

ZÁPADOČESKÁ UNIVERZITA V PLZNI

FAKULTA APLIKOVANÝCH VĚD

Katedra kybernetiky

Podpora tvorby aplikace pro určení kompatibility příznaků
pacient-dárce

Bakalářská práce

Martina Chvojková

Kybernetika a řídicí technika

Vedoucí práce: Ing. Lucie Houdová

Plzeň 2012

ZÁPADOČESKÁ UNIVERZITA V PLZNI
Fakulta aplikovaných věd
Akademický rok: 2011/2012

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: Martina CHVOJKOVÁ
Osobní číslo: A10B0845P
Studijní program: B3918 Aplikované vědy a informatika
Studijní obor: Kybernetika a řídicí technika
Název tématu: Podpora tvorby aplikace pro určení kompatibility příznaků pacient-dárce
Zadávající katedra: Katedra kybernetiky

Z á s a d y p r o v y p r a c o v á n í :

1. Seznamte se s problematikou HLA systému.
2. Seznamte se s klasifikací a metodami typizace HLA antigenů a významem pro výběr dárce kostní dřeně.
3. Vytvořte pravděpodobnostní modely pro určení kompatibility příznaků pacient-dárce.
4. Provedte zhodnocení výsledků

Rozsah grafických prací: dle potřeby
Rozsah pracovní zprávy: 30-40 stránek A4
Forma zpracování bakalářské práce: tištěná
Seznam odborné literatury:

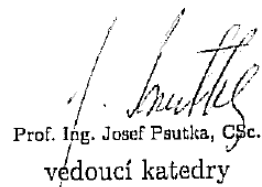
GROFF, J. R. , WEINBERG, P. N. SQL Kompletní průvodce. Computer Press/CP Books, 2001
HURLEY, C. K., ELLIS, J. M. DNA Methods for HLA Typing a Workbook for Beginners. Version 4, Georgetown University, 1993
ŠVOJGROVÁ, M., KOZA, V., HAMPLOVÁ, A. Transplantace kostní dřeně: Průvodce Vaší léčbou. 1. vyd. Plzeň: F. S. Publishing, 2006

Vedoucí bakalářské práce: Ing. Lucie Houdová
Katedra kybernetiky

Datum zadání bakalářské práce: 13. února 2012
Termín odevzdání bakalářské práce: 18. května 2012



Doc. Ing. František Vávra, CSc.
děkan



Prof. Ing. Josef Pautka, CSc.
vedoucí katedry

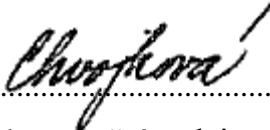
V Plzni dne 13. února 2012

PROHLÁŠENÍ

Předkládám tímto k posouzení a obhajobě bakalářskou práci zpracovanou na závěr studia na Fakultě aplikovaných věd Západočeské univerzity v Plzni.

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně a výhradně s použitím odborné literatury a pramenů, jejichž úplný seznam je její součástí.

V Plzni dne 17. května 2012


.....
vlastnoruční podpis

PODĚKOVÁNÍ

Tímto bych ráda poděkovala všem, kteří mi pomáhali s tvorbou mé práce ať již odbornou radou nebo doporučením literatury, ze které jsem mohla čerpat potřebné informace. Především bych ráda poděkovala vedoucí bakalářské práce paní Ing. Lucii Houdové a panu Danielovi Georgievovi, Ph.D. za cenné konzultace a podněty při realizaci této práce.

Anotace

Bakalářská práce se věnuje vývoji podpůrných statistických aplikací pro určení kompatibility pacient-dárce při transplantaci kostní dřeně. Právě ve dřeni jsou obsaženy důležité molekuly - HLA antigeny, které jsou odpovědné za tkáňovou slučitelnost dvou rozdílných organismů. Práce se zabývá úvodem do problematiky HLA antigenů, vymezuje základní pojmy a potřebné definice statistiky a pravděpodobnosti. Nakonec jsou popisovány podpůrné statistické programy v prostředí MATLAB a Java, které pracují s uváděnými pravděpodobnostními modely.

Klíčová slova

transplantace, MHC, HLA antigeny, pravděpodobnost, testování algoritmu, interaktivní implementace

Annotation

The Bachelor's Thesis discusses the development of supporting statistical applications for determining a patient/donor compatibility of a bone marrow transplantation. Just in a bone marrow is where are contained the important molecules – HLA antigens, which are responsible for the tissue compatibility of two different organisms. This work deals with the introduction to the HLA antigens problems, and defines basic concepts and definitions of statistics and probability. Finally, there are described supporting statistical programs in the MATLAB and Java, which works with a probability models.

Key words

Transplantation, MHC, HLA antigens, probability, testing algorithms, interactive implementation

Seznam zkratek

HLA – Human Leucocyte Antigen – Lidský leukocitový antigen

MHC – Histocompatibility Complex – Hlavní histokompatibilní komplex

HSC – Hematopoetic Stem Cells – Hematopoetické kmenové buňky

DNA – Deoxyribonucleic acid – Deoxyribonukleová kyselina

PCR – Polymerase Chain Reaction – Polymerázová řetězová reakce

NMDP – National Marrow Donor Program

MAC – Multiple Allel Codes.

ČNRDD – Český národní registr dárců dřeně

WHO – World Health Organization – Světová zdravotnická organizace

MATLAB – MATrix LABoratory

GAP – Google App Engine

BMDW – Bone Marrow Donors Worldwide

Obsah

1	Úvod	11
2	Transplantace kostní dřeně	13
2.1	Kostní dřeň.....	13
2.1.1	Funkce kostní dřeně.....	13
2.2	Transplantace.....	14
2.3	Výběr vhodného dárce.....	15
3	HLA antigeny – tkáňové znaky	16
3.1	MHC.....	16
3.1.3	Polymorfismus antigenů	17
3.2	Dědičnost HLA znaků.....	18
3.3	Metody typizace.....	20
3.3.1	Sérologická typizace HLA.....	20
3.3.2	Molekulárně genetická typizace HLA – PCR	20
3.4	Nomenklatura.....	21
3.4.1	NMDP (MAC) kódy	23
3.5	Využití HLA typizace v praxi.....	24
4	Pravděpodobnost	25
4.1	Vymezení základních pojmů	25
4.2	Definice pravděpodobnosti	26
4.3	Podmíněná pravděpodobnost.....	29
4.4	Pravděpodobnostní modely pro práci s databází.....	30
4.4.1	Možnosti využití pravděpodobnostních modelů.....	31
5	Vývoj podpůrných statistických aplikací	34
5.1	Testování algoritmu v prostředí MATLAB.....	35
5.1.1	Koncepce programu – MATLAB.....	35

5.2	<i>Interaktivní implementace v prostředí Java</i>	41
5.2.1	<i>Koncepce programu – Java</i>	42
6	Závěr	47
	<i>Výkladový slovník pojmů</i>	49
	<i>Použité zdroje</i>	51
	<i>Přílohy</i>	54

Seznam obrázků

<i>Obr. 2.1 – Kostní dřev – výroba krvinek</i>	<i>10</i>
<i>Obr. 3.1 – krátké raménko 6. chromozomu</i>	<i>14</i>
<i>Obr. 3.2 – Dědičnost tkáňových znaků</i>	<i>15</i>
<i>Obr. 3.3 – Počet nových alel od roku 1968 do roku 2012, WHO.....</i>	<i>18</i>
<i>Obr. 4.1 – Znázorňující oblast Ω s jevem A</i>	<i>24</i>
<i>Obr. 4.2 – Podmíněná pravděpodobnost</i>	<i>26</i>
<i>Obr. 4.3 – Podmíněná pravděpodobnost pro hod kostkou</i>	<i>26</i>
<i>Obr. 5.1 – Přehled vstupu nových dárců do registrů ČR v průběhu let 2001 – 2011</i>	<i>31</i>
<i>Obr. 5.2 – Četnost jednotlivých alelických skupin antigenu A1</i>	<i>35</i>
<i>Obr. 5.3 – Četnost jednotlivých alelických skupin antigenu A1 pro centrum Plzeň.....</i>	<i>36</i>
<i>Obr. 5.4 – Podmíněná pravděpodobnost sloupce B1 s podmiňujícími prvky A1*1 a DRB1*1*5.....</i>	<i>37</i>
<i>Obr. 5.5 – Geografická statistika HLA</i>	<i>38</i>
<i>Obr. 5.6 – Ukázka převodu dat z Access do MySQL za pomoci skriptu PHP</i>	<i>39</i>
<i>Obr. 5.7 – Geografická statistika – četnosti antigenu A*02 v ČR.....</i>	<i>43</i>

Seznam tabulek

<i>Tab. 3.1 – Platná nomenklatura od 1. dubna 2010</i>	<i>19</i>
<i>Tab. 3.2 – Přesnost rozlišení pro jednotlivé typizace</i>	<i>20</i>
<i>Tab. 3.3 – Relativní riziko onemocnění u osob s určitým typem HLA</i>	<i>21</i>
<i>Tab. 4.1 – Příklad úseku databáze dárců kostní dřevě.....</i>	<i>28</i>
<i>Tab. 5.1 – Příklad úseku databáze v Microsoft Excel</i>	<i>33</i>
<i>Tab. 5.2 – Příklad úseku databáze v MySQL.....</i>	<i>40</i>
<i>Tab. 5.3 – Četnosti antigenu A*02 pro dárcovská centra v ČR</i>	<i>43</i>

1 Úvod

Koncem 50. let 20. století se zrodily zásadní poznatky o imunitním systému lidského těla, konkrétně byly na povrchu buněk objeveny histokompatibilní znaky (HLA znaky), díky kterým bylo umožněno slučovat tkáně dvou rozdílných lidí. Mluvíme o počátku možností transplantací.

Na celém světě najdeme pacienty s chorobami, u nichž je nemožné, je medikamentně léčit. Existuje však způsob, jak jim pomoci k uzdravení, a to díky transplantaci krvetvorných kmenových buněk. Krvetvorné buňky se získávají od vhodných dárců, ať už jde o příbuzné či nikoliv. Aby bylo možné transplantaci uskutečnit, je nutné, aby byla zajištěna mezi dárcem a příjemcem určitá shoda. Ta je ovlivněna tzv. HLA antigeny. Jedná se o tkáňové znaky, které zajišťují slučitelnost tkání. Bohužel 60 – 70 % nemocných nemá nikoho z rodiny, s kým by se jeho HLA znaky shodovaly. Proto vznikají po celém světě tzv. transplantační banky, které v sobě uchovávají informace o dárcích a hlavně o jejich tkáňových znacích.

Získat pro pacienta vhodného dárce není nikterak jednoduché. HLA znaky jsou děděny, z toho vyplývá, že nejčastější možná kompatibilita pacient-dárce je mezi vlastními sourozenci. Přesto to není jediná možnost, jak dárce vypátrat. Jeden z faktorů ovlivňující úspěšnost hledání dárce, je právě zmíněná genetická příbuznost lidí. Nejen v rodině, ale také v regionu či státě. S řešením problémů spojených s vyhledáním nám mohou pomoci tzv. pravděpodobnostní modely. Tyto modely je možné tvořit na základě informací poskytnutých transplantačními bankami.

Současný stav vyhledávacího systému určení kompatibility pacient-dárce není bohužel ve srovnání s registry v Americe či Německu nejefektivnější. Při zpracování dat je statistika prováděna ručně. Dalším problémem je zdlouhavá komunikace transplantačního oddělení s registrem dárců, díky které je prodloužen vyhledávací proces, a tím i stále se zvyšující riziko pro pacienta. Tým Západočeské univerzity po dohodě s Českým národním registrem dárců dřeně (ČNRDD) byl požádán o zlepšení vyhledávacího systému, aby všechny zmíněné nedostatky byly omezeny na minimum. Díky spolupráci, která mi byla umožněna, jsem tak mohla přispět do projektu, kde se mým hlavním cílem a zároveň názvem mé bakalářské práce stala „*Podpora tvorby aplikace pro určení kompatibility příznaků pacient-dárce*“

Jejím cílem je vyvinout podpůrné statistické aplikace pro určení kompatibility pacient-dárce. Práce spočívá v implementaci výše zmíněných pravděpodobnostních modelů do vývojového prostředí MATLAB a programovacího jazyka Java, a tím tak zefektivnit nábor nových dárců do databáze.

Úvod práce se zabývá problematikou transplantací. Významným zdrojem poznatků je obor imunologie, kde je blíže popsána základní funkce kostní dřeně - hematopoéza, která hraje zásadní roli během transplantací. Další kapitola je zaměřena na HLA antigeny, které jsou odpovědné za tkáňovou slučitelnost. Čtvrtá kapitola je věnována pravděpodobnostem, které jsou nezbytnou součástí při hledání dárce. Poslední část práce obsahuje popis programů a rozbor nástrojů využitých při jejich konstrukci v prostředí MATLAB a Java.

2 Transplantace kostní dřeně

2.1 Kostní dřeň

Kostní dřeň (morek) je krvetvorný orgán, vážící přibližně 2600g. Jedná se o rosolovitou tkáň, vyplňující dřeňovou dutinu kostní.

2.1.1 Funkce kostní dřeně

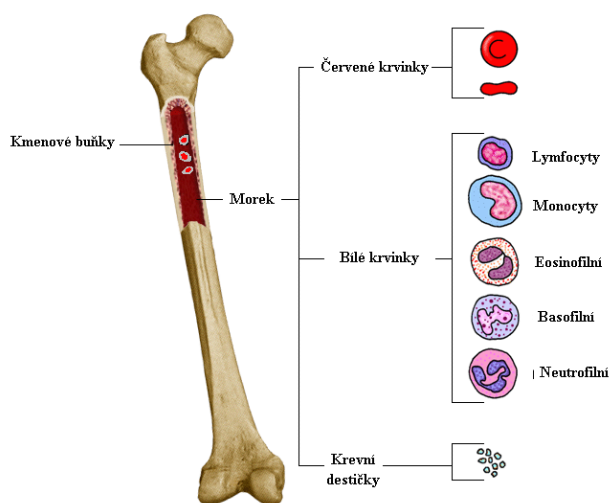
Hlavní funkcí kostní dřeně je hematopoéza neboli krvetvorba. Tento proces je uskutečněn za pomoci tzv. hematopoetických kmenových buněk (HSC – *hematopoietic stem cells*). Původ HSC je v játrech lidského embrya. Ke konci embryonálního vývoje se tyto buňky přemísťují do nově vznikající kostní dřeně. Ze zárodečných HSC vznikají tři hlavní typy krvinek – červené, bílé a krevní destičky (*obr. 2.1*). [1]

Proces krvetvorby začíná v nervové soustavě, kdy mozková centra kontrolují počet krevních buněk v těle. Při ztrátě některého druhu vysílá nervové centrum signál do kostní dřeně, která je okamžitě připravena zvýšit krvetvorbu. To je možné jen v případě, že dřeň disponuje zdravými hematopoetickými buňkami. Ještě nezralé krvinčky se drží v kostní dřeni několik dní, dokud zcela nedospějí. Poté jsou vpuštěny do oběhu, kde konají svoji funkci. Po vyrovnaní složek krve se hematopoéza opět ustálí do běžného provozu.

Krvetvorné buňky se v oběhu ocitají jen ve velmi výjimečných situacích. Například při podávání růstových faktorů v podobě léků či průběhem léčby protinádorových onemocnění. [3]

Postupné stárnutí člověka má za následek, že se kostní dřeň proměňuje. Zhruba po 70 letech lidského života se objem kostní dřeně sníží na polovinu. [4]

Obr. 2.1 – Kostní dřeň – hematopoéza [2]



2.2 Transplantace

Transplantace kostní dřeně je léčebný postup, zejména u hematologických onemocnění (leukémie, aplastická anémie nebo některé druhy imunitních poruch), kdy do těla nemocného vpravujeme zdravé krvetvorné buňky, které nahrazují produkci krve buněk pacientových.

Existují tři možné varianty pro získání dárcovských krvetvorných buněk. První možností je odebrání přímo z kosti, kdy se do hloubky 5-10mm horních lopat kostí kyčelních vnoří jehla až do úrovně, kde se nachází dřevná dutina (*příloha č. 2.1*). [5] „*Vlastní odběr však představuje zátěž jako malý chirurgický výkon srovnatelný s malou plastickou operací. Nejdéle druhý den je dárce propuštěn z nemocnice.*“ [6] Druhá varianta využívá odsátí krvetvorných buněk z periferní krve, kdy jde o odběr přímo z krevního oběhu. Třetí možností je pupečnicková krev odebraná při porodu z pupečnicku (placenty).

Velkou výhodou zmíněného druhu transplantace je, že je kostní dřev v tekutém stavu. K úspěšné transplantaci kostní dřeně stačí pouze malé množství kmenových buněk. Dárci se toto malé množství vpraví přímo do žíly. Krvetvorné buňky si samy v těle dokážou najít pro svůj vývin nejvhodnější místo díky vlastnosti migrace, kterou si osvojily v průběhu dozrávání plodu, kdy musely cestovat z embrya do kostí.

Transplantace kostní dřeně není pro pacienta tak nebezpečná jako u výměny jiných orgánů (srdce, ledviny, atd.). Úskalí číhají až po zákroku, kdy se čeká, jak pacient danou dřev přijme a do jaké míry bude obranyschopnost organismu oslabena. [3]

Typy transplantací

- *Autologní* – auto čili sám – při této transplantaci jsou brány kmenové buňky od samotného pacienta.
- *Alogenní* – kmenové buňky jsou získány od jiného zdravého člověka.
- *Syngenní* – od jednovaječného dvojčete.
- *Xenogenní* – od jedince jiného živočišného druhu.

2.3 Výběr vhodného dárce

Složitost procesu transplantace kostní dřeně není v samotném vpravení krvetvorných buněk do těla, ale v imunitní reakci, která je následuje. [3]

Přirozenou vlastností a úkolem určitých typů bílých krvinek je bránit svůj organismus proti všemu, co do těla nepatří. Bílé krvinky svým receptorem během průchodu organismem kontrolují podezřelé buňky. Pakliže na nějakou cizí narazí, okamžitě se jí pokusí zneškodnit. Proto se během výběru vhodného dárce musíme řídit tzv. tkáňovými znaky (HLA znaky), které jsou obsaženy na všech buňkách, nejen na transplantovaných. Pokud bílá krvinka při putování pacientovým tělem narazí na dárcovu buňku, začne ji testovat na HLA znaky. Pokud se tyto znaky **s pacientovými** do jisté míry shodují, přestane si jí všímat. V opačném případě vyvolá alarm. Buď tělo transplantované buňky vůbec nepřijme, anebo (při menší míře neshody) se transplantované krvinky sice uchytí, ale postupně začnou tělo svého nositele ničit, protože jej vnímají jako „cizí“.

Různých kombinací HLA znaků mezi lidmi jsou bohužel tisíce a tisíce a najít dárce, který by se pro pacienta „hodil“, je nesmírně obtížné. V ČNRDD se „*v současnosti hledá vhodný dárce pro zhruba 250 českých a 450 zahraničních pacientů ročně*“. [6]

Přivykání si na cizí dřeň je dlouhodobé, avšak pokud se vybere podle tkáňových znaků vhodný dárce, je to nejspolehlivější způsob uzdravení.

Abychom pochopili, jak výběr vhodného dárce skutečně probíhá, je nutné si přiblížit, co HLA antigeny jsou a do jaké míry mohou ovlivnit transplantaci kostní dřeně.

3 HLA antigeny – tkáňové znaky

Zkratka HLA je odvozena od anglického názvu – „Human Leucocyte Antigens“ - „lidské leukocytové antigeny“, neboť právě tyto antigeny byly poprvé objeveny na leukocytech, ale jsou přítomny na povrchu všech jaderných buněk. Jejich skladba je vrozená a pro každého člověka zcela jedinečná. HLA antigeny hrají významnou roli v imunitním systému, díky tomu, že spadají pod rozsáhlý komplex – MHC. [7]

Roku 1958 objevil francouzský imunolog J. Dausset první antigen s názvem „Mac“, později HLA-A2 patřícího do rozsáhlého imunitního komplexu. Roku 1965 se právě imunolog Dusset zasloužil o objevení celého systému HLA.

V 70. letech věda přinesla další informace v podobě nových lokusů (místo v DNA, kde se daný gen nachází). Jednalo se o lokusy s názvem HLA -A, -B, -C. V té době bylo dokázáno, že antigeny jsou odpovědné za individuální imunologickou reaktivitu.

V 80. letech byly nalezeny další lokusy HLA -DR, -DQ, -DP a byl pochopen biologický význam HLA společně s jeho genetickými zákony.

90. léta odhalila tzv. neklasické izotypy HLA-E, -F, -G, -H, -J, -K, -X a v současnosti nejenže známe jednotlivá místa, kde se geny nacházejí, ale dokonce může specifikovat i jejich konkrétní formu – alelu. [8]

3.1 MHC

MHC – Major Histocompatibility Complex – hlavní histokompatibilní komplex je soubor 40 – 50 genů seřazených v dlouhém úseku DNA na krátkém raménku 6. chromozomu u lidí. (*obr. 3.1*). MHC se ale pouze u lidí nazývá HLA komplex. MHC geny jsou organizovány do oblastí kódujících tři třídy MHC molekul. [7]

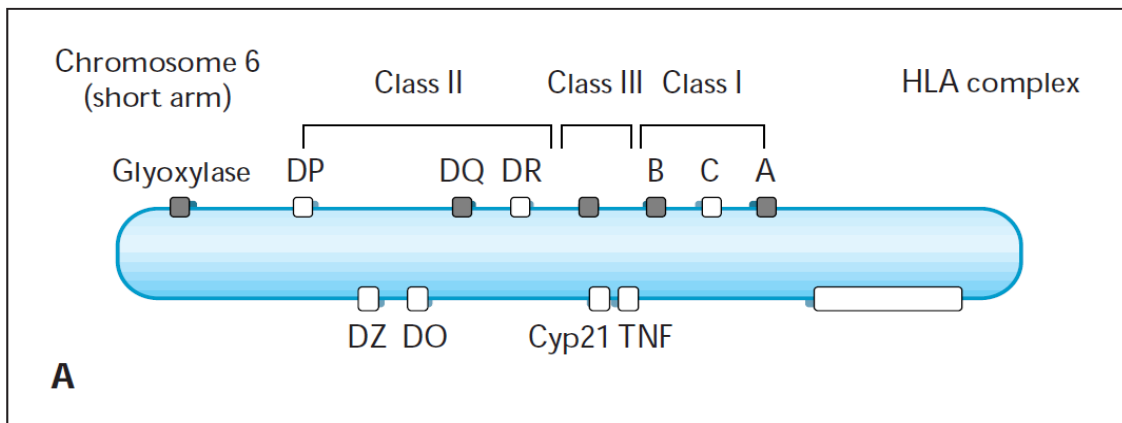
3.1.2 Dělení MHC

První a druhá oblast zahrnuje geny kódující HLA antigeny I. a II. třídy. Zatímco III. třída zahrnuje produkty spojené s imunitními procesy. [7]

V rámci každé třídy MHC existují izotypy:

- **MHC I:** HLA -A, -B, -C: klasické izotypy
HLA -E, -F, -G: neklasické izotypy: ty mají ve srovnání s geny klasickými omezenou buněčnou distribuci a nižší stupeň polymorfismu, z toho HLA-E a HLA-G: přispívají k toleranci plodu v děloze
- **MHC II:** HLA -DR, -DQ, -DP: klasické izotypy
LMP2, LMP7: bílkoviny štěpící proteiny pro HLA I. Třídy
TAP: membránové transportéry
- **MHC III:** C4, C2, TNF (faktor nekrotizující tumory), atd. [9]

Obr. 3.1 – krátké raménko 6. chromozomu [10]



3.1.3 Polymorfismus antigenů

Antigeny jsou velmi polymorfní (variabilní). „Tímto termínem se označuje jev, kdy v lokusu určitého genu (na určitém místě v molekule DNA, kde se tento gen nachází), se u jedinců dané populace vyskytují jeho různé varianty, které se nazývají alely. Aby se určitý gen mohl fenotypově projevit (tj. přepsat se do určitého produktu), musejí se v jeho lokusu nacházet nejméně dvě alely – jedna pochází od matky, druhá od otce. HLA-antigeny však mají ve svých genových lokusech abnormálně velké množství alel, proto tvoří nejpolymorfnější systém v lidském těle.“ [11]

Pokud bude brán v úvahu pro každý gen 100 různých alel, pak teoretická diverzita MHC je 10^{12} . [7]

3.2 Dědičnost HLA znaků

Každý člověk má tzv. fenotyp neboli soubor HLA znaků, který je složen právě ze dvou haplotypů. Každý z haplotypů je tvořen sadou antigenů obsahujících konkrétní alely. Polovinu těchto znaků zdědíme od matky a polovinu od otce.

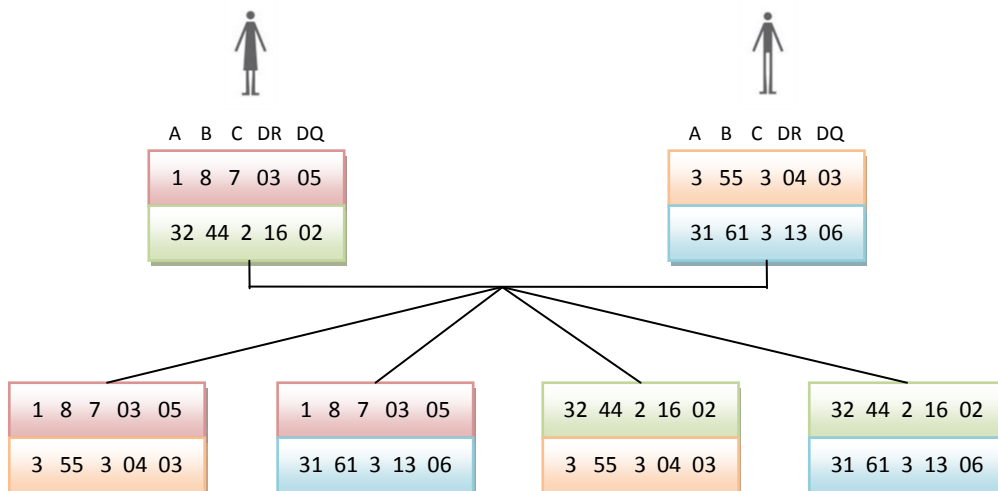
„Z hlediska transplantace se v současné době považují za nejdůležitější (a proto se také nejpřesněji vyšetřují) HLA antigeny I. třídy A, B, C a antigeny II. třídy DR a DQ. Existuje ale řada dalších – tzv. minoritních antigenů, které dosud nejsou dostatečně probádány, a jejich vliv na průběh transplantace se teprve zkoumá.“ [3]

V současnosti je požadavek na míru shody 10/10 neboli v pěti HLA antigenech, konkrétně v (HLA -A, -B, -C, -DRB1 a -DQB1). Nejmenší možná shoda představuje 6/10 v genech (HLA -A, -B, -DRB1), ale zde bohužel pro pacienta vzniká smrtelné riziko odvržení štěpu.

Počet teoreticky možných kombinací HLA znaků u člověka dosahuje několika miliard. Je známo, že některé tkáňové typy (kombinace znaků) se vyskytují v určitém národě či oblasti častěji, jiné jsou extrémně vzácné. Protože se jednotlivé znaky dědí, shodu mezi dvěma jedinci najdeme nejnáze v pokrevním příbuzenstvu. Od rodičů na potomky se příslušná polovina znaků předává obvykle ve zmíněné kompletní sadě (haplotypu).

Pro zjednodušení je uveden příklad, podle kterého je dle genetických zákonů možné dědit jednu ze čtyř možných variant výše zmíněných druhů HLA antigenů mezi sourozenci (obr. 3.2). [3]

Obr. 3.2 – Dědičnost tkáňových znaků [2]



Z obrázku (*obr. 3.2*) vyplývá, že:

- sourozenci mají 25% pravděpodobnost, že budou haploidentičtí,
- 50%, že budou sdílet jeden haplotyp,
- 25%, že budou zcela inkompatibilní.

V širším okruhu rodinných příslušníků (rodič a dítě, sestřenice a bratřenci, strýc a neteř, atd.) je již malá pravděpodobnost úplné shody. Jednu sadu znaků mohou sice sdílet od společných předků, avšak druhá polovina je předávána náhodně z jiné příbuzenské větve, kde se může vyskytnout jen malá míra shody.

Mezi lidmi je tedy velmi obtížně nalézt někoho se shodnými či velmi podobnými tkáňovými znaky, šance je jedna k tisícům až milionům. Možnost objevení takového jedince klesá či stoupá podle skutečnosti, zda je pacientův fenotyp složen ze znaků vyskytujících se v dané populaci častěji nebo naopak vlastní nestandardní kombinaci.

Aby bylo možno uskutečnit vyhledávání, existují nejen v České republice, ale i po celém světě tzv. registry dárců, které shromažďují informace o HLA znacích. [2] Ve světě působí tzv. BMDW – Bone Marrow Donors Worldwide – neboli Světová síť registrů dárců dřeně. V České republice zastřešuje dárce dřeně Český národní registr dárců dřeně – ČNRDD. Registry mezi sebou komunikují v případech, kdy hledají pro pacienta vhodného dárce. Bohužel je někdy možné, že má pacient natolik unikátní složení HLA, že nelze vyhledat vhodného dárce mezi miliony lidí. [2]

3.3 Metody typizace

HLA typizace je rozbor HLA znaků, neboli určení zděděného fenotypu. Typizace konkrétní haplotypy neurčuje, je však možné je vymezit za pomoci k tomu sestavených složitých metod.

Typizaci je možno provádět dvojitým způsobem. Jako první, historicky starší, je sérologické vyšetření. Tato typizace definuje molekuly HLA exprimované na buněčných membránách. Novější metoda je molekulárně genetická typizace. Při níž jsou určovány přímo HLA alely, tj. sekvence nukleotidů kódující HLA antigeny. [12]

3.3.1 Sérologická typizace HLA

Sérologická typizace neboli **mikrolymfocytotoxický test** je typizační metoda sledující reakci protilátek na buněčné membráně. [13]

Bývá uplatňována pro její jednoduchost a nevyžaduje drahé vybavení. Test probíhá poměrně krátkou dobu (během tří hodin). Oproti tomu jsou u sérologické metody i jisté nevýhody. Typizace vyžaduje velké množství krve, životaschopnost lymfocytů a bývá obtížné nalézt vhodné anti-sérum nutné pro reakci se vzácnými antigeny různých populací. Za největší nevýhodu je považována nepřesnost, s jakou jsou HLA typizovány. Proto ji v současné době začínají nahrazovat přesnější molekulárně genetické metody.

3.3.2 Molekulárně genetická typizace HLA – PCR

PCR (polymerázová řetězová reakce) byla vyvinuta v Cetus Corporation v Emeryville v Kalifornii. PCR je metoda umožňující zmnožení určitého úseku DNA do několika kopií v cyklické reakci o třech teplotních fázích. K replikaci slouží velmi malé množství DNA, které může představovat i jedinou molekulu. Abychom na řetězci vymezili místo, které chceme namnožit, osadíme jej po obou stranách tzv. primery. [14] Primer je prvek nukleové kyseliny sloužící jako výchozí bod pro replikaci DNA. [15]

PCR probíhá v termocykleru, tj. zařízení, které dokáže automaticky po jeho naprogramování během několika sekund zvýšit nebo snížit teplotu o několik desítek stupňů °C. [14]

Postup metody:

1. **Denaturace** – Rozpojení řetězců DNA na dva jednovláknové.
2. **Hybridizace** (dosednutí primerů) – Vybrané úseky DNA jsou osazeny primery.
3. **Syntéza DNA** – Při této fázi dochází k opětovnému spojování řetězce.

Kroky 1 – 3 se cyklicky opakují, pro dostatečnou amplifikaci (rozmnožení) původní molekuly DNA obvykle postačuje 30 cyklů. V případě, že na začátku byla ve vzorku pouze jediná molekula DNA, po 32 cyklech teoreticky dostaneme až 1 miliardu nasyntetizovaných molekul. [14]

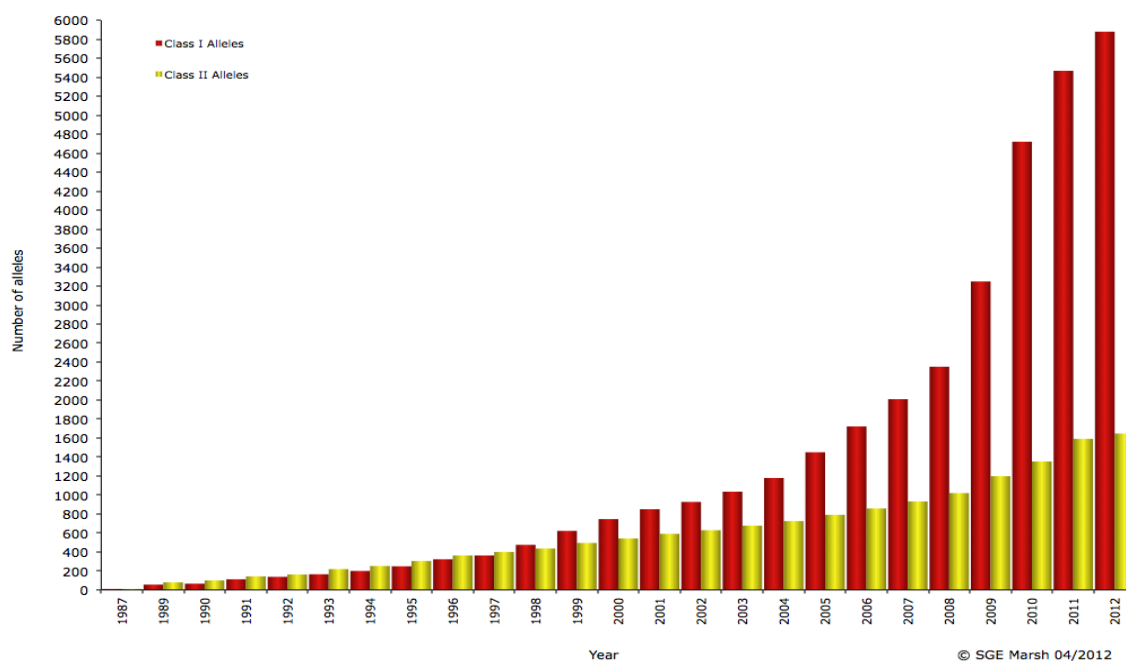
PCR je dále dělena na několik typů:

- a) PCR-SSP = PCR se sekvenčně specifickými primery
- b) PCR-SSO = PCR se sekvenčně specifickými oligonukleotidy
- c) PCR-SBT = Sequencing Based Typing, přímé sekvenování

3.4 Nomenklatura

Od roku 1958, kdy J. Dausset objevil první antigen (viz kap. 3), byl učiněn velký pokrok. Za pomoci nových technologií bylo možné identifikovat další genové varianty. Avšak od počátku si vědci uvědomovali, že polymorfismus alel bude zastoupen širokou škálou antigenů a bude nutné zavést univerzální názvosloví pro určování HLA znaků – Nomenklaturu. [16]

Obr. 3.3 – Počet nových alel od roku 1968 do roku 2012, WHO [17]



Proto byl roku 1968 ustanoven Nomenklaturní výbor Světové zdravotnické organizace (WHO). Tento výbor každoročně označuje nově objevené HLA alely. Aktuální informace pak prezentuje na <http://hla.alleles.org/nomenclature/index.html>. [11] Podle poslední aktualizace nomenklatury z roku 2012 WHO existuje již více než 5500 alel I. třídy a 1600 alel II. třídy (*obr. 3.3*). [17]

Předchozí verze nomenklatury z roku 2004 začala být nevyhovující z důvodu nárůstu počtu nově popsanych alel. Například u nejpočetnějších alelických skupin (HLA-A2, -B15) byla vyčerpána kapacita pojmenování. Nomenklaturní výbor začal na tuto skutečnost reagovat a připravil novou verzi nomenklatury platící od 1. dubna 2010 spolu s oficiální databází sekvencí HLA alel.

Specifické oblasti antigenů jsou nově odděleny dvojtečkou (A*02:101). Dále nomenklatura umožňuje pokrýt HLA alely, které sdílejí identickou sekvenci domén (např. A*02:01:01G) nebo také alely kódující identické vazebné domény HLA molekuly (např. A*02:01P) (*tab. 3.1*).

Přechod od starší verze nomenklatury k nové je směřován ke zjednodušení HLA výsledků pro klinickou praxi, zvláště pak pro využití v typizaci pacientových a dárcovských kmenových buněk. [16]

Tab. 3.1 – Platná nomenklatura od 1. dubna 2010 [18]

Nomenklatura	Označuje:
HLA	HLA region a předpona pro HLA gen
HLA-DRB1	Specifický HLA lokus, např. DRB1
HLA-DRB1*13	Skupina alel, které kódují DR13 antigen
HLA-DRB1*13:01	Specifická HLA alela
HLA-DRB1*13:01N	„Null“ alela (alela, která není exprimována)
HLA-DRB1*13:01:02	Alela, která se liší synonymní mutací
HLA-DRB1*13:01:01:02	Alela s mutací mimo kódující region
HLA-A*24:09N	„Null“ alela
HLA-A*30:14L	Alela kódující protein se signifikantně redukovanou nebo „nízkou“ expresí na buňkách
HLA-A*24:02:01:02L	Alela kódující protein se signifikantně redukovanou nebo „nízkou“ expresí na buňkách a kde se mutace nachází mimo kódující region
HLA-B*44:02:01:02S	Alela kódující protein, který je exprimován pouze jako „sekreční“ molekula
HLA-A*03:01:01Q	Alela s mutací ovlivňující expresi na povrchu buněk. Avšak nebylo dosud jednoznačně potvrzeno a její expresní status zůstává zatím 'Questionable'

Přesnost rozlišení typizace

HLA typizace je schopna rozlišit konkrétní antigen v sérologii nebo antigen i s jeho specifickou alelou v případě DNA typizace.

Sérologická typizace zachycuje „Broady“ a „Splity“. Broad je soubor antigenů, neboli jednotlivých splitů. Split představuje daný antigen.

DNA typizace je už přesnější, pracuje na několika úrovních rozlišení (nízká, střední, vysoká) (tab. 3.2). V praxi nízkou a střední úroveň najdeme při HLA typizaci dárců během vstupu do registrů. Vysoké rozlišení se používá pro určení kompatibility mezi pacientem a dárcem. Bohužel se vysoká úroveň užívá méně často, důvodem je finanční nákladnost.

Tab. 3.2 – Přesnost rozlišení pro jednotlivé typizace [19]

Příklad	Rozlišení		Popis
	Sérologie	DNA	
A9	Broad	-	Nejistý antigen – obvykle soubor splitů
A3	Split	-	Jistý antigen
A*02:XX OR A*02	-	LOW	Možnost rozložit na soubor alel, antigenu
A*02:AB	-	Intermediate	Množina alel s vysokým rozlišením. Kde AB představuje NMDP kód.
A * 0201	-	HIGH	Jistá alela – nejlepší výsledek

3.4.1 NMDP (MAC) kódy

NMDP (National Marrow Donor Program) kódy nebo také MAC – Multiple Allel Codes. Jedná se o kódy „zahrnující skupinu alel získaných DNA typizací, ze kterých je **právě jedna** platná.“ [20]

Tyto kódy jsou zobrazovány pomocí typizačního přístroje. Příklad MAC kódu a v něm obsažená množina alel:

- AB – {01, 02}
- XA – {02, 21}
- DKRA – {10, 16, 69}
- HFNT – {01,08,11,13,14,17}

3.5 Využití HLA typizace v praxi

Antigeny mají v lékařství velký význam. Příkladem jsou jedinci, kteří díky určitým HLA molekulám mají „lepší“ obranyschopnost organismu. Především se to týká těch, kteří mají znak HLA-DR3. Pakliže se tyto lidé setkají s nákazou, jejich imunitní systém na ni zareaguje a zneškodní ji, než daný jedinec onemocní. Oproti tomu obranyschopnost u lidí se znakem HLA-DR7 není příliš výjimečná. Tělo proti infekci zakročí, avšak ne nějak mimořádně.

Právě HLA molekuly mají význam při vzniku autoimunitních onemocnění. Například výše zmíněný znak HLA-DR3 se u autoimunit nachází velmi často. Bohužel je to stinná stránka tohoto znaku. Síla imunitní reaktivity je rozvinuta na úkor přesnosti rozlišení vlastního od cizího. Ačkoli jedinec s DR3 (dále také se znaky A1, B8) snáze překoná nápor infekce, je to za cenu toho, že se v určitém okamžiku obranyschopnost obrátí proti vlastním orgánům i tkáním.

Jasnou souvislost s HLA má nemoc zvaná Bechtěrevova choroba (ankylózuující spondylartritida). Při této nemoci je použita sérologická typizace, kdy se zkoumá přítomnost znaku HLA-B27. Právě tento znak se nachází až u 98 % pacientů s touto chorobou. Výskyt HLA-B27 není znamením, že pacient onemocní ankylózuující spondylartritidou. U kavkazské populace je Bechtěrevova choroba zastoupena u 9%, z toho jedinců s HLA-B27 onemocní asi jen 5 – 20 %. [21]

Relativní riziko různých autoimunitních onemocnění v závislosti na HLA antigenech obsažených v krvi znázorňuje tabulka níže (*tab. 3.3*).

Tab. 3.3 – Relativní riziko onemocnění u osob s určitým typem HLA [21]

Onemocnění	HLA	Relativní riziko*
Ankylózuující spondylartritida (Bechtěrevova choroba)	B27	87,4
Goodpastureův syndrom	DR2	15,9
Pemphigus	DR4	14,4
Uveitida	B27	10,0
Roztroušená skleróza	DR2	4,8
Revmatoidní artritida	DR4	4,2

* Relativní riziko udává, kolikrát častěji se dané onemocnění vyskytuje u jedinců s danou alelickou formou HLA než u jedince, kteří nenesou tuto alelu.

4 Pravděpodobnost

Pravděpodobnostní přístup byl v této práci využíván pro výpočty genetické diverzity dárců kostní dřeně v ČNRDD. Následující kapitola je citována z učebních textů [22] a [23].

4.1 Vymezení základních pojmů

Náhodný pokus je proces, který má alespoň dva možné výsledky. A právě jeden z těchto výsledků není jednoznačně určen podmínkami, za kterých je prováděn – hod kostkou, sázka do loterie, zhotovení výrobku, ...

Množina všech možných výsledků náhodného pokusu zahrnuje všechny možné jevy, které mohou nastat, označuje se Ω – u kostky $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Jev jistý je označení pro celou množinu Ω .

Náhodný jev se značí jako A, B, C, \dots ($A, B, C, \dots \subset \Omega$). Reprezentuje výsledek náhodného pokusu. U kostky lze náhodné jevy označit: $A \dots$ padne 1, $B \dots$ padne 2, ...

Jev nemožný je představován prázdnou množinou \emptyset .

Základní operace s jevy: protože jevy mají charakter podmnožin, je možné k nim přistupovat pomocí množinových operací, používat množinovou symboliku a graficky znázornit pomocí Vennových diagramů.

- Sjednocení jevů $A \cup B$ (nastane jev A nebo jev B , nebo oba jevy zároveň).
- Průnik jevů $A \cap B$ (nastane jev A a zároveň jev B).
- Pokud $A \cap B = \emptyset$ označujeme jevy jako **neslučitelné**.
- Rozdíl jevů $A \setminus B$ (nastane jev A , ale nikoliv jev B).
- $A \subset B$ jev A je podjevem jevu B .
- $A \equiv B$ rovnocenné jevy.
- Doplněk jevu (jev opačný) A je jev $\bar{A} \equiv \Omega \setminus A$
- Platí $A \cap \bar{A} = \emptyset$ a $A \cup \bar{A} = \Omega$

De Morganova pravidla

- $\overline{A_1 \cup A_2 \cup \dots \cup A_n} = \overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_n}$
- $\overline{A_1 \cap A_2 \cap \dots \cap A_n} = \overline{A_1} \cup \overline{A_2} \cup \dots \cup \overline{A_n}$

Úplný systém neslučitelných jevů: jevy A_1, A_2, \dots, A_n nazveme úplným systémem neslučitelných jevů, pokud pro ně platí:

- jsou neslučitelné, tj. $\forall i, j (i \neq j)$ platí $A_i \cap A_j = \emptyset$
- $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$

Např. na kostce je úplným systémem neslučitelných jevů $A_i = \text{padne } i, i = 1, 2, \dots, 6$.

Jev nazveme **elementárním jevem**, pokud neexistují dva jevy $A_1 \neq A$ a $A_2 \neq A$ tak, že by $A = A_1 \cup A_2$

4.2 Definice pravděpodobnosti

Def. 1: Pravděpodobnost jevu

Každému jevu A přiřazujeme reálné číslo $P(A)$. Pravděpodobnost lze chápat jako předpověď poměrných četností výsledku při mnohonásobném opakování daného pokusu; pravděpodobnost lze chápat jako kvantitativní ohodnocení stupně jistoty. Existuje několik možností jak pravděpodobnost zavést – klasická definice, geometrická definice, statistická definice a axiomatická definice.

Def. 2: Klasická definice pravděpodobnosti

Předpoklady:

- Prostor elementárních jevů Ω se skládá z konečného počtu elementárních jevů $E_1, E_2, \dots, E_N, 0 < N < +\infty$
- p_1, p_2, \dots, p_n , jsou nezáporná čísla splňující $\sum_{i=1}^N p_i = 1$
- Každý elementární jev má stejnou možnost výskytu a jejich pravděpodobnosti jsou tedy shodné $p_i = \frac{1}{n}, i = 1, 2, \dots, N$
- Každý jev A lze popsat množinou jevů $\{E_{i_1}, E_{i_2}, \dots, E_{i_k}\}$, kde E_i jsou výsledky pokusu příznivé jevu A

Pak pravděpodobnost jevu A je definována jako:

$$P(A) = \frac{N_A}{N} \quad (4.1)$$

kde N_A je počet výsledků příznivých jevu,

N je počet všech možných výsledků.

Příklad 4.1 Jev A ...na kostce padne sudé číslo. Elementární jevy jsou v tomto případě: E_1 ...na kostce padne číslo 1, ..., E_6 ...na kostce padne číslo 6. Pravděpodobnost jevu A určíme tedy jako poměr jevů, které jev A obsahuje (padne 2, 4, nebo 6) ke všem možným jevům (těch je 6). Celkem tedy dostáváme:

$$P(A) = \frac{3}{6} = 0,5$$

Def. 3: Statistická definice pravděpodobnosti

Náhodný pokus opakujeme N -krát a pozorujeme výskyt jevu A v M případech. Poměr $\frac{M}{N}$ se označuje jako poměrná (relativní) četnost jevu A . Pravděpodobnost jevu A je pak:

$$P(A) = \lim_{N \rightarrow \infty} \frac{M}{N} \quad (4.2)$$

Def. 4: Geometrická definice pravděpodobnosti

Předpoklady:

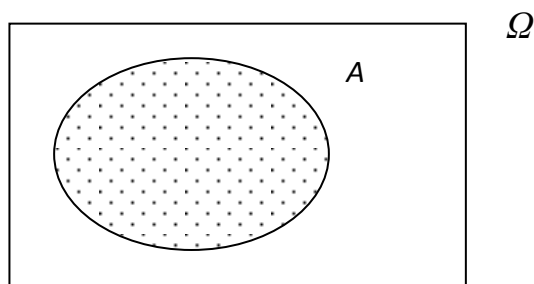
- Ω jsme schopni vyjádřit jako neprázdnou omezenou oblast v \mathbb{R}^n
- Jev A jsme schopni vyjádřit jako podoblast Ω

Pak pravděpodobnost jevu vyjádříme jako:

$$P(A) = \frac{\lambda(A)}{\lambda(\Omega)} \quad (4.3)$$

Kde $\lambda(A)$ je míra oblasti A (např. délka, obsah) a $\lambda(\Omega)$ je míra oblasti Ω (obr. 4.1).

Obr. 4.1 – Znázorňující oblast Ω s jevem A



Axiómy pravděpodobnosti

- Pro jev A platí $0 \leq P(A) \leq 1$.
- Pro neslučitelné jevy $A_1, A_2 \dots$ platí: $P(\cup_i A_i) = \sum_i P(A_i)$;
- platí, že $P(\Omega) = 1$, kde Ω je jev jistý a $P(\emptyset) = 0$, kde je jev nemožný.

Def. 5: Pravděpodobnost opačného jevu

Z definice jevu \bar{A} opačného k jevu A plyne, že:

- tyto jevy jsou neslučitelné $A \cap \bar{A} = \emptyset$,
- jejich sjednocení je jev jistý $A \cup \bar{A} = \Omega$.

Odtud dostáváme vztah pro pravděpodobnost opačného jevu pomocí pravděpodobnosti jevu původního:

$$P(\bar{A}) = 1 - P(A) \quad (4.4)$$

Využití pravděpodobnosti opačného jevu se opírá o skutečnost, že výpočet opačného jevu je někdy jednodušší, než výpočet jevu původního.

Def. 6: Pravděpodobnost sjednocení jevů

Pro pravděpodobnost sjednocení jevů A a B platí následující vzorec:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (4.5)$$

Pro **neslučitelné jevy** platí $P(A \cap B) = 0$, a proto pro sjednocení neslučitelných jevů A a B platí vzorec:

$$P(A \cup B) = P(A) + P(B) - 0 = P(A) + P(B) \quad (4.6)$$

Def. 7: Sdružená pravděpodobnost

Sdružená pravděpodobnost nezávislých diskrétních veličin jest součinem marginálních pravděpodobností.

$$P(A|B) = P(A)P(B|A) = P(B)P(A|B) \quad (4.7)$$

4.3 Podmíněná pravděpodobnost

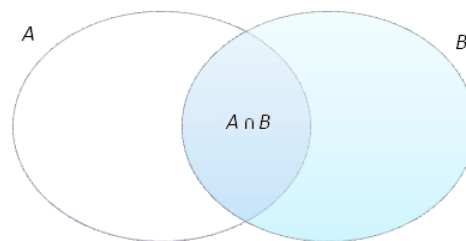
Def. 8: Podmíněná pravděpodobnost

Podmíněná pravděpodobnost jevu A , za podmínky, že **již nastal jev B** ($P(B) > 0$) je definována jako:

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (4.8)$$

Ilustrace je uvedena na následujícím obrázku (obr. 4.2). Jev B již nastal (určitě tedy jev, který nastane, bude v množině B), proto je pravděpodobnost toho, že nastane i jev A rovna poměru plochy průniku k ploše kterou zaujímá jev B .

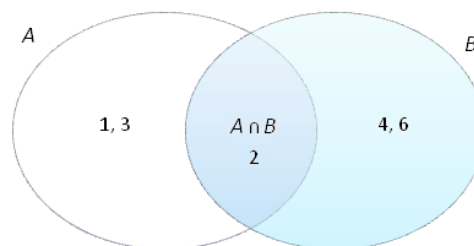
Obr. 4.2 – Podmíněná pravděpodobnost



Příklad 4.2 Na kostce si označíme jevy: A ... padne 1, 2, nebo 3
 B ... padne sudé číslo

Ptáme se, kolik je $P(A|B)$, tedy jaká je pravděpodobnost toho, že nastal jev A (ten nastal v případě, že na kostce padlo 1, 2, nebo 3), za předpokladu, že víme, že již nastal jev B (tj. na kostce padlo sudé číslo). $P(A \cap B)$ tvoří jev: padne číslo 2. Grafické znázornění najdeme na obrázku níže (obr. 4.3).

Obr. 4.3 – Podmíněná pravděpodobnost pro hod kostkou



$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3}$$

4.4 Pravděpodobnostní modely pro práci s databází

Pravděpodobnostní modely této podkapitoly jsou využívány při interaktivní implementaci v prostředí Java. Vzorce 4.9 – 4.15 slouží k výpočtu genetické diverzity dárců dřeně v ČNRDD. Pravděpodobností vzorce jsou čerpány z odborných textů [19].

Jednoduchá pravděpodobnost:

$$P_D(A = x) = P(A = x | MAS_x) P(MAS_x) \quad (4.9)$$

$$P(MAS_x) = \frac{|\{d \in MAS_x\} DNR|}{|DNR|} \quad (4.10)$$

$$P(A = x | MAS_x) = \frac{P_c(A = x)}{\sum_{x' \in MAS_x} P_c(A = x')} \quad (4.11)$$

Podmíněná pravděpodobnost:

$$P_D(A = x | B = y) = \frac{P(A = x, B = y)}{P_c(B = y)} \quad (4.12)$$

$$P(A = x, B = y) = P(A = x, B = y | MAS_{xy}) P(MAS_{xy}) \quad (4.13)$$

$$P(MAS_{xy}) = \frac{|\{d \in MAS_{xy}\} DNR|}{|DNR|} \quad (4.14)$$

$$P(A = x, B = y | MAS_{xy}) = \frac{P_c(A = x, B = y)}{\sum_{x', y' \in MAS_{xy}} P_c(A = x', B = y')} \quad (4.15)$$

Kde:

$P_c(-)$ – pravděpodobnost založená na výskytu daných alel (či haplotypu) v kavkazské populaci

$P_D(-)$ – posteriorní pravděpodobnost výskytu alel vycházející z dat dárců ČNRDD

MAS_h – skupina MAS obsahující alelu (nebo např. i haplotyp) h

DNR – skupina dárců v databázi ČNRDD

$/S/$ – mohutnost datového souboru

4.4.1 Možnosti využití pravděpodobnostních modelů

Ke vzorcům 4.9 – 4.15 jsou vytvořeny krátké příklady, které využívají údaje z databáze dárců v podobě informací o jejich transplantačních HLA znacích. V příkladu 4.3 je počítána pravděpodobnost výskytu pro gen A, v příkladu 4.4 je počítána podmíněná pravděpodobnosti pro geny A a B, kdy se gen A stává podmiňujícím.

Tabulka (tab. 4.1) znázorňuje úsek dat z databáze, kdy v prvním řádku jsou vypsané názvy jednotlivých genů a v ostatních řádcích pak před dvojtečkou jejich specifické antigeny př.: 05:08/12/24 s jejich jednotlivými možnostmi výskytu alel za dvojtečkou – př.: 05:08/12/24.

Tab. 4.1 – Příklad úseku databáze dárců kostní dřeně

A	B	DRB1
05:08/12/24	06:28	09:15
05:08	06:AD	09:KL
05:08/11	06:01/02/28	
05:MTI	06:28	
05:08		
05:08		
05:08		
05:08/11		

NMDP kódy: MTI – {08, 12, 24}

AD – {01, 28}

KL – {15, 24}

Příklad 4.3 Výpočet pravděpodobnosti

Výpočet pravděpodobnosti dle vzorce 4.9, kdy se v databázi hledá požadovaný gen, antigen a jeho specifická alela. Příklad hledané hodnoty: A*05:08.

Kde: A – gen,

05 – antigen,

08 – specifická alela.

Pak za A je dosazena do vzorce 4.9 hledaná hodnota $A^*05:08$. Dále je počítáno dle vzorce 4.11 v případě $a)$ a vzorce 4.10 v případě $b)$.

$$P_D(A^*05:08) = \underbrace{P(A^*05:08|MAS_x)}_a) \cdot \underbrace{P(MAS_x)}_b)$$

$$a) P(A^*05:08|MAS_x) = \frac{P_c(A^*05:08)}{\sum_{x' \in MAS_x} P_c(A^*05:08')} = \frac{1}{n} = \frac{1}{4} \quad \text{kde } n = \{08,11,12,24\}$$

$$b) P(MAS_x) = \frac{|d \in MAS_x| DNR|}{|DNR|} = \frac{8}{|DNR|}$$

$$P_D(A^*05:08) = \frac{1}{4} \cdot \frac{8}{|DNR|} = \frac{2}{|DNR|}$$

Kde:

n – z alel příslušících danému antigenu (př.: 05) je utvořena množina A , pak n je množina B obsahující unikátní prvky množiny A

Příklad 4.4 Výpočet podmíněné pravděpodobnosti

Výpočet podmíněné pravděpodobnosti dle vzorce 4.12, kdy se v databázi hledá požadovaný gen, antigen a jeho specifická alela za podmínky, že je informace o $A^*05:08$ už známá. Hledané $B^*06:28$ za podmínky, že již nastalo $A^*05:08$.

Kde: B – gen,
06 – antigen,
28 – specifická alela.

Pak za A je do vzorce dosazena hledaná hodnota – $B^*06:28$ a za B hodnotu, která je již známá – $A^*05:08$. Dále je počítáno dle vzorce 4.9 v případě $a)$ a vzorců 4.13-15 v případě $b)$.

$$P(B^*06:28|A^*05:08) = \frac{\overbrace{P(B^*06:28, A^*05:08)}^{b)}}{\underbrace{P_c(A^*05:08)}_{a)}}$$

a) Výpočet pravděpodobnosti pro $A^*05:08 \rightarrow P(A^*05:08) = \frac{2}{|DNR|}$ (viz výpočet příkladu 4.3)

b) $P(B^*06:28, A^*05:08) =$

$$= \underbrace{P(B^*06:28, A^*05:08|MAS_{xy})}_{b_1)} \underbrace{P(MAS_{xy})}_{b_2)} = \frac{1}{8} \cdot \frac{4}{DNR} = \frac{1}{2DNR}$$

$$b_1) \quad P(B^*06:28, A^*05:08|MAS_{xy}) = \frac{1}{n} = \frac{1}{8}$$

$$n = ([08/12/24] \cdot [28]) + ([08] \cdot [01/28]) + ([08/11] \cdot [01/02/28]) \\ + ([08/12/24] \cdot [28]) = 8$$

$$b_2) \quad P(MAS_{xy}) = \frac{|d \in MAS_{xy} \setminus DNR|}{|DNR|} = \frac{4}{DNR}$$

$$P(B^*06:28|A^*05:08) = \frac{\frac{1}{2DNR}}{\frac{1}{DNR}} = \frac{DNR}{2DNR} = \frac{1}{2}$$

Kde:

n – kartézský součin množin x_n a y_n , - $(x_n \times y_n)$, kde x_n je množina tvořená po řádcích výčtem prvků za dvojtečkou př.: 05:08/12/24 v konkrétním sloupci (sloupec A) a y_n je množina druhého sloupce (B), pak n je množina unikátních kartézských součinů

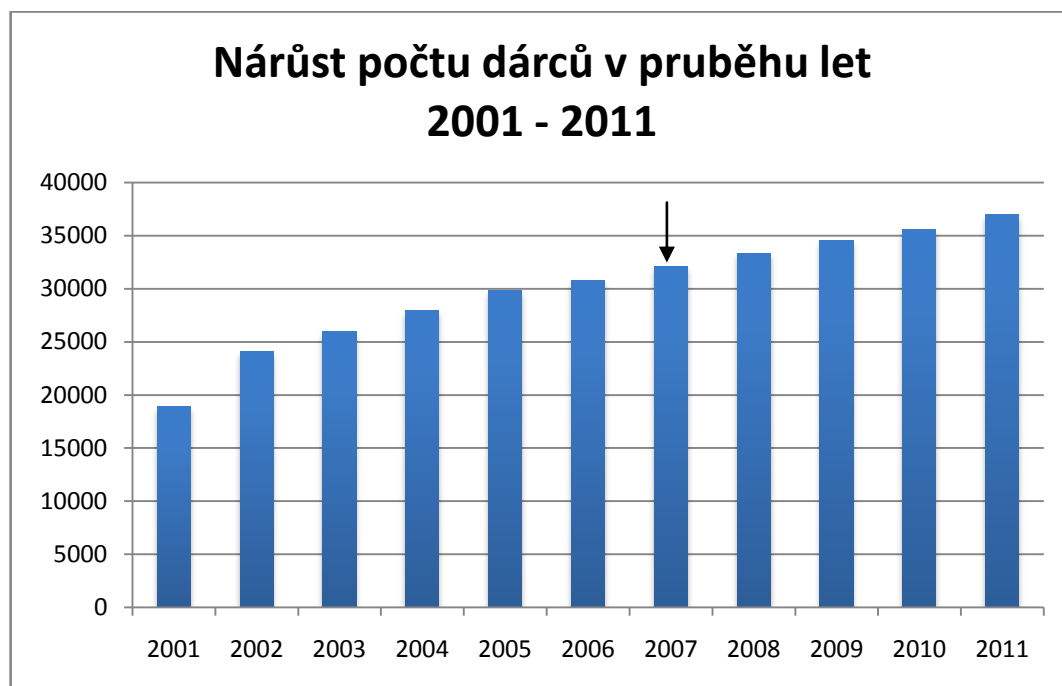
5 Vývoj podpůrných statistických aplikací

Na počátcích projektu byla zvolena práce na prototypu ve vývojové prostředí MATLAB. Cílem funkčnosti programu bylo prohledávání databáze na úrovni antigenů, kdy jednotlivé výstupy měli graficky zobrazovat výčet četností antigenů a podmíněných pravděpodobností. Projekt pro ČNRDD byl dále rozšířen o implementaci v programovacím jazyce Java. Program v Javě, stejně jako MATLAB, pracuje na bázi výpočtu pravděpodobností rozšířené o vyhledávání na úrovni alel.

Velkou výhodou toho projektu byla možnost práce nad skutečnými daty. K dispozici jsem měla databázi s více než 32 tisíci dárci k roku 2007 (*obr. 5.1*), které poskytl ČNRDD. Data mi byla předána v databázovém softwaru Microsoft Access. Pro implementaci v MATLABu byla data převedena do tabulkového procesoru Microsoft Excel. Pro práci s Javou byla upravena tak, aby byl usnadněn vyhledávací proces.

Data byla poskytnuta registrem v podobě fenotypů. Pro určování kompatibility pacient-dárce je možné využít fenotypů, v praxi se však častěji využívá haplotypů. Proto je celá databáze dárců kostní dřeně pomocí nástroje MATLAB a Java prohledávána na haplotypové úrovni.

Obr. 5.1 – Přehled vstupu nových dárců do registrů ČR v průběhu let 2001 – 2011 [24]



5.1 Testování algoritmu v prostředí MATLAB

Cílem prototypu v prostředí MATLAB bylo vhodně naimplementovat pravděpodobnostní vzorce k výpočtu genetické diverzity. Algoritmus vyhledává konkrétní HLA geny, a dále pak počítá pravděpodobnosti za pomoci vzorců 4.1 a 4.8. Program umožňuje vyhledat četnosti daných alelických skupin, počítá jednoduchou a podmíněnou pravděpodobnost pro proměnný tvar vstupních parametrů a následně zobrazuje výstupy v grafické podobě.

Z důvodu citlivosti dat poskytnutých ČNRDD, která nemohou být dále šiřitelná, jsou grafy vykresleny nad testovacími daty. Bylo vygenerováno 1000 položek prezentujících dárce s jejich transplantačními HLA znaky pomocí MATLABu funkcí *rand()* s rovnoměrným rozložením.

MATLAB

Systém MATLAB, který vyvinula v roce 1984 firma The Mathworks, Inc. v USA, je výkonné prostředí určené pro vědecké a inženýrské výpočty a vizualizaci dat. Název MATLAB vznikl z anglického MATrix LABoratory. Program umožňuje manipulaci s maticemi, graficky zobrazovat funkce, možnost vytvářet vlastní uživatelská rozhraní a implementaci algoritmů.

5.1.1 Koncepce programu – MATLAB

Databáze

Pro práci s MATLABem bylo vhodné uložit databázi z Accessu do formátu Excel. Důvodem přeložení byla komplikovaná komunikace MATLABu s Accessem, která vyžadovala instalaci ODBC ovladače pro propojení. V případě, že bylo spojení prováděno Windows XP – 32 bitů a verzí MATLABu 2007b – 32 bitů, ovladač fungoval bez jakýchkoli potíží. Problémy nastaly po spuštění v systému Windows7 64 bitů, kdy nebylo možné uskutečnit spojení z důvodu mixování 32/64 bitů. Proto bylo vhodnější uložit databázi do formátu *.xls, aby bylo zamezeno zbytečným komplikacím, kdyby byl program spouštěn na jiném počítači s jinou systémovou verzí Windows.

Databáze obsahuje 32 tisíc dárců. Kde každá řádka reprezentuje HLA typizaci dárce a jednotlivé sloupce představují určitý HLA antigen. Ukázka dat je zobrazena v tabulce níže (tab. 5.1).

Tab. 5.1 – Příklad úseku databáze v Microsoft Excel

A1S	A2	A2S	B1	B1S	B2	B2S	DRB1*1	DRB1*2
	8		35		30		07:01	11:04
	6		16	39	21		13	19
	2		78		36		03	7
23	9	13	5		19		12:MT	03

Popis programu

Soubory pro práci s programem:

loadXLSFile.m – Funkce pro načtení vstupních dat z Excelu. Jako výstup jsou přiřazeny tři proměnné. V *m* je uložena matice v číselném formátu, *dict* jsou umístěny hlavičky sloupců (názvy HLA antigenů) v textovém formátu, *vektorMest* obsahuje zkratky všech dárcovských center databáze.

viceprvkovaPpst.m – Funkce pro výpočet podmíněné pravděpodobnosti dvou a více prvků. Do funkce je možné zadávat proměnný počet vstupních parametrů, kdy podmiňující prvek je zadán na posledním místě.

viceprvkovaPpstZadanehoSloupce.m – Funkce pro výpočet podmíněné pravděpodobnosti. Do funkce je možné zadávat proměnný počet vstupních parametrů, kdy se podmiňujícím prvkem stává celý sloupec. Díky tomu je výstup funkce počítán pro všechny možné varianty podmíněných prvků.

vyberCentra.m – Funkce pro reprezentaci vybraného dárcovského centra. Při zadávání je zvolen řetězec představující dárcovské centrum (Praha – ‘AA’, Plzeň – ‘PL’). Funkce je zadávána ve tvaru: `[mRedukovana] = vyberCentra(m, vektorMest, 'AA')`, kdy výstupním parametrem je soubor dat daného odběrového centra v podobě matice `[mRedukovana]`.

polohaMesta.m – Pomocná funkce pro hledání dárcovského centra v celkové matici, kdy vrací indexy řádků matice a je volána funkcí *vyberCentra.m*.

grafTick.m – Funkce pro vykreslení grafu

elementCount2.m – Výpočet jednotlivých četností pro zadané hodnoty.

kontrolaDat.m – Sloužící pro kontrolu sérologie s DNA typizací. U pacienta, který má vyplněnou DNA typizaci, a chybí sérologie, je vypsána -1.

Postup vyhledání

Po spuštění programu MATLAB je jako průvodce vytvořena funkce *moznosti.m*. Po jejím zavolání jsou ukázány všechny možnosti, jak s programem pracovat.

Nejprve je nutné vytvořit spojení s databází pomocí funkce *loadXLSFile* ve tvaru: `[m, dict, vektorMest] = loadXLSFile('TestovaciData.xlsx')`. Načtená data jsou uložena do tří zmíněných proměnných.

Další volbou může být vyhledání četnosti daného prvku pro určitý antigen nebo zobrazení všech prvků antigenu funkcemi: `c=elementCount2(m, dict, 9, 'DRB1*2')`, `[mc] = elementCounts(m, dict, 'DRB1*1')`. Následující možnosti je vypsání podmíněné pravděpodobnosti ve dvou variantách. První nám umožní podmíněnou pravděpodobnost počítat pro zadaný konkrétní prvek př: `ppstPoslednihoPrvku = viceprvkovaPpst(m, dict, 'DA1', 2, 'DB1', 7)`. Druhá varianta zadání je pro výpočet podmíněné pravděpodobnosti ne pro jeden prvek, ale pro celý sloupec: `podmPpstZadanehoSloupce(m, dict, 16, 'DRB1*1', 'DRB1*2')`.

V databázi je uložena informace, z kterého dárcovského centra se daný poskytovatel nachází. Proto je v programu vytvořena funkce, která zadané centrum separuje do nové matice: `[mRedukovana] = vyberCentra(m, vektorMest, mesto)`, se kterou je možné dále pracovat.

Výstupy

Výstupy jsou zobrazovány pomocí:

- a) konzole,
- b) grafu.

Příklad 5.1 – Zadání do funkce programu: Vypočtená četnost celého sloupce A1 ze souboru *TestovaciData.xlsx* (obr. 5.2).

Zadání:

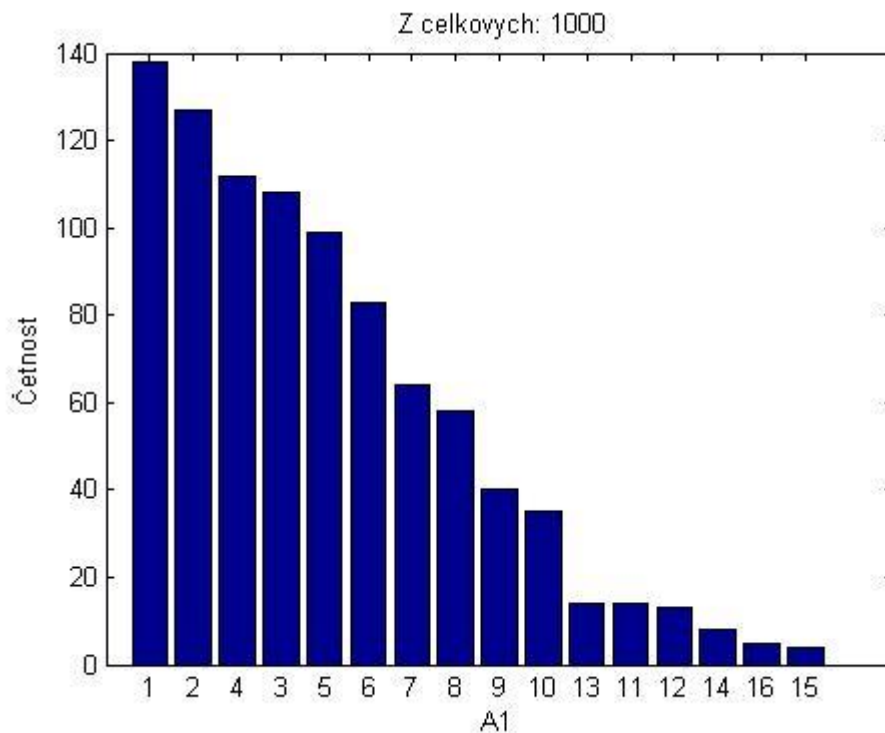
```
[m, dict, vektorMest] = loadXLSFile('TestovaciData.xlsx');  
[mc] = elementCounts(m, dict, 'A1')
```

Výstup:

Vypis cetnosti sloupce A1:

Hodnota 0 se vyskytuje 78x.
Hodnota 1 se vyskytuje 138x.
Hodnota 2 se vyskytuje 127x.
Hodnota 3 se vyskytuje 108x.
Hodnota 4 se vyskytuje 112x.
Hodnota 5 se vyskytuje 99x.
Hodnota 6 se vyskytuje 83x.
Hodnota 7 se vyskytuje 64x.
Hodnota 8 se vyskytuje 58x.
Hodnota 9 se vyskytuje 40x.
Hodnota 10 se vyskytuje 35x.
Hodnota 11 se vyskytuje 14x.
Hodnota 12 se vyskytuje 13x.
Hodnota 13 se vyskytuje 14x.
Hodnota 14 se vyskytuje 8x.
Hodnota 15 se vyskytuje 4x.
Hodnota 16 se vyskytuje 5x.

Obr. 5.2 – Četnost jednotlivých alelických skupin antigenu A1



Příklad 5.2 – Zadání práce s programem: Vypočtená četnost celého sloupce A1 ze souboru *TestovaciData.xlsx* z vybraného odběrového centra - pro Plzeň (obr. 5.3).

Zadání:

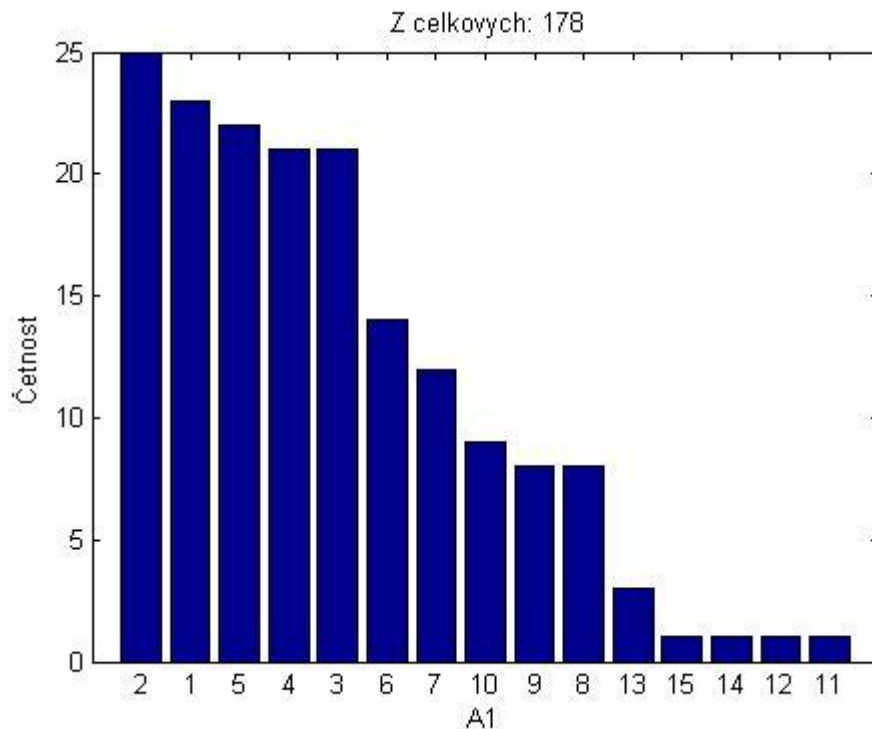
```
[m, dict, vektorMest] = loadXLSFile('TestovaciData.xlsx');  
[mRedukovana] = vyberCentra(m, vektorMest, 'AA');  
[mc] = elementCounts(mRedukovana, dict, 'A1')
```

Výstup:

Vypis cetnosti sloupce A1: PL

Hodnota 0 se vyskytuje 8x.
Hodnota 1 se vyskytuje 23x.
Hodnota 2 se vyskytuje 25x.
Hodnota 3 se vyskytuje 21x.
Hodnota 4 se vyskytuje 21x.
Hodnota 5 se vyskytuje 22x.
Hodnota 6 se vyskytuje 14x.
Hodnota 7 se vyskytuje 12x.
Hodnota 8 se vyskytuje 8x.
Hodnota 9 se vyskytuje 8x.
Hodnota 10 se vyskytuje 9x.
Hodnota 11 se vyskytuje 1x.
Hodnota 12 se vyskytuje 1x.
Hodnota 13 se vyskytuje 3x.
Hodnota 14 se vyskytuje 1x.
Hodnota 15 se vyskytuje 1x.

Obr. 5.3 – Četnost jednotlivých alelických skupin antigenu A1 pro centrum Plzeň



Příklad 5.3 – Zadání práce s programem: Výpočet podmíněné pravděpodobnosti pro prvek B1*19, kde podmiňujícími prvky jsou antigeny A1*1 a DRB1*1*5. A výpočet podmíněná pravděpodobnost pro všechny prvky celého sloupce B1 se stejnými podmiňujícími prvky. Příklad počítá s daty ze souboru *TestovaciData.xlsx*. (obr. 5.4).

Zadání:

```
[m, dict, vektorMest] = loadXLSFile('TestovaciData.xlsx');
ppstPoslednihoPrvku= viceprvkovaPpst(m,dict,'A1',1,'DRB1*1',5,'B1',19)
viceprvkovaPpstZadanehoSloupce(m, dict, 'A1', 1, 'DRB1*1', 5, 'B1')
```

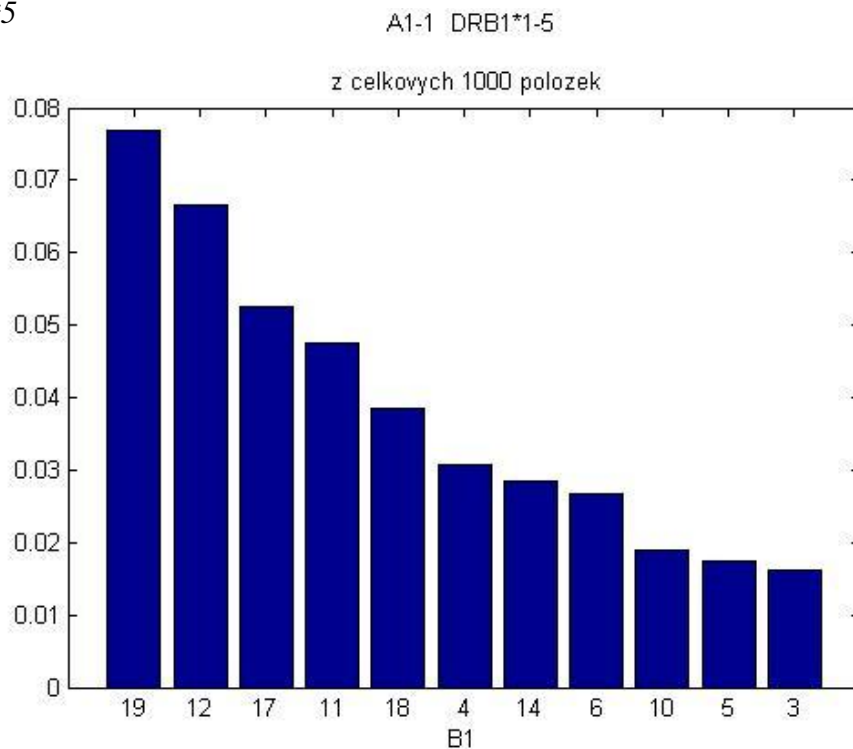
Výstup:

```
ppstPoslednihoPrvku =
    0.0769

viceprvkovaPpstZadanehoSloupce =

    3.0000    0.0161
    4.0000    0.0308
    5.0000    0.0175
    6.0000    0.0267
   10.0000    0.0189
   11.0000    0.0476
   12.0000    0.0667
   14.0000    0.0286
   17.0000    0.0526
   18.0000    0.0385
   19.0000    0.0769
```

Obr. 5.4 – Podmíněná pravděpodobnost sloupce B1 s podmiňujícími prvky A1*1 a DRB1*1*5

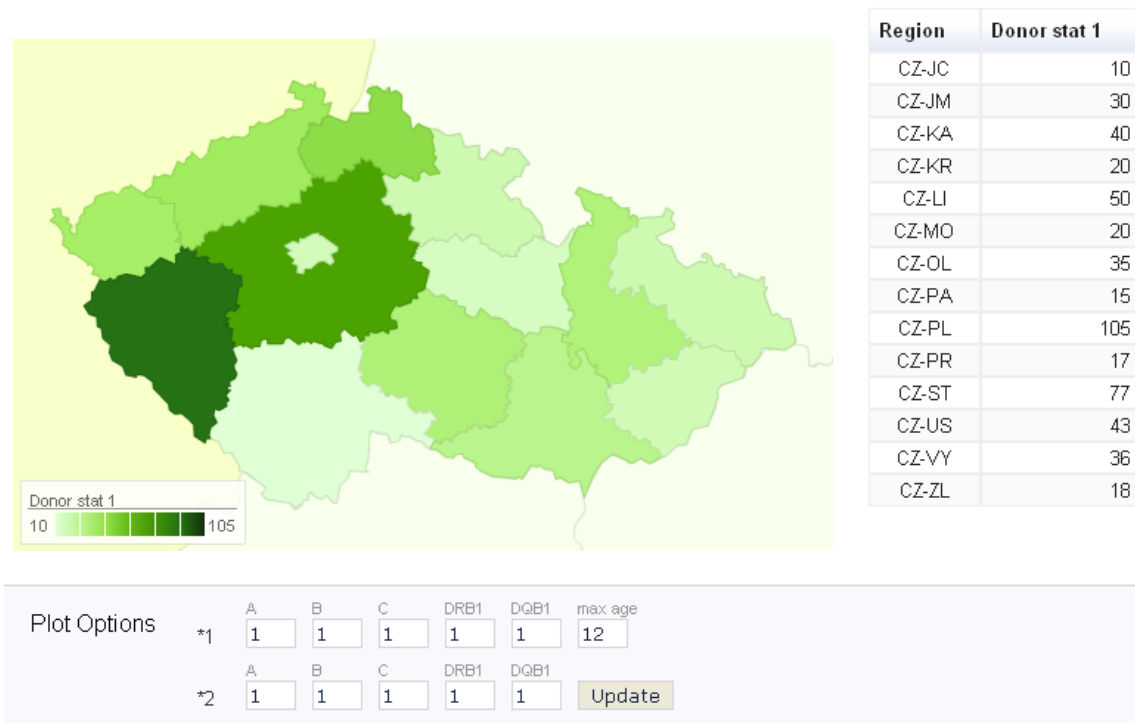


5.2 Interaktivní implementace v prostředí Java

Program v prostředí Java má za cíl prohledávat databázi dárců, nalézt v ní zadanou formu dotazu v podobě antigenu a specifické alely a vrátit výsledné pravděpodobnosti, které jsou řízeny vzorci z podkapitoly 4.4. Pro snazší implementaci byla databáze uložena lokálně v databázovém systému MySQL, nyní je externě uložena na webové síti Google App Engine s níž vytvořený program komunikuje.

Program je určen k vizualizaci dat pro webovou aplikaci na stránkách http://ccy.zcu.cz/registr/HLA_stats.html (obr. 5.5). Aplikace znázorňuje výskyt HLA antigenů, které jsou pro vykreslení zadávány v nastavení „Plot Options“, pravá tabulka obrázku značí četnosti HLA v jednotlivých krajích. Aplikace bude dále komunikovat prostřednictvím souboru *Index.jsp*, který bude čerpat z databáze dárců na úložišti Google App Engine a díky němuž budou v databázi dárců ČNRDD vyhledávány geny, antigeny i konkrétní alely, počítány četnosti a podmíněné pravděpodobnosti. Všechny zmíněné úkony mohou být prováděny nad celou databází nebo nad vyseparovaným dárcovským centrem.

Obr. 5.5 – Geografická statistika HLA



Java

Programovací jazyk Java vznikl roku 1995. Po jeho vpuštění na trh zaznamenal velký ohlas díky jeho vlastnostem jako je podpora objektově orientované programování. Dále je Java multiplatformní, tzn. spustitelná systémy Windows, Linux, Unix, atd. [25]

Google App Engine

Google App Engine (GAP) vznikl roku 2008. Jedná se o prostředí umožňující tvorbu a hostování webových aplikací. GAP původně k tvorbě aplikací využíval programovacího jazyka Python. Na žádost veřejnosti se společnost Google o rok později rozhodla GAP rozšířit o podporu pro jazyk Java. [26]

Výhody aplikace spočívají v pověsti společnosti Google, která je zárukou dlouhodobého vývoje GAP. Dále je umožněna správa dat zcela zdarma. „*Tyto aplikace běží na serverové architektuře Googlu, není tak nutné se starat o správu systému, či investovat spousty prostředků do nákupu hardwaru. Google se postará o veškerou údržbu, takže se lze soustředit přímo na aplikaci samotnou. Bezpečnost, soukromí, ochrana dat a aplikace samotné je s App Engine na nejvyšší úrovni.*“ [27]

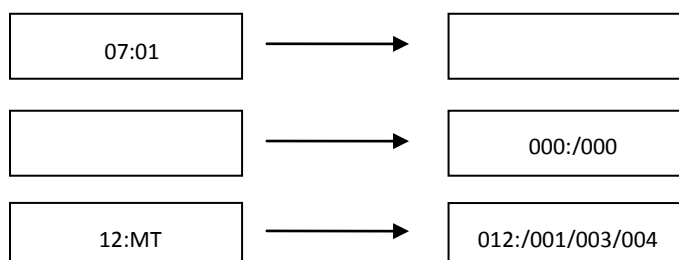
5.2.1 Koncepce programu – Java

Databáze

Zkušební Java verze pracovala s lokální databází, jejím úložištěm se stal databázový systém MySQL. Systém byl vybrán z důvodu volně šiřitelného softwaru a jeho verze jsou multiplatformní.

Data z Microsoft Access byla uložena do formátu *.csv a následně importována do MySQL pomocí PHP skriptu. Kde byly číselné hodnoty upraveny do formátu pro snazší vyhledávání a jednodušší skládání dotazů v databázi MySQL. Prázdné kolonky byly následně přepsány do tvaru nul a NMDP kódy byly nahrazeny jim příslušející množinou alel (*obr. 5.6*).

Obr. 5.6 – Ukázka převodu dat z Access do MySQL za pomoci skriptu PHP



Databáze je nyní uložena na Google App Engine se stejným formátem dat jako v MySQL (tab. 5.2).

Tab. 5.2 – Příklad úseku databáze v MySQL

A1S	A2	B1	B1S	B2	B2S	DRB1*1	DRB1*2
000:/000	008:/000	035:/000	000:/000	030:/000	000:/000	007:/001	011:/004
000:/000	006:/000	016:/000	039:/000	021:/000	000:/000	013:/000	019:/000
000:/000	002:/000	078:/000	000:/000	036:/000	000:/000	003:/000	007:/000
023:/000	009:/000	005:/000	000:/000	019:/000	000:/000	012:/001/003/004	003:/000

Popis programu

Probability.java – Třída pro výpočet jednoduchých a podmíněných pravděpodobností. Pracuje se vzorci 4.9 – 4.15 a vznikají v ní SQL dotazy.

SearchItem.java – Vytváří jednotlivé položky ze zadaných řetězců.

Record.java – Přepisování položek z MySQL databáze na číselné hodnoty, které jsou dále využívány při práci ve třídě *Probability*.

Index.jsp – Soubor pro zobrazování spočtené hodnoty třídou *Probability* na webové stránce.

Postup vyhledání pravděpodobnosti pacient-dárce

Pro počítání podmíněné pravděpodobnosti je nejprve sestaven požadovaný řetězec ve formátu "DRB1-1;016:016|DB1;015:015" přiřazený do metody *vypocti()* v souboru *Index.jsp*. Řetězec je složen z genu DRB1*1, jeho antigenu (16) a alely (16) podmíněný genem DB*1, antigenem (15) a alelou (15). Dále je možné do souboru *Index.jsp* vložit řetězec s požadovanou zkratkou dárcovského centra. Řetězec např. pro Plzeň je vložen do metody *setDonorCenter("PL")*. Výpočet je pak počítán nad daty konkrétního centra.

Zadaný řetězec "DRB1-1;016:016|DB1;015:015", který je argumentem čili zadáním výpočtu je předán do třídy *Probability*, kde je rozdělen podle vztahu 4.12 do podoby:

$$P(DRB1*1 = 016:016|DB1 = 015:015) = \frac{P(DRB1*1=016:016,DB1=015:015)}{P_c(DB1=015:015)}$$

V této podobě je dále převeden do třídy *SearchItem*, která jej rozloží na jednotlivé řetězové položky ve tvaru DRB1-1;016:/016 obsahující například hledaný antigen či konkrétní alelu. Volba tvaru vzniklého řetězce je odvozena podle prvků databáze pro snazší hledání.

Nově vytvořené řetězce putují opět do třídy *Probability*, která z nich poskládá SQL dotaz, který je zaslán databázi. V případě zadání dárcovského centra je SQL dotaz uzpůsoben tak, aby data byla nejprve vyseparována podle daného centra.

Databáze následně vrací výsledky nalezených shod, které jsou třídou *Record* přepsány na číselné hodnoty, s kterými třída *Probability* dále pracuje. *Probability* zpracovává číselné hodnoty podle podkapitoly 4.4. Důležitou částí třídy je vytváření kartézského součinu, kde je využita metoda *hash()*. Do metody vstupují prvky, ona sama si pak dokáže vyloučit opakující se prvky a následně vrací pouze ty z nich, které jsou unikátní. Poté je již dosazeno do vzorů, jejichž výstupním parametrem je nalezená hodnota.

Zdrojový kód *Index.jsp* je přeložen na servlet a následně kompilovaný. Výsledkem je servlet generující HTML, které zobrazuje hodnotu výpočtu.

Výstupy

Výstupy jsou zobrazovány pomocí:

- a) skládání SQL dotazů,
- b) hodnoty pravděpodobnostních dotazů,
- c) geografická vizualizace dat.

a) Skládání SQL dotazů

Ukázka SQL dotazu vytvořeného metodou *createSQLWhere()* k uvedenému řetězci DRB1-1;016:016|DB1;015:015":

```
SELECT * FROM `darci-working-copy` WHERE 1 = 1 AND `darci-working-copy`.`DRB1-1` like ('016%') AND `darci-working-copy`.`DRB1-1` like ('%/016/%') AND `darci-working-copy`.`DB1` like ('015%') AND `darci-working-copy`.`DB1` like ('%/015/%')
```

Při skládání SQL dotazů v programu Java je použita ze třídy *Probability()* metoda *createSQLWhere()* využívající položky ze seznamu *List<SearchItem>*.

```
/*
 * Metoda pro vytvoreni SQL dotazu, položky jsou získávány ze seznamu
 * SearchItem.
 * String sql vrací SQL dotaz v podobě řetězce.
 */
public String createSQLWhere(List<SearchItem> search) {
    String sql = "SELECT * FROM `" + tableName + "` WHERE 1 = 1";

    for(SearchItem si : search) {
        if(si.hasHLAProtein()) {
            sql += " AND `" + tableName + "`.`" + si.column + "`
like ('" + si.alleleGroup + "%') ";

            sql += " AND `" + tableName + "`.`" + si.column + "`
like ('%/" + si.specificHLAProtein + "/%') ";
        }
        else {
            sql += " AND `" + tableName + "`.`" + si.column + "`
like ('" + si.alleleGroup + "%') ";
        }
    }

    if(isDonorCenter()) {
        sql += " AND `" + tableName + "`.`kod` = '" + donorCenter + "' ";
    }

    return sql;
}
```

b) Hodnoty pravděpodobnostních dotazů

Příklad výpočtu podmíněné pravděpodobnosti programem Java pro zadaný příklad.

$$P(DRB1*1 = 016:016|DB1 = 015:015) = \mathbf{0.0952381}$$

c) Geografická vizualizace dat

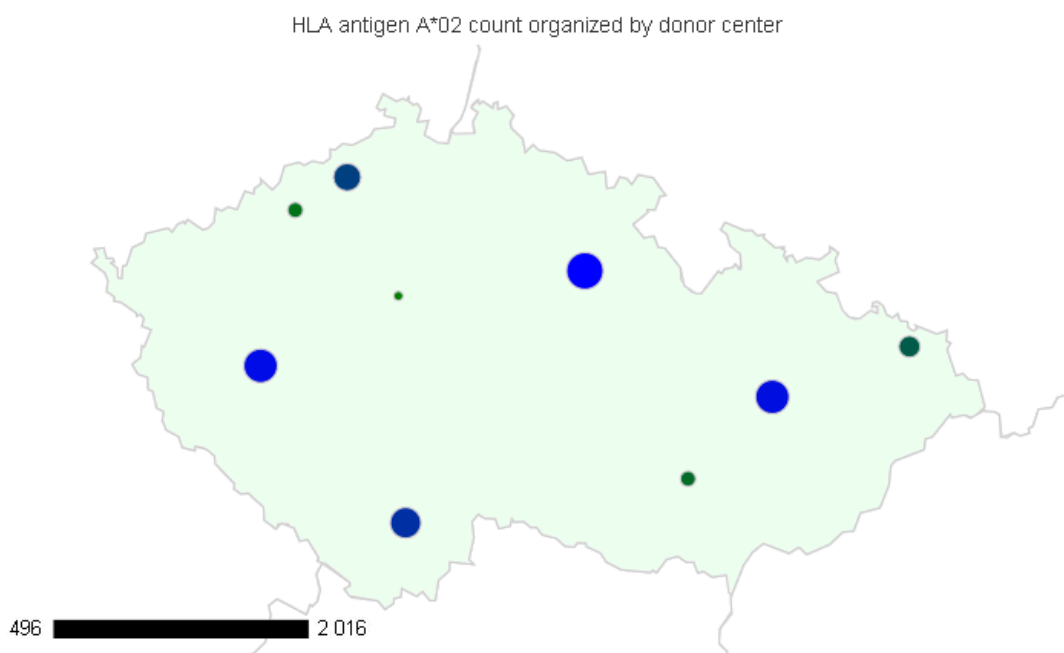
Program je možné využít pro geografickou vizualizaci dat. V níže uvedeném příkladu 5.4 jsou zaznamenány četnosti dárcovských center v databázi ČNRDD pro antigen A*02 v České republice.

Příklad 5.4 Četnost antigenu A*02 pro jednotlivá dárcovská centra jsou zaznamenána v tabulce (tab. 5.3). Četnosti jsou zobrazeny v obrázku (Obr. 5.7), který znázorňuje geografickou statistiku antigenu A*02 v České republice. Body na mapě prezentují odběrová centra, velikost těchto bodů se odvíjí od výskytu antigenu v lokálních databázích.

Tab. 5.3 – Četnosti antigenu A*02 pro dárcovská centra v ČR

Odběrové centrum	Četnost prvku A*02 ve vybraném centru
AA – Praha	496
BM – Brno	728
CB – České Budějovice	1466
HK – Hradec Králové	2016
MO – Most	645
OL – Olomouc	1834
OS – Ostrava	924
PL – Plzeň	1873
UL – Ústí nad Labem	1257
	11239

Obr. 5.7 – Geografická statistika – četnosti antigenu A*02 v ČR



6 Závěr

Bakalářská práce s názvem „*Podpora tvorby aplikace pro určení kompatibility pacient-dárce*“, je zaměřena na tvorbu programů umožňujících prohledávání databáze Českého národního registru dárců kostní dřeně (ČNRDD) a následné zpracování dat pomocí pravděpodobnostních modelů.

Hlavním důvodem zvoleného tématu bylo stát se součástí velkého projektu a práce v týmu, který vytvářel vyhledávací algoritmus a pro něj potřebné pravděpodobnostní modely pro zefektivnění vyhledávacího procesu pro ČNRDD. Práce v týmu pro mě byla neocenitelnou zkušeností nejen z hlediska praxe. Dalším důvodem byla praktická využitelnost tohoto projektu. Velkou výhodou byla možnost práce nad skutečnými daty. K dispozici jsem měla databázi s více než 32 tisíci dárci k roku 2007, kde byly zaznamenány informace o jejich HLA znacích v podobě haplotypů. Data nám poskytli ČNRDD ze všech dárcovských center kostní dřeně v České republice.

Bakalářská práce je členěna do několika kapitol. Úvod práce je zaměřen na transplantaci kostní dřeně. V této kapitole jsem se věnovala charakteristice kostní dřeně, kde se v lidském těle nalézá a blíže se specifikuje její nejvýznamnější funkce – hematopoéza. Další kapitola uvádí do problematiky HLA antigenů, kde jsou vymezeny základní pojmy čerpající z oboru imunologie. Ve čtvrté kapitole jsem se zabývala úvodem do pravděpodobností, které jsou nedílnou součástí interaktivní implementace. Jsou zde popisovány pravděpodobnostní modely pracující s NMDP kódy. Poslední kapitola líčí programovou část prototypu v prostředí MATLAB a implementaci v prostředí Java, kde jsou znázorněny výstupy zmíněných programů.

Práce je komponována tak, aby byla využitelná pro vizualizaci statistických dat, která jsou potřebná k vyhledávacímu procesu. Díky sestavení pravděpodobnostních modelů a vizualizaci dat je možné zefektivnit nábor nových dárců do databáze, a tím tak ulehčit lékařům vyhledávací proces. Podle prototypu v MATALBu již byly provedeny statistiky pro rok 2012.

Prvním cílem této práce bylo studium informací týkajících se problematiky transplantace kostní dřeně spojenou s důležitostí HLA antigenů. Dozvěděla jsme se, že úspěšnost transplantace je významně ovlivněna výběrem vhodného dárce. Je tedy nutné, aby se pacientovi tkáňové znaky s dárcovými do jisté míry shodovaly. V současnosti je požadavek na míru shody 10/10 neboli v pěti HLA genech (HLA -A, -B, -C, -DRB1 a

-DQB1). Nejmenší možná shoda představuje 6/10 v genech (HLA -A, -B, -DRB1), ale zde už bohužel pro pacienta vzniká smrtelné riziko odvržení štěpu.

Druhým cílem bylo získání teoretických poznatků z oboru pravděpodobnosti. V těchto kapitolách jsem se snažila zaměřit na pravděpodobností vztahy, které měly co nejvíce souviset s programovou částí.

Poslední cíl byl dělen na dva samostatné úseky. První z nich měl za úkol vytvoření prototypu vyhledávacího programu v prostředí MATLAB, který měl nalézt zadané hodnoty v podobě konkrétních HLA genů v databázi. Po lokalizaci vstupních parametrů je možné zjistit četnosti alelických skupin, jejich vzájemnou podmíněnost a následné výstupy zobrazit graficky. Výhodou je zadávání proměnného počtu vstupních parametrů a specifikace dárcovského centra.

Ve druhém úseku posledního cíle vznikla interaktivní implementace programu v prostředí Java. Program je konstruován tak, aby prohledával databázi dárců, našel v ní zadanou formu dotazu ve tvaru antigenu a specifické alely a vrátil výsledné hodnoty v podobě pravděpodobností podle vzorců podkapitoly 4.4. Program dále komunikuje s databází, která je externě uložena na webové síti Google App Engine.

Spolupráce na tak obsáhlém projektu mě velmi obohatila. Přinesla mi nejen nové poznatky týkající se prohloubení znalostí pravděpodobnosti a programové implementace, ale také jsem se dozvěděla o závažnosti tématu „transplantace kostní dřeně“, která bohužel není v dnešní době příliš známá v povědomí lidí.

Výkladový slovník pojmů

Alela – jedna ze dvou nebo více forem určitého genu, který zodpovídá za určitou fenotypovou vlastnost a je přítomný v určitém lokusu na určitém chromosomu.

Alogenní příjemce – geneticky podmíněný rozdíl mezi antigeny dárce transplantátu a jeho příjemcem

Amplifikace – rozmnožení

Ankylozující spondylartritida – Bechtěrevova choroba – skupinu zánětlivých revmatických onemocnění, postihuje spojení na páteři

BMDW – Bone Marrow Donors Worldwide

ČNRDD – Český národní registr dárců dřeně

Fenotyp – je soubor všech pozorovatelných vlastností a znaků živého organismu. Představuje výsledek spolupůsobení genotypu a prostředí, čili to, jak organismus v daném znaku (znacích) skutečně vypadá.

Gen – základní jednotka dědičnosti lokalizovaná v určitém místě chromosomu.

Genom – soubor genů v buňce, souhrn všech genů v daném organismu, lidský genom obsahuje 20 – 25.000 genů.

Genotyp – genetická výbava určitého organismu, která představuje soubor alel specificky uspořádaných v jeho genomu.

Goodpastureův syndrom – vzácné autoimunitní onemocnění, charakterizované přítomností autoprotilátek proti bazální membráně ledvinových glomerulů a alveolárních kapilár. Projevuje se rychle postupujícím zánětem parenchymu ledvin (glomerulonefritidou) a krvácením do plic.

Haplotyp – je kombinace alel dvou nebo více těsně vázaných lokusů na jednotlivém chromozomu. Každý jedinec dostane dva soubory alel, od každého rodiče jeden. Každý soubor alel se označuje jako haplotyp. Čili jedinec dědí jeden haplotyp od otce, druhý od matky.

Chromosom – vláknitá struktura buněčného jádra, v níž je v podobě DNA obsažena dědičná informace.

Kostní dřev – morek – rosolovitá tkáň uvnitř kosti

Leukémie – leukémie je onemocnění krvetvorby, jenž se projevuje zmnožením nezralých bílých krvinek.

Lokus – specifické místo na chromosomu v molekule DNA, kde se nachází určitý gen. V jednom lokusu může být jen jedna z možných alel daného genu.

Myelodysplastický syndrom – onemocněním krvetvorné tkáně. Porucha krvetvorby vede k nedostatku krvinek a k projevům anémie.

Pemphigus – autoimunitní puchýřnaté onemocnění postihující kůži i sliznice po celém těle.

Revmatoidní artritida – je zánětlivé autoimunitní onemocnění, které se projevuje především zánětem a bolestivostí kloubů, postihuje však i ostatní orgány.

Roztroušená skleróza – chronické autoimunitní onemocnění, při kterém lidský imunitní systém napadá centrální nervovou soustavu (mozek a míchu).

Syngenní vztah – genetická shoda transplantátu a příjemce v histokompatibilitních antigenech.

Uveitida – zánětlivé oční onemocnění uvey (duhovky, cévnatky a řasnatého tělíska)

WHO – World Health Organization – Světová zdravotnická organizace

Použité zdroje

1. Hematopoetická kmenová buňka. [Online] [Citace: 29.11.2011]
http://cs.wikipedia.org/wiki/Hematopoetická_kmenová_buňka.
2. Wellthy Choices Network. [Online] [Citace: 12.11.2011]
<http://www.wellthychoices.net/2010/10/27/adya-clarity-baths-article-by-matt-monarch/>.
3. ŠVOJGROVÁ, M., KOZA, V., HAMPLOVÁ, A. Transplantace kostní dřeně: Průvodce Vaší léčbou. 1. vyd. Plzeň : F. S. Publishing, 2006. ISBN 80-903560-2-8.
4. ČIHÁK, R. Anatomie I. Praha : Grada, 2001. Sv. 2. ISBN 978-80-247-3817-8.
5. PENKA, M., TESAŠOVÁ, E., et al. Hematologie a transfuzní lékařství. 1. vyd. Praha : Grada, 2011. ISBN-13: 978-80-247-3459-0.
6. SYSLOVÁ, J. iDNES.cz. Co můžete darovat ze sebe: krev, vajíčka i orgány. [Online] [Citace: 11.4.2012] http://ona.idnes.cz/co-muzete-darovat-ze-sebe-0hj-/zdravi.aspx?c=A120410_131956_zdravi_pet.
7. KOPECKÝ, J. Imunologie. [Online] [Citace: 6.2.2012]
<http://rum.prf.jcu.cz/public/kopecky/>.
8. KREJSEK, J., KOPECKÝ, O. Klinická imunologie. Hradec Králové : Nucleus HK, 2004. ISBN 80-86225-50-X.
9. HOŘEJŠÍ, V., BARTUŇKOVÁ, J. Základy imunologie. 4. vyd. Praha : Triton, 2009. ISBN: 978-80-7387-280-9.
10. LEBECK, L. K., GARAVOY, M. R. Histocompatibility Testing and Organ Sharing. [Online] [Citace: 6.8.2011] <http://www.kidneyatlas.org/book5/adk5-08.pdf>.
11. FERENČÍK, M. Imunitní systém: informace pro každého. 1. vyd. Praha : Grada. ISBN: 80-247-1196-6.
12. PETŘEK, M. Medicina.cz. [Online] [Citace: 5.2.2012]
http://www.medicina.cz/odborne/clanek.dss?s_id=4730.

13. **KÁBRT, L.** HLA genetická příbuznost Čechů s ostatními národy. *Praha, 2009. Diplomová práce. České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra kybernetiky.*
14. **ČAPKOVÁ FRIDRICHOVÁ, R., A KOL.** Kurz základních metod molekulární biologie. [Online] [Citace: 2.4.2012]
<http://www.bc.cas.cz/doc/ekotech/study/ZMMB4.pdf>.
15. *Wikipedia.* Primer (molecular biology). [Online] [Citace: 12.4.2012]
[http://en.wikipedia.org/wiki/Primer_\(molecular_biology\)](http://en.wikipedia.org/wiki/Primer_(molecular_biology)).
16. **MRÁZEK, F., AMBRŮZOVÁ, Z., PETŘEK, M.** OHD - Olomoucké hematologické dny. [Online] [Citace: 21.3.2012]
<http://www.olhemdny.cz/abstrakta/byauthor.php?year%5BBrocnik%5D=2010&abstracts%5Bprijmeni%5D=Mr%Elzek&abstracts%5Bjmeno%5D=Franti%9Aek>.
17. **ROBINSON, J., et al.** The IMGT/HLA database. *Nucleic Acids Research.* [Online] [Citace: 25.2.2012] <http://hla.alleles.org/>.
18. **JINDRA, P.** Kostní dřeň. HLA data dárců - pravidla reportování, záznam do databáze ČNRDD. [Online] [Citace: 21.3.2012]
<http://www.kostnidren.cz/registr/odbornici/specificka-cast.php>.
19. **TĚHNÍK, J.** Geographical Analysis of Databases of Stem Cell Donor Registries. *Praha, 2011. Diplomová práce. České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra kybernetiky.*
20. **HOUDOVÁ, L., GEORGIEV, D.** Návrh metod pro vyhledávání dárce (ČNRDD). *Plzeň : Západočeská univerzita v Plzni, 2011.*
21. **SHOENFELD, Y., FUČÍKOVÁ, T., BARTŮŇKOVÁ, J.** Autoimunita: vnitřní nepřítel. 1. vyd. *Praha : Grada, 2007. ISBN:978-80-247-2044-9.*
22. **REIF, J., KOBEDA, Z.** Úvod do pravděpodobnosti a spolehlivosti. 2. vyd. *Plzeň : Západočeská univerzita v Plzni, 2006. ISBN 80-7043-333-7.*
23. **MAREK, P.** Učební materiály: Základní pojmy, definice pravděpodobnosti. *Plzeň, 2009.*

24. *Kostní dřevěň*. Vývoj databáze dárců ČNRDD. [Online] [Citace: 6.3.2012]
<http://www.kostnidren.cz/registr/vysledky/vyvoj.php>.
25. *Java*. [Online] [Citace: 16.4.2012] <http://www.java.com>.
26. *Google Developers*. Google App Engine. [Online] [Citace: 18.3.2012]
<http://code.google.com/intl/cs-CZ/appengine/>.
27. *Komtesa*. [Online] [Citace: 18.3.2012] [http://www.komtesa.com/google-apps-
implementace-a-vyvoj-pomoci-app-engine/](http://www.komtesa.com/google-apps-implementace-a-vyvoj-pomoci-app-engine/).
28. *University of Maryland Medical Center*, *Bone marrow aspiration*. [Online] [Citace: 12.11.2011]
http://www.umm.edu/patiented/articles/bone_marrow_aspiration_000171.htm.

Přílohy

Příloha 2.1 – Transplantace kostní dřeně z kosti pánevní [28]

