YEARS            | 20 - 30 YEARS | 30 - 40 YEARS | 40 - 50 YEARS | > 50 YEARS     |           |       |       |
| PERCENTAGE (%)                                                                    | 2,96                  | 27,30         | 35,59         | 28,65         | 5,50           |           |       |       |
| NUMBER                                                                            | 154                   | 1.401         | 1.826         | 1.470         | 280            |           |       |       |
|                                                                                   |                       |               |               |               |                |           |       |       |
|                                                                                   | O R I G I N   A R E A |               |               |               |                |           |       |       |
|                                                                                   | WESTERN GREECE        | SOUTH AEGEAN  | ATTICA        | PELOPONNESE   | CENTRAL GREECE | MACEDONIA | CRETE | OTHER |
| PERCENTAGE (%)                                                                    | 36,68                 | 20,46         | 18,22         | 7,70          | 5,67           | 5,00      | 3,90  | 2,37  |
| NUMBER                                                                            | 1.882                 | 1.050         | 935           | 394           | 291            | 257       | 200   | 122   |

**TABLE 2.** Results of the observed antigens and alleles number.

| <b>LOCUS</b> | <b>NUMBER OF<br/>UNIQUE ANTIGENS</b> | <b>NUMBER OF<br/>UNIQUE ALLELES</b> |
|--------------|--------------------------------------|-------------------------------------|
| HLA-A        | 20                                   | 64                                  |
| HLA-B        | 35                                   | 107                                 |
| HLA-DRB1     | 14                                   | 65                                  |

**TABLE 3.** Results of antigen frequencies.

| HLA-A   |               | HLA-B   |               | HLA-DRB1 |               |
|---------|---------------|---------|---------------|----------|---------------|
| ANTIGEN | FREQUENCY (%) | ANTIGEN | FREQUENCY (%) | ANTIGEN  | FREQUENCY (%) |
| 02      | 26,092        | 35      | 17,355        | 11       | 25,764        |
| 24      | 14,958        | 51      | 13,504        | 16       | 15,232        |
| 01      | 10,391        | 18      | 11,635        | 13       | 9,145         |
| 03      | 9,575         | 44      | 7,525         | 04       | 9,094         |
| 11      | 7,955         | 08      | 4,573         | 03       | 7,688         |
| 32      | 7,257         | 07      | 3,916         | 15       | 7,336         |
| 26      | 5,118         | 40      | 3,769         | 01       | 6,621         |
| 68      | 3,666         | 39      | 3,413         | 07       | 6,480         |
| 30      | 3,236         | 15      | 3,340         | 14       | 5,938         |
| 23      | 2,486         | 55      | 3,036         | 10       | 1,898         |
| 33      | 2,317         | 13      | 2,967         | 08       | 1,563         |
| 29      | 1,892         | 49      | 2,850         | 12       | 1,524         |
| 31      | 1,628         | 38      | 2,786         | 09       | 0,313         |
| 25      | 0,750         | 52      | 2,737         |          |               |
| 69      | 0,507         | 27      | 2,572         |          |               |
| 66      | 0,497         | 14      | 2,326         |          |               |
| 74      | 0,117         | 41      | 1,937         |          |               |
| 34      | 0,010         | 57      | 1,559         |          |               |
| 80      | 0,010         | 58      | 1,517         |          |               |
|         |               | 37      | 1,419         |          |               |
|         |               | 50      | 1,367         |          |               |
|         |               | 47      | 0,536         |          |               |
|         |               | 53      | 0,517         |          |               |
|         |               | 56      | 0,429         |          |               |
|         |               | 73      | 0,352         |          |               |
|         |               | 45      | 0,224         |          |               |
|         |               | 34      | 0,166         |          |               |
|         |               | 48      | 0,097         |          |               |
|         |               | 42      | 0,029         |          |               |
|         |               | 54      | 0,020         |          |               |
|         |               | 78      | 0,020         |          |               |
|         |               | 31      | 0,010         |          |               |
|         |               | 59      | 0,010         |          |               |
|         |               | 81      | 0,010         |          |               |

**TABLE 4.** Results of allele frequencies.

| HLA-A  |               | HLA-B  |               |        |               | HLA-DRB1 |               |
|--------|---------------|--------|---------------|--------|---------------|----------|---------------|
| ALLELE | FREQUENCY (%) | ALLELE | FREQUENCY (%) | ALLELE | FREQUENCY (%) | ALLELE   | FREQUENCY (%) |
| 02:01  | 23,400        | 51:01  | 12,442        | 15:02  | 0,029         | 11:04    | 15,676        |
| 24:02  | 14,147        | 18:01  | 10,644        | 27:12  | 0,029         | 16:01    | 12,876        |
| 01:01  | 9,992         | 35:01  | 7,816         | 35:36  | 0,029         | 11:01    | 8,500         |
| 03:01  | 8,564         | 08:01  | 4,531         | 42:01  | 0,029         | 03:01    | 7,529         |
| 11:01  | 7,800         | 35:03  | 4,290         | 47:03  | 0,029         | 07:01    | 6,375         |
| 32:01  | 7,101         | 44:02  | 4,025         | 13:01  | 0,020         | 14:01    | 5,332         |
| 26:01  | 4,852         | 07:02  | 3,109         | 15:10  | 0,020         | 01:01    | 4,789         |
| 68:01  | 3,163         | 55:01  | 2,951         | 15:71  | 0,020         | 15:01    | 4,713         |
| 23:01  | 2,438         | 13:02  | 2,923         | 18:33  | 0,020         | 13:01    | 4,201         |
| 30:01  | 1,972         | 49:01  | 2,812         | 37:38  | 0,020         | 13:02    | 3,074         |
| 31:01  | 1,589         | 38:01  | 2,750         | 41:29  | 0,020         | 04:03    | 2,675         |
| 33:01  | 1,585         | 52:01  | 2,692         | 44:06  | 0,020         | 15:02    | 2,272         |
| 02:05  | 1,452         | 35:02  | 2,637         | 50:34  | 0,020         | 04:05    | 2,207         |
| 29:01  | 1,055         | 39:01  | 2,594         | 54:01  | 0,020         | 16:02    | 2,105         |
| 03:02  | 0,799         | 35:08  | 2,383         | 55:02  | 0,020         | 10:01    | 1,900         |
| 25:01  | 0,731         | 14:02  | 2,162         | 08:09  | 0,010         | 04:02    | 1,827         |
| 33:03  | 0,711         | 40:02  | 2,026         | 08:37  | 0,010         | 01:02    | 1,724         |
| 29:02  | 0,682         | 44:03  | 1,831         | 13:09  | 0,010         | 12:01    | 1,495         |
| 30:04  | 0,624         | 58:01  | 1,479         | 15:29  | 0,010         | 13:03    | 1,265         |
| 30:02  | 0,604         | 15:01  | 1,460         | 15:46  | 0,010         | 11:03    | 1,085         |
| 69:01  | 0,507         | 44:05  | 1,411         | 18:07  | 0,010         | 08:01    | 0,789         |
| 24:03  | 0,487         | 37:01  | 1,391         | 18:11  | 0,010         | 04:01    | 0,760         |
| 66:01  | 0,487         | 41:01  | 1,348         | 31:01  | 0,010         | 04:04    | 0,743         |
| 68:02  | 0,411         | 50:01  | 1,339         | 35:05  | 0,010         | 04:07    | 0,555         |
| 02:02  | 0,312         | 27:05  | 1,277         | 35:11  | 0,010         | 14:04    | 0,518         |
| 01:02  | 0,224         | 57:01  | 1,277         | 37:03  | 0,010         | 08:04    | 0,448         |
| 26:08  | 0,166         | 40:01  | 1,218         | 39:10  | 0,010         | 13:05    | 0,322         |
| 29:10  | 0,156         | 27:02  | 0,968         | 44:04  | 0,010         | 09:01    | 0,313         |
| 02:06  | 0,156         | 15:17  | 0,863         | 44:17  | 0,010         | 08:03    | 0,214         |
| 32:59  | 0,156         | 07:05  | 0,692         | 49:03  | 0,010         | 15:99    | 0,205         |
| 74:01  | 0,088         | 51:08  | 0,647         | 49:11  | 0,010         | 03:85    | 0,175         |
| 02:11  | 0,049         | 41:02  | 0,567         | 50:02  | 0,010         | 15:03    | 0,146         |
| 02:85  | 0,049         | 47:01  | 0,507         | 51:04  | 0,010         | 04:08    | 0,127         |
| 02:17  | 0,039         | 53:01  | 0,497         | 51:65  | 0,010         | 07:27    | 0,097         |
| 31:78  | 0,039         | 40:06  | 0,469         | 53:05  | 0,010         | 11:02    | 0,088         |
| 02:08  | 0,029         | 56:01  | 0,409         | 56:05  | 0,010         | 01:60    | 0,078         |
| 24:17  | 0,029         | 18:05  | 0,401         | 56:40  | 0,010         | 08:02    | 0,078         |
| 74:03  | 0,029         | 73:01  | 0,352         | 81:01  | 0,010         | 13:21    | 0,068         |
| 01:03  | 0,020         | 18:03  | 0,341         |        |               | 01:03    | 0,049         |
| 02:07  | 0,020         | 15:18  | 0,302         |        |               | 04:06    | 0,029         |
| 02:20  | 0,020         | 39:06  | 0,292         |        |               | 11:15    | 0,029         |
| 03:05  | 0,020         | 39:24  | 0,224         |        |               | 16:05    | 0,029         |
| 33:07  | 0,020         | 45:01  | 0,224         |        |               | 04:11    | 0,020         |
| 68:13  | 0,020         | 15:03  | 0,214         |        |               | 04:52    | 0,020         |
| 01:42  | 0,010         | 15:08  | 0,166         |        |               | 04:96    | 0,020         |
| 02:04  | 0,010         | 27:07  | 0,166         |        |               | 07:11    | 0,020         |
| 02:12  | 0,010         | 34:24  | 0,166         |        |               | 08:10    | 0,020         |
| 02:22  | 0,010         | 51:07  | 0,156         |        |               | 11:13    | 0,020         |
| 02:40  | 0,010         | 57:03  | 0,156         |        |               | 11:39    | 0,020         |
| 02:58  | 0,010         | 14:01  | 0,146         |        |               | 12:41    | 0,020         |
| 02:64  | 0,010         | 18:92  | 0,146         |        |               | 13:11    | 0,020         |
| 03:65  | 0,010         | 39:31  | 0,146         |        |               | 14:07    | 0,020         |
| 23:53  | 0,010         | 15:24  | 0,107         |        |               | 14:17    | 0,020         |
| 24:07  | 0,010         | 18:18  | 0,107         |        |               | 15:06    | 0,020         |
| 24:34  | 0,010         | 48:01  | 0,097         |        |               | 04:10    | 0,010         |
| 24:52  | 0,010         | 44:29  | 0,088         |        |               | 04:13    | 0,010         |
| 26:32  | 0,010         | 57:02  | 0,088         |        |               | 11:06    | 0,010         |
| 30:10  | 0,010         | 27:03  | 0,068         |        |               | 11:10    | 0,010         |
| 30:29  | 0,010         | 39:05  | 0,068         |        |               | 11:12    | 0,010         |
| 33:72  | 0,010         | 55:60  | 0,068         |        |               | 11:29    | 0,010         |
| 34:02  | 0,010         | 44:25  | 0,059         |        |               | 12:02    | 0,010         |
| 68:24  | 0,010         | 51:05  | 0,059         |        |               | 13:14    | 0,010         |
| 68:55  | 0,010         | 15:16  | 0,049         |        |               | 13:88    | 0,010         |
| 80:01  | 0,010         | 52:32  | 0,049         |        |               | 14:05    | 0,010         |
|        |               | 07:10  | 0,039         |        |               | 14:16    | 0,010         |
|        |               | 15:09  | 0,039         |        |               |          |               |
|        |               | 27:09  | 0,039         |        |               |          |               |
|        |               | 57:66  | 0,039         |        |               |          |               |
|        |               | 58:47  | 0,039         |        |               |          |               |

**TABLE 5.** Results of the estimated frequency values of the 100 most frequent haplotypes.

| A/A | HLA-A | HLA-B | HLA-DRB1 | FREQUENCY (%) |
|-----|-------|-------|----------|---------------|
| 1   | 02:01 | 18:01 | 11:04    | 2,634         |
| 2   | 01:01 | 08:01 | 03:01    | 2,155         |
| 3   | 02:01 | 51:01 | 16:01    | 1,302         |
| 4   | 24:02 | 18:01 | 11:04    | 1,123         |
| 5   | 02:01 | 51:01 | 11:01    | 1,035         |
| 6   | 24:02 | 35:02 | 11:04    | 0,942         |
| 7   | 33:01 | 14:02 | 01:02    | 0,750         |
| 8   | 32:01 | 40:02 | 16:02    | 0,712         |
| 9   | 02:01 | 51:01 | 11:04    | 0,636         |
| 10  | 11:01 | 35:01 | 01:01    | 0,630         |
| 11  | 02:01 | 51:01 | 14:01    | 0,616         |
| 12  | 24:02 | 51:01 | 11:04    | 0,609         |
| 13  | 01:01 | 52:01 | 15:02    | 0,588         |
| 14  | 02:01 | 35:01 | 14:01    | 0,514         |
| 15  | 02:01 | 39:01 | 16:01    | 0,512         |
| 16  | 11:01 | 52:01 | 15:02    | 0,499         |
| 17  | 03:01 | 35:01 | 01:01    | 0,491         |
| 18  | 02:01 | 13:02 | 07:01    | 0,490         |
| 19  | 02:01 | 44:05 | 16:01    | 0,474         |
| 20  | 30:01 | 13:02 | 07:01    | 0,451         |
| 21  | 02:01 | 51:01 | 13:01    | 0,439         |
| 22  | 24:02 | 08:01 | 03:01    | 0,433         |
| 23  | 02:01 | 27:02 | 16:01    | 0,407         |
| 24  | 02:01 | 44:02 | 11:04    | 0,401         |
| 25  | 02:01 | 44:02 | 16:01    | 0,392         |
| 26  | 03:01 | 07:02 | 15:01    | 0,383         |
| 27  | 32:01 | 44:02 | 11:04    | 0,381         |
| 28  | 02:01 | 35:03 | 14:01    | 0,365         |
| 29  | 01:01 | 57:01 | 07:01    | 0,356         |
| 30  | 03:01 | 18:01 | 11:04    | 0,354         |
| 31  | 01:01 | 37:01 | 16:01    | 0,344         |
| 32  | 03:01 | 18:05 | 16:01    | 0,337         |
| 33  | 02:01 | 18:01 | 16:01    | 0,314         |
| 34  | 11:01 | 18:01 | 11:04    | 0,308         |
| 35  | 23:01 | 49:01 | 11:01    | 0,306         |
| 36  | 02:01 | 35:02 | 11:04    | 0,305         |
| 37  | 32:01 | 51:01 | 11:01    | 0,303         |
| 38  | 23:01 | 44:03 | 07:01    | 0,300         |
| 39  | 24:02 | 39:01 | 16:01    | 0,300         |
| 40  | 02:01 | 35:03 | 11:04    | 0,299         |
| 41  | 32:01 | 18:01 | 11:04    | 0,289         |
| 42  | 11:01 | 35:01 | 14:01    | 0,289         |
| 43  | 03:01 | 07:02 | 16:01    | 0,285         |
| 44  | 02:01 | 18:01 | 14:01    | 0,285         |
| 45  | 24:02 | 18:01 | 16:01    | 0,280         |
| 46  | 02:01 | 08:01 | 03:01    | 0,279         |
| 47  | 03:01 | 41:01 | 04:05    | 0,278         |
| 48  | 24:02 | 51:01 | 04:03    | 0,276         |
| 49  | 11:01 | 51:01 | 11:04    | 0,273         |
| 50  | 01:01 | 18:01 | 11:04    | 0,265         |
| 51  | 02:01 | 73:01 | 04:05    | 0,265         |
| 52  | 26:01 | 51:01 | 16:01    | 0,262         |
| 53  | 11:01 | 35:01 | 16:01    | 0,261         |
| 54  | 03:01 | 13:02 | 07:01    | 0,255         |
| 55  | 68:01 | 18:01 | 11:04    | 0,250         |
| 56  | 03:01 | 51:01 | 11:01    | 0,245         |
| 57  | 25:01 | 18:01 | 15:01    | 0,242         |
| 58  | 02:01 | 51:01 | 01:01    | 0,241         |
| 59  | 01:01 | 51:01 | 11:04    | 0,240         |
| 60  | 24:02 | 35:03 | 11:04    | 0,233         |
| 61  | 03:01 | 38:01 | 13:01    | 0,233         |
| 62  | 02:01 | 51:01 | 07:01    | 0,232         |
| 63  | 02:01 | 51:01 | 04:03    | 0,229         |
| 64  | 26:01 | 55:01 | 16:01    | 0,226         |
| 65  | 24:02 | 35:03 | 14:01    | 0,225         |
| 66  | 02:01 | 18:01 | 11:01    | 0,225         |
| 67  | 24:02 | 55:01 | 16:01    | 0,224         |
| 68  | 02:05 | 50:01 | 07:01    | 0,210         |
| 69  | 01:01 | 35:01 | 01:01    | 0,209         |
| 70  | 11:01 | 18:03 | 11:04    | 0,208         |

|     |       |       |       |         |
|-----|-------|-------|-------|---------|
| 71  | 24:02 | 13:02 | 07:01 | 0,207   |
| 72  | 02:01 | 27:05 | 01:01 | 0,205   |
| 73  | 24:02 | 51:01 | 11:01 | 0,204   |
| 74  | 23:01 | 49:01 | 11:04 | 0,200   |
| 75  | 11:01 | 35:01 | 11:01 | 0,194   |
| 76  | 24:02 | 38:01 | 13:01 | 0,193   |
| 77  | 02:01 | 35:08 | 16:01 | 0,189   |
| 78  | 02:01 | 51:01 | 04:07 | 0,186   |
| 79  | 68:01 | 18:01 | 16:01 | 0,184   |
| 80  | 24:02 | 35:03 | 12:01 | 0,183   |
| 81  | 24:02 | 07:02 | 15:01 | 0,175   |
| 82  | 01:01 | 40:06 | 14:04 | 0,174   |
| 83  | 26:01 | 51:01 | 11:01 | 0,173   |
| 84  | 26:01 | 52:01 | 15:02 | 0,171   |
| 85  | 26:01 | 39:01 | 16:01 | 0,171   |
| 86  | 02:01 | 07:02 | 15:01 | 0,171   |
| 87  | 03:01 | 35:01 | 11:01 | 0,170   |
| 88  | 26:01 | 38:01 | 04:02 | 0,166   |
| 89  | 24:02 | 44:02 | 11:04 | 0,166   |
| 90  | 24:02 | 35:01 | 01:01 | 0,165   |
| 91  | 11:01 | 55:01 | 16:01 | 0,164   |
| 92  | 33:03 | 58:01 | 13:02 | 0,163   |
| 93  | 24:02 | 18:01 | 15:01 | 0,162   |
| 94  | 33:01 | 14:02 | 11:04 | 0,160   |
| 95  | 30:01 | 39:01 | 16:01 | 0,159   |
| 96  | 24:02 | 52:01 | 15:02 | 0,158   |
| 97  | 03:01 | 07:02 | 11:01 | 0,157   |
| 98  | 01:01 | 44:02 | 11:04 | 0,156   |
| 99  | 33:03 | 58:01 | 03:01 | 0,155   |
| 100 | 68:01 | 51:01 | 13:02 | 0,155   |
|     |       |       | SUM   | 36,937% |

**TABLE 6.** Results of the 10 higher positive and negative values of linkage disequilibrium for the observed allele pairs of the three analyzed HLA loci.

| HLA-A-B |       |        |         | HLA-B-DRB1 |       |         | HLA-A-DRB1 |       |         |
|---------|-------|--------|---------|------------|-------|---------|------------|-------|---------|
| A/A     | A     | B      | D>0     | B          | DRB1  | D>0     | A          | DRB1  | D>0     |
| 1       | 02:01 | 51:01  | 0,0228  | 18:01      | 11:04 | 0,0410  | 01:01      | 03:01 | 0,0159  |
| 2       | 01:01 | 08:01  | 0,0217  | 08:01      | 03:01 | 0,0345  | 02:01      | 16:01 | 0,0119  |
| 3       | 11:01 | 35:01  | 0,0161  | 52:01      | 15:02 | 0,0179  | 24:02      | 11:04 | 0,0113  |
| 4       | 02:01 | 18:01  | 0,0158  | 13:02      | 07:01 | 0,0162  | 33:01      | 01:02 | 0,0075  |
| 5       | 33:01 | 14:02  | 0,0115  | 35:01      | 01:01 | 0,0154  | 32:01      | 16:02 | 0,0075  |
| 6       | 32:01 | 40:02  | 0,0105  | 35:02      | 11:04 | 0,0141  | 02:01      | 14:01 | 0,0065  |
| 7       | 03:01 | 07:02  | 0,0096  | 14:02      | 01:02 | 0,0137  | 11:01      | 01:01 | 0,0062  |
| 8       | 30:01 | 13:02  | 0,0090  | 51:01      | 11:01 | 0,0132  | 02:01      | 11:04 | 0,0053  |
| 9       | 24:02 | 35:02  | 0,0090  | 39:01      | 16:01 | 0,0124  | 24:02      | 04:03 | 0,0039  |
| 10      | 23:01 | 49:01  | 0,0075  | 07:02      | 15:01 | 0,0117  | 26:01      | 16:01 | 0,0038  |
| HLA-A-B |       |        |         | HLA-B-DRB1 |       |         | HLA-A-DRB1 |       |         |
| A/A     | A     | B      | D<0     | B          | DRB1  | D<0     | A          | DRB1  | D<0     |
| 1       | 29:02 | 15:01  | -0,1000 | 15:03      | 15:01 | -0,0996 | 69:01      | 10:01 | -0,0973 |
| 2       | 25:01 | 41:01  | -0,0986 | 14:02      | 08:04 | -0,0985 | 26:08      | 14:01 | -0,0965 |
| 3       | 02:02 | 07:02  | -0,0973 | 73:01      | 04:03 | -0,0976 | 68:02      | 15:02 | -0,0961 |
| 4       | 74:01 | 18:01  | -0,0965 | 44:03      | 14:04 | -0,0975 | 02:02      | 13:02 | -0,0957 |
| 5       | 30:01 | 40:06  | -0,0964 | 57:01      | 04:01 | -0,0973 | 66:01      | 10:01 | -0,0935 |
| 6       | 29:02 | 44:05  | -0,0964 | 27:05      | 04:01 | -0,0970 | 24:03      | 10:01 | -0,0934 |
| 7       | 33:03 | 41:01  | -0,0958 | 40:01      | 08:01 | -0,0959 | 25:01      | 13:03 | -0,0933 |
| 8       | 03:01 | 15:24  | -0,0954 | 44:25      | 11:04 | -0,0949 | 69:01      | 04:02 | -0,0933 |
| 9       | 03:01 | 18:18  | -0,0954 | 51:05      | 11:04 | -0,0949 | 32:01      | 04:08 | -0,0932 |
| 10      | 03:01 | 44:182 | -0,0953 | 55:01      | 09:01 | -0,0948 | 01:02      | 13:01 | -0,0921 |

**TABLE 7.** Frequencies of the 15 most frequent common haplotypes of the Greek and the German registry.

| <b>FREQUENCY (%)<br/>GREEK REGISTRY</b> | <b>HLA HAPLOTYPE</b>       | <b>FREQUENCY (%)<br/>GERMAN REGISTRY</b> |
|-----------------------------------------|----------------------------|------------------------------------------|
| 2,634%                                  | A*02:01,B*18:01,DRB1*11:04 | 0,344%                                   |
| 2,155%                                  | A*01:01,B*08:01,DRB1*03:01 | 5,826%                                   |
| 0,750%                                  | A*33:01,B*14:02,DRB1*01:02 | 0,429%                                   |
| 0,630%                                  | A*11:01,B*35:01,DRB1*01:01 | 0,423%                                   |
| 0,491%                                  | A*03:01,B*35:01,DRB1*01:01 | 1,540%                                   |
| 0,490%                                  | A*02:01,B*13:02,DRB1*07:01 | 0,822%                                   |
| 0,451%                                  | A*30:01,B*13:02,DRB1*07:01 | 0,648%                                   |
| 0,407%                                  | A*02:01,B*27:02,DRB1*16:01 | 0,244%                                   |
| 0,383%                                  | A*03:01,B*07:02,DRB1*15:01 | 3,843%                                   |
| 0,356%                                  | A*01:01,B*57:01,DRB1*07:01 | 0,901%                                   |
| 0,300%                                  | A*23:01,B*44:03,DRB1*07:01 | 0,829%                                   |
| 0,279%                                  | A*02:01,B*08:01,DRB1*03:01 | 0,580%                                   |
| 0,242%                                  | A*25:01,B*18:01,DRB1*15:01 | 0,472%                                   |
| 0,210%                                  | A*02:05,B*50:01,DRB1*07:01 | 0,271%                                   |
| 0,207%                                  | A*24:02,B*13:02,DRB1*07:01 | 0,282%                                   |

**TABLE 8.** Estimation values of quantities  $p(n)$  and  $d(n)$  for each of the three scenarios, based on the initial registry of 5131 donors

| INITIAL DONOR REGISTRY SIZE (n = 5131)                     |                                                                                                                                       |                                                                                                                                                               |
|------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SCENARIO                                                   | <i>Probability of finding at least one fully matched donor for the next patient of the registry in need of a transplant</i><br>$p(n)$ | <i>Number of new donors who must be recruited at donor registry size n in order to increase the matching probability for the next patient by 1%</i><br>$d(n)$ |
| (A) High Resolution HLA-A, HLA-B, HLA-DRB1                 | 0,0517                                                                                                                                | 2.796                                                                                                                                                         |
| (B) Low Resolution HLA-A, HLA-B, HLA-DRB1                  | 0,1569                                                                                                                                | 960                                                                                                                                                           |
| (Γ) Low Resolution HLA-A, HLA-B & High Resolution HLA-DRB1 | 0,0809                                                                                                                                | 1.689                                                                                                                                                         |

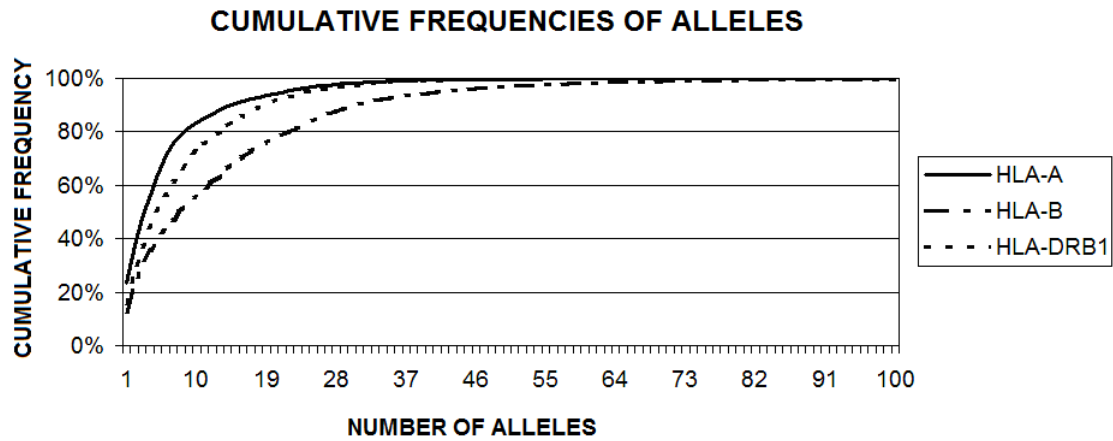


FIGURE 1. Graph of cumulative frequencies of alleles.

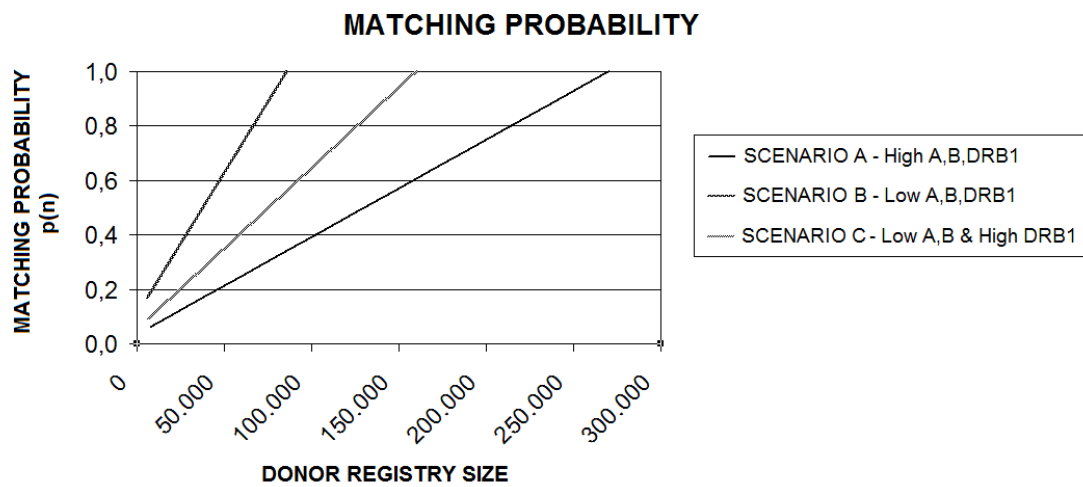


FIGURE 2. Graph of matching probability  $p(n)$  by donor registry size – Linear Model.

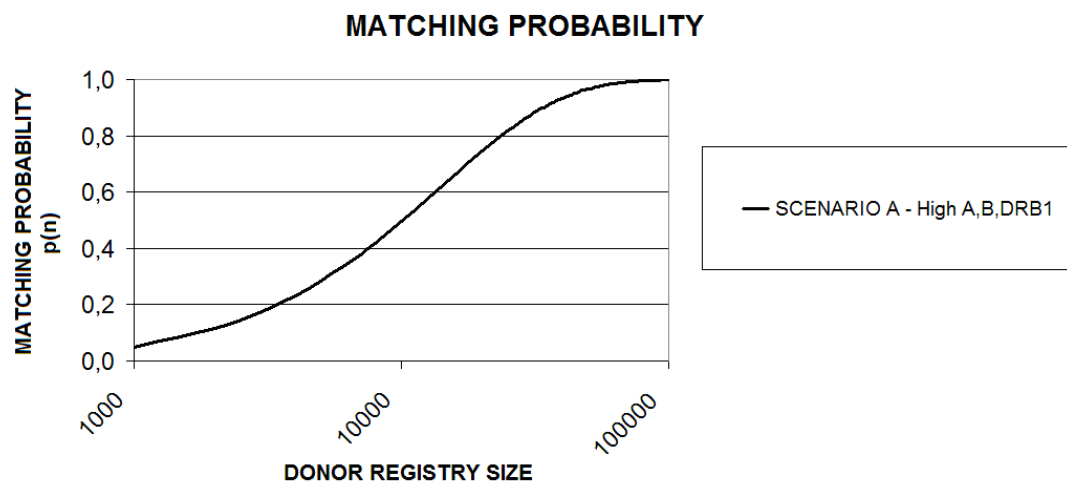
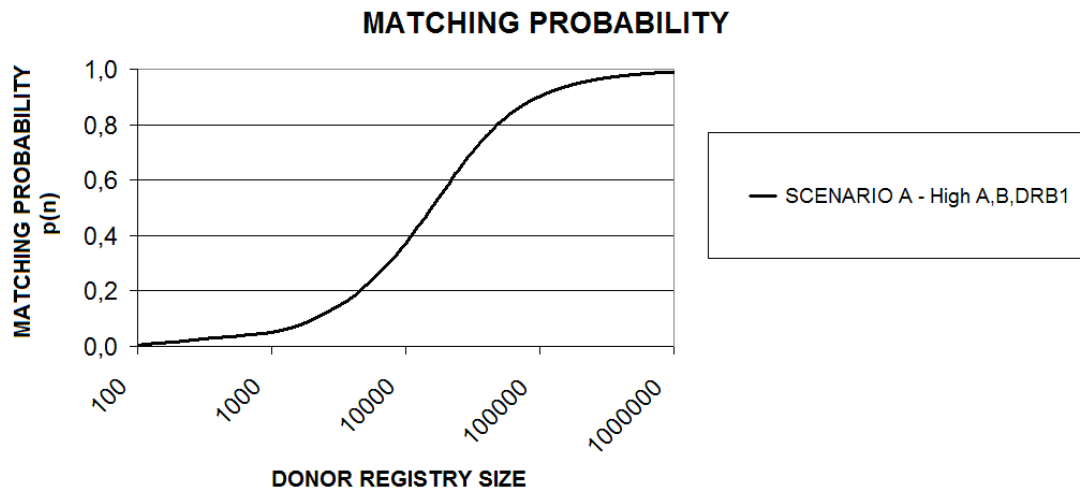
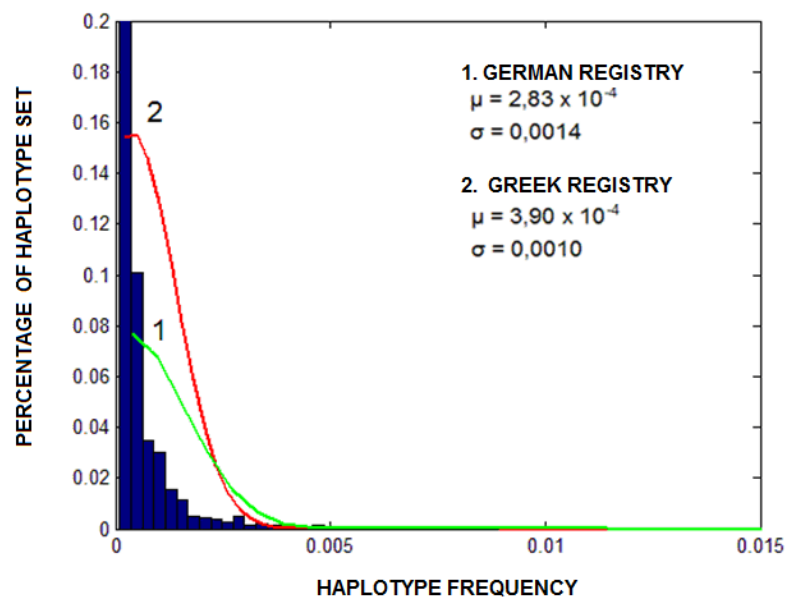


FIGURE 3. Graph of matching probability  $p(n)$  by donor registry size – Discrete Model.



**FIGURE 4.** Graph of matching probability  $p(n)$  by donor registry size – Continuous Model.



**FIGURE 5.** Normal distribution density function curve of the haplotype frequencies of German & Greek registries

# 26<sup>ο</sup>

ΕΛΛΗΝΙΚΗ ΑΙΜΑΤΟΛΟΓΙΚΗ ΕΤΑΙΡΕΙΑ

**Πανελλήνιο Αιματολογικό Συνέδριο**

12 | 13 | 14 Νοεμβρίου 2015, Αθήνα, Ξενοδοχείο HILTON



Αθήνα, 2 Οκτωβρίου 2015

**ΔΙΟΙΚΗΤΙΚΟ ΣΥΜΒΟΥΛΙΟ ΕΑΕ  
ΚΑΙ ΟΡΓΑΝΩΤΙΚΗ ΕΠΙΤΡΟΠΗ**

Πρόεδρος:

Δημήτριος Καρακάσης

Αντιπρόεδροι:

Παναγιώτης Παναγιωτίδης  
Γεώργιος Πατεράκης

Γεν. Γραμματέας:

Ιωάννης Μπαλταδόκης

Ειδικός Γραμματέας:

Μαρία Γκανίδου

Ταμίας:

Μαρία Παγώνη

Μέλη:

Άννα Κολοβού  
Ιωάννης Μπάτσας  
Ευάγγελος Τέρπος  
Παναγιώτης Τσιριγώτης

**ΕΠΙΣΤΗΜΟΝΙΚΗ ΕΠΙΤΡΟΠΗ**

Πρόεδρος:

Ελισάβετ Γρουζή

Μέλη:

Αθηνά-Νόρα Βύνιου  
Μαρία Γαβαλάκη  
Ολγα Κατσαρού  
Παναγιώτης Παναγιωτίδης  
Ιωάννης Παπασωτηρίου  
Ευάγγελος Τέρπος  
Αλέξανδρος Σπυριδωνίδης  
Αικατερίνη Φοίφα

ΓΡΑΜΜΑΤΕΑΣ Ε.Ε.:

Ιωάννης Μπαλταδόκης

Αξιότιμοι κύριοι,

Είμαστε στην ευχάριστη θέση να σας γνωρίσουμε ότι η εργασία σας με τίτλο:

**ΕΚΤΙΜΗΣΗ ΣΥΧΝΟΤΗΤΑΣ ΤΩΝ ΗΛΑ ΑΛΛΗΛΟΜΟΡΦΩΝ & ΑΠΛΟΤΥΠΩΝ  
ΣΕ ΣΥΝΟΛΟ 5.131 ΕΛΛΗΝΩΝ ΕΘΕΛΟΝΤΩΝ ΔΟΤΩΝ ΜΥΕΛΟΥ ΤΩΝ ΟΣΤΩΝ  
ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΓΙΑ ΣΤΡΑΤΗΓΙΚΟ ΣΧΕΔΙΑΣΜΟ ΜΗΤΡΩΟΥ ΔΟΤΩΝ**

έχει επιλεγεί να παρουσιαστεί ως **προφορική ανακοίνωση** στη Συνεδρία του **Σαββάτου 14 Νοεμβρίου 2015 από 16.15 έως 17.15 στην αίθουσα Γ.**

Το **ID παρουσίασης** της εργασίας είναι: **52**

Η ανάπτυξη του θέματος θα είναι διάρκειας **8 λεπτών και θα υπάρχει χρόνος 2 λεπτών για συζήτηση**. Ο χρόνος για κάθε παρουσίαση θα τηρηθεί αυστηρά, ώστε η Συνεδρία να ολοκληρωθεί μέσα στον προβλεπόμενο χρόνο και να αρχίσει η επόμενη, στην καθορισμένη της ώρα.

Με εκτίμηση,

Για την Οργανωτική και Επιστημονική Επιτροπή

Οι Πρόεδροι

Δημήτριος Καρακάσης  
Δ.Σ. ΕΑΕ &  
Οργανωτικής Επιτροπής

Ελισάβετ Γρουζή  
Επιστημονικής  
Επιτροπής

Γραμματεία: **VitaCongress** B. BOYPAZEPH & ΣΙΑ Ε.Ε.

Παπαδιαμαντοπούλου 4 & Βασ. Σοφίας, 115 28 Αθήνα, Τηλ.: 210 72 54 360, Fax: 210 72 54 363

e-mail: info@vitacongress.gr • web: www.vitacongress.gr

**ΕΚΤΙΜΗΣΗ ΣΥΧΝΟΤΗΤΑΣ ΤΩΝ ΗΛΑ ΑΛΛΗΛΟΜΟΡΦΩΝ & ΑΠΛΟΤΥΠΩΝ ΣΕ ΣΥΝΟΛΟ 5.131 ΕΛΛΗΝΩΝ ΕΘΕΛΟΝΤΩΝ ΔΟΤΩΝ ΜΥΕΛΟΥ ΤΩΝ ΟΣΤΩΝ ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΓΙΑ ΣΤΡΑΤΗΓΙΚΟ ΣΧΕΔΙΑΣΜΟ ΜΗΤΡΩΟΥ ΔΟΤΩΝ**

Π. Κορμπάκης<sup>1</sup>, Α. Σαρμά<sup>1</sup>, Γ. Οικονομοπούλου<sup>1</sup>, Ε. Κεφάλα<sup>1</sup>, Δ. Μαλλιώρα<sup>1</sup>, Α. Μουζάκη<sup>1</sup>, Γ. Ανδρουλάκης<sup>2</sup>, Α. Σπυριδωνίδης<sup>1</sup>.

<sup>1</sup>Κέντρο Ενημέρωσης & Προσέλευσης Εθελοντών Δοτών Μυελού των Οστών Πανεπιστημίου Πατρών (ΚΕΔΜΟΠ), Πανεπιστήμιο Πατρών, Ρίο, email: [cbmdpatras@upatras.gr](mailto:cbmdpatras@upatras.gr),

<sup>2</sup>Τμήμα Διοίκησης Επιχειρήσεων, Πανεπιστήμιο Πατρών, Ρίο, email: [gandroul@upatras.gr](mailto:gandroul@upatras.gr)

ΣΚΟΠΟΣ ΤΗΣ ΜΕΛΕΤΗΣ: Έλεγχος ποιότητας του Ελληνικού μητρώου εθελοντών δοτών μυελού των οστών «ΚΕΔΜΟΠ» και υπολογισμός πιθανοτήτων ιστοσυμβατότητας για την βελτιστοποίηση του μητρώου.

ΥΛΙΚΟ & ΜΕΘΟΔΟΣ: Τα δεδομένα συλλέχθηκαν από 5.131 δότες οι οποίοι εγγράφηκαν στο μητρώο του ΚΕΔΜΟΠ μεταξύ 2010-2013 και τυποποιήθηκαν για HLA-A,-B,-DRB1 σε υψηλή ανάλυση. Οι αναλύσεις των HLA δεδομένων πραγματοποιήθηκαν με το ειδικό πακέτο λογισμικού βιοστατιστικής ανάλυσης Arlequin 3.5. Ο προσδιορισμός πιθανοτήτων συμβατότητας μεταξύ ατόμων του μητρώου έγινε με την εφαρμογή του αλγορίθμου EM. Για την εκτίμηση του μεγέθους του μητρώου που θα διασφαλίζει μια πιθανότητα εύρεσης συμβατότητας σε 80% των ατόμων χρησιμοποιήθηκε το διακριτό μοντέλο το οποίο λαμβάνει υπόψη του το πλήθος των μοναδικών φαινοτύπων στο υπό ανάλυση αρχικό μητρώο.

ΑΠΟΤΕΛΕΣΜΑΤΑ: Τα χαρακτηριστικά του δείγματος ως προς το φύλο, την ηλικία και την περιοχή προέλευσης φαίνονται στον πίνακα. Ανευρέθηκαν 77 αλληλόμορφα για το γονίδιο HLA-A, 125 για το HLA-B και 73 για το HLA-DRB1. Οι μεγαλύτερες τιμές της ανισορροπίας γενετικής σύνδεσης (LD) βρέθηκαν μεταξύ των HLA-B και HLA-DRB1 αλληλίων με το ζεύγος B\*18:01/DRB1\*11:04 να εμφανίζει την υψηλότερη σύνδεση ( $D=0,0410$ ). Το πλήθος των διαφορετικών μοναδικών απλοτύπων υπολογίστηκε σε 2.511 με αθροιστική συχνότητα των 100 συχνότερων απλοτύπων ίση με 37%. Σε σύγκριση με δημοσιευμένα στοιχεία του γερμανικού μητρώου DKMS (Human Immunology 2009:895-902), το μητρώο του ΚΕΔΜΟΠ εμφανίζει μεγαλύτερο άνοιγμα της καμπύλης συνάρτησης πυκνότητας απλοτύπων ως δείκτης ποιότητας ενός αρχείου (Εικόνα). Η πιθανότητα εύρεσης στο μητρώο ενός HLA-A,B,DRB1 συμβατού δότη σε επίπεδο αλληλίου (8/8 allele matched) για ένα άτομο του αρχείου (άτομο ελληνικής καταγωγής) είναι 5,17% με την απαίτηση προσθήκης στο μητρώο 2.796 νέων δοτών έτσι ώστε η συγκεκριμένη πιθανότητα να αυξηθεί κατά 1%. Αύξηση του μεγέθους του μητρώου σε 148.000 δότες οδηγεί σε πιθανότητα εύρεσης συμβατού 8/8 δότη για Έλληνες ασθενείς που προσεγγίζει το 80%.

ΣΥΜΠΕΡΑΣΜΑΤΑ: Παρά την ανθρωπογεωγραφία του δείγματος που είναι επικεντρωμένη κατά κύριο λόγο στην Δυτική Ελλάδα και Πελοπόννησο, παρατηρείται μεγάλο πλήθος απλοτύπων οι οποίοι έχουν πολύ μικρή συχνότητα ή παρατηρούνται κατά μοναδικότητα καταδεικνύοντας την ποιότητα του μητρώου. Αύξηση του αριθμού των δοτών σε 148.000 με προσέλκυσή τους από το σύνολο των περιοχών της ηπειρωτικής αλλά και της νησιωτικής Ελλάδας, ώστε να διατηρηθούν τα ποιοτικά χαρακτηριστικά του μητρώου, μπορεί να οδηγήσει σε αυτοεξυπηρέτηση της Ελλάδας σε μοσχεύματα κατά 80%.

| ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ 5.131 ΕΛΛΗΝΩΝ<br>ΕΘΕΛΟΝΤΩΝ ΔΟΤΩΝ ΜΗΤΡΩΟΥ ΚΕΔΜΟΠ |       |                               |       |
|----------------------------------------------------------------|-------|-------------------------------|-------|
| ΦΥΛΟ                                                           |       | ΠΕΡΙΟΧΗ                       |       |
| ΑΝΔΡΕΣ                                                         | 47,6% | ΔΥΤ. ΕΛΛΑΔΑ /<br>ΠΕΛΟΠΟΝΝΗΣΟΣ | 44,4% |
| ΓΥΝΑΙΚΕΣ                                                       | 52,3% |                               |       |
| ΗΛΙΚΙΑ (έτη)                                                   |       | ΑΙΓΑΙΟ                        | 20,4% |
| < 20                                                           | 2,9%  | ΑΤΤΙΚΗ                        | 18,2% |
| 20-30                                                          | 27,3% | ΣΤΕΡΕΑ ΕΛΛΑΔΑ                 | 5,6%  |
| 30-40                                                          | 35,6% | ΜΑΚΕΔΟΝΙΑ                     | 5,0%  |
| 40-50                                                          | 28,6% | ΚΡΗΤΗ                         | 3,9%  |
| > 50                                                           | 5,5%  | ΛΟΙΠΑ                         | 2,4%  |

