

**Prirodoslovno-matemati ki fakultet**

**Biološki odsjek**

**Zavod za molekularnu biologiju**

**SEMINARSKI RAD**

**Evolucija kompleksa gena tkivne podudarnosti: model nastanka i nestanka gena**

**Evolution of the major histocompatibility complex: birth-and-death evolution model**

**Mila Baši**

**Mentor:**

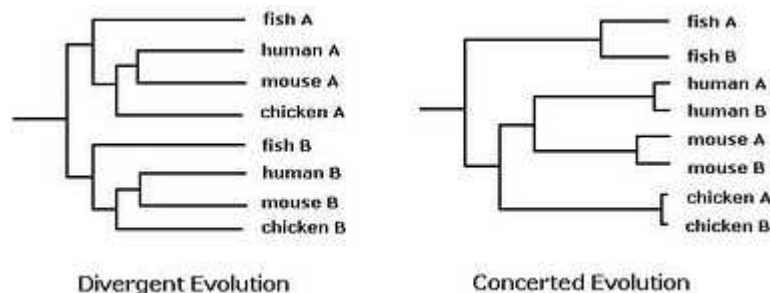
**doc. dr. sc. Ana Galov**

**Zagreb, srpanj 2011.**

## 1. UVOD

Multigenske porodice su one u kojima članovi potječu od zajedničkog pretka, a imaju istu ili sličnu funkciju i DNA slijed. Do 1970. godine smatralo se da članovi multigenske porodice evoluiraju na način da se geni dupliciraju i zatim evoluiraju neovisno jedan o drugom i tako stječu nove funkcije. Svaki od dupliciranih gena ima zasebnu filogeniju, ali su one slične (slike 1 i 2). Taj model je nastao na temelju proučavanja  $\alpha$ ,  $\beta$ , i  $\gamma$  lanaca hemoglobina i mioglobina. Nazvan je modelom divergentne evolucije.

Oko 1970. godine, proučavaju i ponavljaju se gene koje kodiraju rRNA roda *Xenopus* otkriveno je da su nukleotidni slijedovi intergenskih regija slični unutar jedne vrste nego između dvije srodne vrste. To nije bilo moguće objasniti pomoću dotadašnjeg modela. Uveden je novi model evolucije multigenskih porodica koji uzrok sličnosti nalazi u usklađenoj evoluciji gena (Liao 1999). Zato geni nemaju dvije nezavisne filogenije i slični su unutar vrste, u usporedbi s članovima iste porodice gena u drugim vrstama (slike 1 i 2). Taj model nazvan je modelom usklađene evolucije (engl. *concerted evolution*) (Nei i sur. 2005). Primjer za evoluciju gena prema tom modelu je evolucija rRNA ili snRNA gena (Nei i sur. 2005).

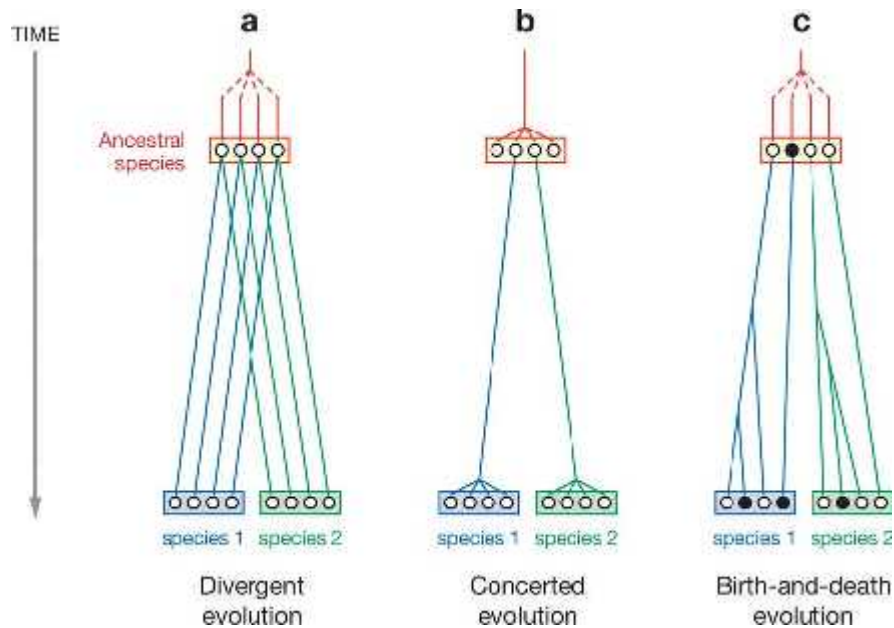


**Slika 1.** Prikaz filogenije prema divergentnom modelu i prema modelu usklađene evolucije

(preuzeto s <http://sandwalk.blogspot.com/2007/01/evolution-of-gene-families.html>)

Nakon uvođenja tog modela, kako je otkriveno sve više multigenskih porodica, ija je evolucija pratila upravo taj model, primijenjen je i na proučavanje evolucije gena koji

kodiraju MHC molekule. Me utim, kod evolucije tih gena, taj model nije generalno primjenjiv. Javila se potreba za uvo enjem novog modela koji je kombinacija modela divergentne evolucije i uskla ene evolucije. Nazvan je modelom nastanka i nestanka gena (engl. *birth-and-death*) (slika 2). Novi geni posljedica su duplikacija i kao takvi nalikuju jedan na drugog, kao i u modelu uskla ene evolucije. Neki od tih gena se 'sa uvaju', odnosno nastane novi gen, a neki se deletiraju ili inaktiviraju delecijским mutacijama, odnosno gen 'nestane'. Stoga e broj gena u porodici ostati isti. Delecije i inaktivacije gena objašnjavaju zašto se neki geni prisutni u pretku ne pojavljuju u vrstama koje su evoluirale iz tog pretka, a neki od 'sa uvanih' geni su mnogo sli niji nego što se o ekuje prema modelu uskla ene evolucije (Nei i sur. 1997). I danas postoje polemike oko toga koji model je ispravan - uskla ene evolucije ili nastanka i nestanka gena.



**Slika 2.** Tri različita modela evolucije gena multigenijskih porodica, bijeli krugovi predstavljaju funkcionalne gene, a crni pseudogene (Nei i sur. 2005).

## 2. KOMPLEKS GENA TKIVNE PODUDARNOSTI KOD ČOVJEKA- HLA (engl. *human leukocyte antigen*)

### 2.1. Uloga MHC molekula

Osnovna komponenta humoralne imunosti su protutijela koja pružaju zaštitu preko izvanstaničnih molekula. Međutim, što je s antigenima koji se nalaze unutar stanica? Za takvu zaštitu odgovorna je stanicama posredovana imunost. Ista je osnovna komponenta limfociti T koji preko svojih receptora (TCR, engl. *T cell receptors*) prepoznaju peptidne fragmente antigena. Ti peptidni fragmenti antigena vezani su preko posebnih antigen predodajnih molekula, nazvanih MHC molekule, koje su izražene na površini antigen predodajne stanice (APC, engl. *antigen presenting cells*). Limfociti T, odnosno TCR, prepoznaju te peptidne fragmente samo ako su vezani na MHC molekule na površini APC i ako dođe do bliske međustanične interakcije (Male i sur. 2003).

### 2.2. Podjela

Postoje dvije vrste MHC molekula, to su MHC molekule skupine I (skraćeno MHC-I) i MHC molekule skupine II (skraćeno MHC-II) (Male i sur. 2003). MHC molekule su antigen predodne molekule koje vežu kratke peptidne fragmente antigena. Peptidni fragmenti koje vežu MHC-I nastali su djelovanjem proteaza u citosolu. Prijenos fragmenata iz citosola u lumen endoplazmatske mrežice vrše prijenosni proteini. U lumenu može se uz pomoć nekoliko proteina formirati MHC-I-antigeni kompleks. Kompleks može u slučaju prisutnosti antigena biti prenesen na površinu stanične membrane Golgijevim sklopom.

MHC-II, s druge strane, prezentiraju peptidne fragmente nastale djelovanjem vezikularnih proteaza. Spajanjem vezikula koje nose MHC-II i endosoma, koji sadrže fragmentirane peptide antigena, formira se MHC-II-antigeni kompleks. Spajanje se također događa u lumenu endoplazmatske mrežice (Andreiss i sur. 2004).

Duljine fragmenata antigena koje predodaju kreću se od osam do deset aminokiselina za MHC-I i 13-24 aminokiselina za MHC-II. Vezno mjesto MHC-II molekula otvorenije je

od veznog mjesta MHC-I molekula što je uzrok sposobnosti vezanja duljih fragmenata (Male i sur. 2003).

### 2.3. Građa

MHC-I građene su od teškog glikoziliranog lanca (45 kDa) nekovalentno povezanog s  $\beta_2$ -mikrotubulinom ( $\beta_2m$ ), također polipeptidnim lancem (12 kDa). Teški glikozilirani lanac sastoji se od 3 ekstracelularne domene ( $D_1$ ,  $D_2$ ,  $D_3$ , svaka po ~90 aminokiselina), transmembranske domene (~25 aminokiselina) i citoplazmatskog repa (30-40 aminokiselina) (slika 3). Tri izvanstanične domene su te koje su glikozilirane. Treća domena,  $D_3$ , nekovalentno povezuje teški lanac s  $\beta_2m$ .  $D_1$  i  $D_2$  domene teškog lanca tvore antigen-vezuće mjesto; platformu 8 antiparalelnih lanaca koje podupiru dva heliksa. Aminokiseline koje tvore antigen-vezuće mjesto uzrok su visoke polimorfnosti (visoke strukturne raznolikosti) i velikih razlika u obliku utora na koji se antigen veže (Male i sur. 2003).

MHC-II građene su od teškog lanca (30-34 aminokiselina) i lakog (26-29 aminokiselina) lanca koji su nekovalentno vezani. Razlika u molekularnoj masi proizlazi iz različite glikoziliranosti.  $D_1$ ,  $D_2$  i  $D_3$  su N-glikozilirani, dok  $D_4$  domena nije. Izvanstanična regija sastoji se od dvije domene ( $D_1$  i  $D_2$  kod lanca, te  $D_1$  i  $D_2$  kod lanca) na koje se nadovezuju transmembranska (~30 aminokiselina) i citoplazmatska regija (10-15 aminokiselina), slično kao i kod molekula skupine I (Male i sur. 2003) (slika 3).

Strukturno gledajući građu MHC-I i MHC-II, uočava se sličnost u građi  $D_3$  domene MHC-I i  $D_2$  i  $D_4$  domena MHC-II koje sve skupa strukturno nalikuju konstantnoj regiji imunoglobulina. Gledajući cjelokupnu trodimenzionalnu strukturu, MHC-I i MHC-II su dosta slični (Male i sur. 2003).

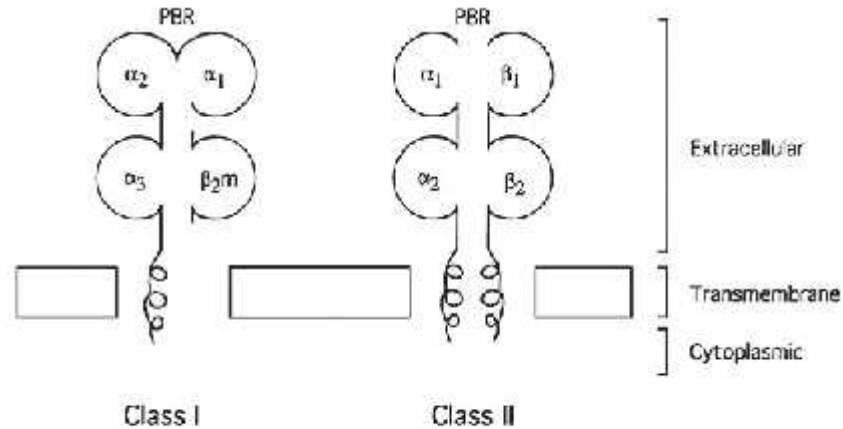


Figure 7. Schematic presentation of the structure of MHC class I and class II molecules. PBR = peptide-binding region. (Reprinted, with permission, from the Annual Review of Genetics, Vol. 24 ©1998 by Annual Reviews, www.annualreviews.org).

**Slika 3.** Prikaz gra e MHC molekula (preuzeto s [http://www.funpecrp.com.br/gmr/year2006/vol4-5/gmr0218\\_full\\_text.htm](http://www.funpecrp.com.br/gmr/year2006/vol4-5/gmr0218_full_text.htm))

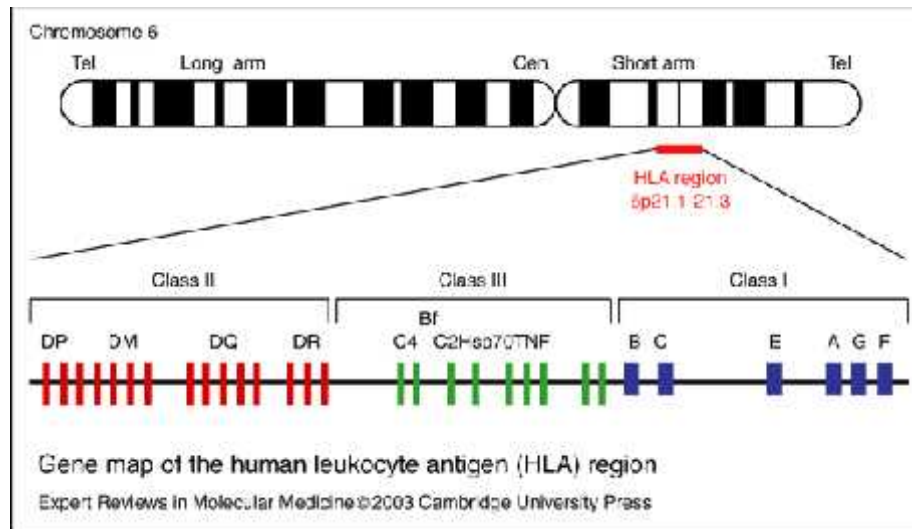
#### 2.4. Genski lokusi

HLA zauzima dio kromosoma 6, segment od ~4,000 kilobaza (kb) koje predstavljaju 84 gena, od kojih ~15 kodira MHC-I, ~19 kodira MHC-II (Klein i sur. 1993) (slika 4).

Geni koji kodiraju MHC-I skupljeni su u klaster i zauzimaju regiju dugu ~2,000 kb, dok geni koji kodiraju MHC-II, koji su tako er skupljeni u klaster, zauzimaju nešto kra u regiju dugu ~1,000 kb. Postoji i regija od ~1,000 kb koja se nalazi izme u regija koje kodiraju MHC-I i MHC-II i zbog toga se naziva MHC-III (Klein i sur. 1993).

Glavna tri genska lokusa MHC- I su I lokusi: HLA-A, HLA-B i HLA-C, pri emu svaki kodira teški lanac klasi ne MHC-I (Male i sur. 2003), dok je  $\beta_2$ -mikrotubulin ( $\beta_2m$ ) kodiran genom koji se nalazi izvan MHC regije (Nei i sur. 1997).

Geni koji kodiraju MHC-II (najmanje 6 lanaca i 10 lanaca) nalaze se u regiji HLA-D, a glavna tri genska lokusa su slijede a: DP, DR i DQ, koji kodiraju barem jedan i barem jedan lanac (Male i sur. 2003).



**Slika 4.** Raspored HLA genskih lokusa (preuzeto s

[http://journals.cambridge.org/fulltext\\_content/ERM/ERM5\\_07/S1462399403005957sup002.htm](http://journals.cambridge.org/fulltext_content/ERM/ERM5_07/S1462399403005957sup002.htm))

## 2.5. Polimorfizam

Jedno od glavnih obilježja MHC molekula jest izrazito visok stupanj polimorfizma. Unutar odre ene MHC molekule, bilo skupine I ili II, polimorfizam je ograni en na odre ene regije. Kod MHC-I to su tri regije  $\alpha_1$  i  $\alpha_2$  domene.  $\alpha_2$  domena je, ini se, više konzervirana. Kod MHC-II polimorfizam ovisi o podregiji i polipeptidnom lancu. Ve ina polimorfnih aminokiselina, kod obje skupine molekula, ograni ene su na vrh molekula oko antigen-vežu eg mjesta, tako da polimorfizam utje e na sposobnost MHC molekula da vežu razli ite peptidne fragmente razli itih antigena i proširuje spektar antigenskih fragmenta koje mogu prezentirati (Male i sur. 2003).

### 3. EVOLUCIJA GENA MULTIGENSKIH PORODICA

#### 3.1. Usklađena evolucija

Iznenađujuće veliku frakciju genoma zauzimaju repetitivne sekvence. One su većinom nekodirajuće regije, a manjim dijelom kodirajuće. Neovisno o svojstvu kodiranja pokazuju zajedničko, neotkriveno svojstvo da evoluiraju usklađeno, stoga nije otkriveno da je i evolucija MHC molekula proučavana na tom modelu evolucije (Liao 1999). Pretpostavlja se da su uzrok evolucije spontane mutacije koje se dogode u jednom genu, ali se prenesu na sve gene porodice tako da geni i dalje evoluiraju usklađeno. Glavni mehanizmi usklađene evolucije su nejednoliki krosingoveri ili genska konverzija. Genska konverzija je vrsta nerekipročne rekombinacije u kojoj se segment gena kopira, 'obradi' u drugi, a molekularni mehanizam do danas nije razjašnjen. I nejednoliki krosingover i genska konverzija imaju slični efekti. Jedina je razlika što nejednoliki krosingover ima za posljedicu povećanje ili smanjenje broja gena, dok genska konverzija ne. Brojni su problemi vezani uz ovaj model evolucije kod gena MHC molekula. Mala je vjerojatnost da se genska konverzija događa između dva diferencirana lokusa u to određeno vrijeme u evoluciji. (Nei i sur. 2005). Također, kada bi geni evoluirali modelom usklađene evolucije, očekivali bi da aleli svakog polimorfnog lokusa ne tvore nužno monofiletski klaster u filogenetičkom stablu, što isto nije slučaj (Nei i sur. 2005). Nije iznenađujuće, stoga, što se pojavio drugi model koji objašnjava evoluciju gena ove multigenske porodice. Iako nisu sporne pojave krosingovera i genske konverzije, ono što je sporno jest njihova učestalost i utjecaj na evoluciju gena koji kodiraju MHC molekule. Ipak, postoje i danas autori koji zagovaraju ovaj model evolucije (Nei i sur. 2005).

#### 3.2. Model nastanka i nestanka gena

Jedna od prvih zamjerki teoriji usklađene evolucije bila je što neki geni unutar multigenske porodice nisu nužno slični drugim genima unutar iste vrste, nego nalikuju genima drugih vrsta. Neki autori zato smatraju da je evolucija multigenske porodice MHC molekula tekla prema modelu nastanka i nestanka gena prema kojem su novi nastanu duplikacijama, a dio njih 'nestane', bilo inaktivacijom ili delecijom. (Nei i sur. 1997). 'Nestanak' gena može biti nasumičan, a može



biti i posljedica selekcije. Taj model objašnjava zašto nekih gena nema u nekih vrsta koje potječu od istog zajedničkog pretka, a oni prisutni, 'sačuvani', su mnogo sličniji nego što se očekuje prema modelu usklađene evolucije. Sve u estalija otkrivanja sljedova DNA u različitim vrstama omogućila su brojne filogenetičke analize. One su, primjerice, vršene nad ljudskim genomom, ali i onim gorila, impanzi itd. Filogenetičke analize pokazale su da je evolutivski obrazac različit od onog koji bi bio očekivan modelom usklađene evolucije; različite porodice ili rodovi sisavaca nemaju 'prave' ortologne gene. Ortologni geni su homologni geni u različitim vrstama koji su donekle slični jer potječu iz posljednjeg zajedničkog pretka (Nei i sur. 2005). Nakon razdvajanja dvije nove vrste iz posljednjeg zajedničkog pretka, imale su različitu evoluciju, što je uzrok različitosti u slijedu DNA. Jedino logično objašnjenje za tu pojavu jest nastanak odnosno nestanak gena koji su odlike ovog modela. On ujedno objašnjava i visoki postotak inaktiviranih gena, pseudogena, u itavoj MHC regiji (Nei i sur. 2005). Za visok stupanj polimorfnosti ponuđeno je drugo objašnjenje – hipoteza heterozigotne prednosti. Selekcija vrši se nad antigen-vežućim mjestu MHC molekula, što je alternativa teoriji da je za to zaslužna genska konverzija. Na elo hipoteze heterozigotne prednosti jest da su heterozigoti višeg fitnesa od homozigota. U ovom slučaju, heterozigoti će biti sposobni obraniti se od dvije različite vrste patogena, a homozigoti samo od jedne. Zato će heterozigoti imati viši stupanj preživljenja - viši fitnes (Nei i sur. 2005). Ipak, ta saznanja ne isključuju mogućnost uključivosti genske konverzije ili nejednolikog krosingovera u evoluciji gena MHC molekula, ali čini se da je taj doprinos zanemarivo mali. Nisu geni koji kodiraju MHC molekule jedini koji pokazuju ovaj obrazac evolucije. U tu skupinu ubrajamo i ostale imunoglobulinske gene i njima srodne gene, gene olfaktornih receptora, gene feromonskih receptora i gene zaslužne za sintezu feromona (Nei i sur. 2005).

#### 4. ZAKLJUČAK

Studije o evoluciji gena koji kodiraju MHC molekule uglavnom se svode na razmatranje dva modela evolucije: model usklađene evolucije i model nastanka i nestanka gena. Iako postoje dokazi koji idu u prilog oba modela, model nastanka i nestanka gena bolje odgovara obrascu evolucije gena koji kodiraju MHC molekule. Iako je mehanizam modela usklađene evolucije, nejednoliki krosingover i genska konverzija, dokazan, ne postoje dokazi da su ti događaji značajni u evoluciji gena koji kodiraju MHC molekule kao takvih, odnosno, njihov doprinos je zanemarivo mali. Istraživanja su pokazala da je model nastanka i nestanka gena vjerojatniji, tj. bolje odgovara obrascu evolucije, što je dokazano sve češćim filogenetičkim analizama koje se vrše nad brojnim vrstama.

Ipak, i danas postoje autori koji zagovaraju teoriju usklađene evolucije. Za razrješenje ove polemike važno je napraviti jednoznačnu studiju evolucijskih modela gena imunološkog sustava. Srećom, količina podataka o DNA sekvencama svakodnevno raste, uglavnom zbog projekata sekvenciranja genoma koje se provodi nad različitim organizmima. Sigurno je da će s novim saznanjima, s više sekvenciranih regija iz različitih vrsta biti lakše dokazati kako je tekla evolucija gena koji kodiraju MHC molekule, ali i ostalih gena multigenskih porodica u genomu.

## 5. LITERATURA

1. Igor Andreiss, Drago Batini , Filip ulo, Danka Gr evi , Matko Maruši , Milan Taradi, Dora Višnji , 2004. Imunologija. Šesto izdanje. Medicinska naklada, Zagreb.
2. Etienne Joly, 2006. Various hypotheses on MHC evolution suggested vy the concerted evolution of CD94L and MHC class Ia molecules, *Biology direct*. **1**:3.
3. Jan Klein, Colm O'hUigin, Felipe Figueroa, Werner E. Mayer, and Dagmar Klein, 1993. Different Modes od Mhc Evolution in Primates. *Mol. Biol. Evol.* **12**, 48-59.
4. Daiqing Liao, 1999. Concerted Evolution: Molecular Mechanism and Biological Implications. *Am. J. Hum. Genet.* **64**, 24-30.
5. David Male, Jonathan Brostoff, David Roth, and Ivan Roitt 2003. Immunology. 7th edition. Alden Group, Oxford.
6. Masatoshi Nei and Alejandro P. Rooney, 2005. Concerted and Birth-and Death Evolution of Multigene Families. *Annu Rev Genet.* **39**, 121-152.
7. Masatoshi Nei, Xun Gu, and Tatyana Sitnikova, 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc. Natl. Acad. Sci. USA* **94**, 7799-7806.
8. [http://journals.cambridge.org/fulltext\\_content/ERM/ERM5\\_07/S1462399403005957sup002.htm](http://journals.cambridge.org/fulltext_content/ERM/ERM5_07/S1462399403005957sup002.htm)
9. [http://www.funpecrp.com.br/gmr/year2006/vol4-5/gmr0218\\_full\\_text.htm](http://www.funpecrp.com.br/gmr/year2006/vol4-5/gmr0218_full_text.htm)
10. <http://sandwalk.blogspot.com/2007/01/evolution-of-gene-families.html>

## 6. SAŽETAK

MHC molekule važna su komponenta imunosti. Njihova uloga jest predočavanje fragmenata antigena limfocitima T. Geni koji kodiraju MHC molekule članovi su multigenske porodice. To znači da imaju zajedničko porijeklo i sličnog su DNA slijeda. MHC regija podijeljena je na tri regije, regiju koja kodira MHC molekule skupine I, regiju koja kodira MHC molekule skupine II, a između se nalazi tzv. MHC regija III. Evolucija gena multigenskih porodica, kao što su porodica rRNA gena, histona, imunoglobulinska porodica, porodica MHC molekula, proučavana je već dugi niz godina, pa su se tako javili različiti modeli koji objašnjavaju njihovu evoluciju. Nekada se vrijeme smatralo da su geni koji kodiraju MHC molekule evoluirali prema modelu usklađene evolucije. Prema tom modelu uzrok evolucije su spontane mutacije koje se kopiraju i šire po porodici gena nizom uzastopnih nejednolikih krosingovera ili genskom konverzijom. Iako prisutnost tih pojava nije upitna, nema dokaza da su ti događaji dovoljno učestali i važni da su uzrok evolucije cijele porodice. Geni koji kodiraju MHC molekule, međutim, evoluirali su prema modelu nastanka i nestanka gena gdje geni nastaju duplikacijama. Neki se geni 'sačuvavaju', a neki se deletiraju ili inaktiviraju, odnosno nestanu. Filogenetičke analize, koje se danas vrše zahvaljujući sve većem broju sekvenciranih slijedova DNA u različitim vrstama, potvrđuju upravo taj obrazac evolucije. Daljnjim sekvenciranjem gena koje kodiraju MHC molekule u što više različitih vrsta može se dobiti potpunija slika evolucije tih gena, ali i gena multigenskih porodica općenito.

## 7. SUMMARY

The MHC molecules are an important part of immunity. Their function is presenting peptide fragments of antigens to T-cells. Genes coding for the MHC are a part of the multigene family. This means that they are of the same origin and have a similar DNA sequence. The MHC region is divided into three regions, one that codes for MHC-I, second that codes for MHC-II and the third one, MHC-III, which is positioned between the first two regions. Evolution of multigene families, like the rRNA genes, histone genes, immunoglobulin gene family, MHC gene family, has been studied for a long time, consequently different models explaining their evolution appeared. For a certain time, it was believed that the evolution of the MHC followed a pattern called concerted evolution. According to this model, the cause of evolution are spontaneous mutations arising in a repeat unit and are then spread rapidly throughout the multigene family by unequal crossing-over or gene conversion. Although these occurrences are not questionable, there is no evidence that they are occurring often enough or even important enough to have a role in gene evolution of the entire multigene family. The MHC genes apparently evolved following birth-and-death evolution, where new genes are created by gene duplication. While some genes are maintained for a long time, others are deleted or become nonfunctional by deleterious mutations. Phylogenetic analyses, which are run thanks to an increasing number of identified DNA sequences in different species, confirm this model of evolution. Continued application of phylogenetic analysis produces not only a broader picture of evolution of these genes, but also evolution of multigene families in general.