

UUS VERERÕHU REGULATSIOONI MÕJUTAV
GEEN *NAT8* (N-atsetüültransferaas 8)

PEETER JUHANSON

MAGISTRITÖÖ BIOMEDITSIINIS

Juhendajad:

Erakorraline professor Maris Laan, Ph.D.

Dotsent Mai Rosenberg, M.D., Sc.D.

Tartu Ülikool
2007

SISUKORD

LÜHENDID JA MÕISTED	4
SISSEJUHATUS	6
1. KIRJANDUSE ÜLEVAADE	7
1.1 Komplekshaigused	7
1.2 Arteriaalne hüpertensioon	8
1.2.1 Essentsiaalne hüpertensioon kui komplekshaigus	9
1.2.2 Teadaolevad essentsiaalse hüpertensiooni monogeensed vormid	11
1.2.3 Neerud ja nende roll hüpertensiooni kujunemisel	11
1.2.4 Essentsiaalse hüpertensiooni kui komplekshaiguse päriliku komponendi kaardistamine	14
1.2.4.1 Essentsiaalse hüpertensiooni uuringud aheldusanalüüsil ..	14
1.2.4.2. Essentsiaalse hüpertensiooni assotsiatsiooniuuringud	18
1.2.4.3. Essentsiaalse hüpertensiooni uuringud loomudelitel abil	22
2. TÖÖ EESMÄRGID	24
3. TÖÖ SKEEM	25
4. MATERJAL JA METOODIKA	26
4.1 <i>NAT8</i> geeni amplifikatsioon inimese genoomist	26
4.2 PCR produktide analüüs ja töötlus	29
4.3. <i>NAT8</i> geeni resekveneerimine	29
4.4. DNA järjestuste esmane analüüs ja visualiseerimine	30
4.5 Statistiline analüüs	31
4.6 <i>In silico</i> analüüs	33
4.7 Assotsiatsiooniuuringus osalevad indiviidid	33
5. TULEMUSED	36
5.1 Uute hüpertensiooni kandidaatgeenide tuvastamine ahelduse analüüsil positiivseteks osutunud piirkondadest	36
5.2 Kromosoomi 2p piirkonna iseloomustus	36
5.3 <i>NAT8</i> iseloomustus ja konserveeritus	37
5.4 <i>NAT8</i> genoomse piirkonna detailsem iseloomustus ja spetsiifiline amplifikatsioon	39
5.5 <i>NAT8</i> geeni resekveneerimine ja polümorfne varieeruvus	41
5.6 <i>NAT8</i> alleelse assotsiatsiooni muster	43
5.7 <i>NAT8</i> nukleotiidne varieeruvus	44

5.8 Kahekümne üheksa (29) hüpertensiooni kandidaatgeeni promootorala resekveneerimine ja nukleotiidiline mitmekesisus*	45
5.9 <i>NAT8</i> SNP-de assotsiatsioonianalüüs vererõhu, vere seerumi	46
biomarkerite ja GF-iga	46
4.10 <i>NAT8</i> promootorala SNP alleelide potentsiaalne mõju geeniekspressioonile	48
6. ARUTELU	50
6.1 Andmebaasides olemasolevate SNP-de usaldusväärsus	50
6.2 <i>NAT8</i> kui võimalik hüpertensiooni kandidaatgeen	51
6.3 <i>NAT8</i> potentsiaalse promootorpiirkonna SNP-d	53
KOKKUVÕTE	56
SUMMARY	57
TÄNUAVALDUSED	58
KIRJANDUSE VIITED	59
LISA 1. <i>NAT8</i> ja <i>NAT8B</i> järjestuste joondus	64

LÜHENDID JA MÕISTED

A	adeniin
Ah	aminohape
bp	aluspaar (<i>base pair</i>)
C	tsütosiin
cM	centi Morgan
dATP	desoksüadenosiintrifosfaat
dCTP	desoksütsütidiintrifosfaat
dGTP	desoksüguanosiintrifosfaat
dTTP	desoksütümidintrifosfaat
EH	essentsiaalne hüpertensioon
Exo I	eksonukleas I
G	guaniin
GF	glomerulaar-filtratsioon (<i>glomerular filtration rate GFR</i>)
<i>In silico</i>	arvutisimulatsioonil läbiviimine (<i>In or by means of a computer simulation</i>)
kb	kiloaluspaar
KD	kilodalton
LD	ahelduse tasakaalustamatus (<i>linkage disequilibrium</i>)
LOD skoor	statistikaline hinnang, mis väljendab kahe lookuse aheldatuse tõenäosust (<i>log of the odds</i>)
<i>NAT8</i>	N-atsetüültransferaas 8
<i>NAT8B</i>	N-atsetüültransferaas 8 duplikaatgeen
OR	haigusalleeliga indiviidide arvu suhe haigusalleelita indiviidide arvu (<i>odd ratio</i>)
p	kromosoomi lühike õlg
PCR	polümeraasi ahelreaktsioon (<i>polymerase chain reaction</i>)
PSV	paraloogsed järjestuse variandid (<i>paralogous sequence variants</i>)
q	kromosoomi pikk õlg

rcf	relatiivne tsentrifugaaljõud (<i>Relative Centrifugal Force</i>)
rpm	pööret minutis (<i>rotation per minute</i>)
SNP	ühenukleotiidne polümorfism (<i>single nucleotide polymorphism</i>)
sAP	kreveti aluseline fosfataas (<i>Shrimp Alkaline Phosphatase</i>)
taq	<i>Thermus aquaticus</i> 'e DNA polümeraas
u	ühik (<i>unit</i>)
URL	dokumentide või muude ressurside aadress internetis (<i>uniform resource locator</i>)
UTR	mittetransleeritav piirkond (<i>untranslated region</i>)

SISSEJUHATUS

Tänu teaduse ja tehnoloogia kiirele arengule suurenevad üha enam teadmised inimeste haiguste geneetilistest põhjustest, mis võimaldavad mitmeid haiguseid juba varakult ennetada või efektiivsemalt ravida. Suureks abiks on olnud inimese genoomi täielik sekveneerimine, millest pärinevate andmete töötlusel ja analüüsil koostöös bioinformaatikutega on loodud põhjalikud andmebaasid, võimaldades kiiret ja paindlikku geneetilise informatsiooni kättesaamist.

Valdav enamus inimestel esinevatest haigustest on kompleks- ehk multifaktoriaalsed haigused, st. haiguse kujunemises osaleb rohkem kui üks geen ning olulised on ka keskkonnafaktorid ja elustiil. Antud tööd võib vaadelda osana projektist, mille eesmärgiks on uurida pärilikult kõrge vererõhu geneetilist tausta. Südameveresoonehaigused, s.h arteriaalne hüpertensioon, ja selle tüsistused on arenenud riikides, s.h Eestis, sagedaimaks surma põhjuseks, mistõttu on nende geneetilise tausta tundmaõppimine äärmiselt vajalik nii haiguse ennetamise seisukohalt kui ka hilisemate tüsistuste, näiteks neerukahjustuste vältimiseks.

Käesoleva magistritöö kirjanduse ülevaate esimeses osas selgitatakse arteriaalse hüpertensiooni kui komplekshaiguse olemust ja selle peamisi tekkemehhanisme neerude osalusel. Eraldi peatutakse ka juba tänaseks teadaolevatel geneetilistel põhjustel, mis osalevad hüpertensiooni või vererõhu regulatsioonis. Teises osas antakse ülevaade põhilistest multifaktoriaalsete haiguste geneetilise kaardistamise meetoditest ning peatutakse seni teostatud essentsiaalse hüpertensiooni geneetilistel uuringutel.

Töö esimese praktilise osa käigus resekveneeriti 42-l Eesti populatsiooni indiviidil uus hüpertensiooni kandidaatgeen *NAT8* ning kirjeldati geeni polümorfset varieeruvust lähtudes Eesti populatsioonist. Teise praktilise osana viidi läbi *NAT8* geeni potentsiaalse promootorala SNP-dega assotsiatsiooniuuringud, kus vaadeldi promootorala SNP-de seost vererõhuga, vere seerumi biomarkerite ja neeru funktsiooni iseloomustava glomerulaarfiltratsiooniga.

1. KIRJANDUSE ÜLEVAADE

1.1 Komplekshaigused

Viimaste aastate jooksul on hakatud üha intensiivsemalt uurima geenide ja erinevate keskkonnafaktorite koosmõju, mis on paljude sagedasti esinevate ning samas ühiskonnale suhteliselt tuttavate haiguste tekke põhjusteks või eeldusteks. Sellest tulenevalt võimaldavad edusammud komplekshaiguste geneetilise tausta uurimisel koostöös teiste teadusharudega paljusid haigusi ennetada või efektiivsemalt ravida.

Aastakümneid on edukalt tegeletud monogeensete ehk mendeliaarsete haiguste uurimise ning kaardistamisega. Erinevalt komplekshaigustest on Mendeli seaduste kohaselt päranduvad haigused (näiteks tsüstiline fibroos, fenüülketonuuria) tavaliselt põhjustatud mutatsiooni(de)st mingi konkreetse geeni kodeerivas alas ja reeglina on välja kujunenud kindel genotüübi-fenotüübi vaheline seos. Kuigi sellised haigused, võrreldes näiteks kõrgvererõhutõvega, hõlmavad inimpopulatsioonist suhteliselt väikest osa, on nad suure tähtsusega vaadates konkreetset perekonda ja selle liikmeid. Komplekshaigused ehk multifaktoriaalsed haigused nagu näiteks kõrgvererõhutõbi, Parkinsoni tõbi, diabeet, vähktõbi, jne., ei pärandu Mendeli seaduste kohaselt. Tegemist on haigustega, mille väljakujunemisel osalevad mitmed geenid ja nende erivormid koos keskkonnafaktoritega (s.h. mitmete haiguste puhul ka konkreetse indiviidi elustiil) (Weiss and Terwilliger 2000). Klassikaliselt on multifaktoriaalseid haiguseid iseloomustatud järgmiselt (Lander and Schork 1994):

1. haiguse osaline penetrantsus
2. oligo- või polügeenne pärilikkus
3. geneetiline heterogeensus
4. haigust soodustavate alleelide sage esinemine populatsioonis
5. keskkonnamõjude oluline roll

Kõik need on uurimist raskendavad tegurid ning samuti ka põhjused, miks tulemused komplekshaiguste kaardistamises on märksa tagasihoidlikumad kui monogeensete haiguste puhul.

1.2 Arteriaalne hüpertensioon

Hüpertensioonina (haiguslikult kõrge vererõhuna) diagnoositakse inimorganismi seisundit, kus korduvatel mõõtmistel organismi puhkeolekus süstoolse ja diastoolse vererõhu väärtused on üle normi ehk siis vastavalt >140 ja/või >90 mmHg. Optimaalseks süstoolse ja diastoolse vererõhu väärtusteks peetakse vastavalt <120 / <80 mmHg (2003), (tabel 1). Kõrgenenud süstoolset vererõhku aga omakorda peetakse üheks tõsiseks riskifaktoriks südame ja veresoonkonna haiguste, insuldi, silma võrkkesta kahjustuste ning neerupuudulikkuse väljakujunemisel (Vikrant *et al.* 2001). Konkreetse indiviidi summaarne risk tuleneb siiski kõigi riskifaktorite ühisest koosmõjust ning võib olla sama kõrge vaatamata vererõhuväärtustele. Näiteks kuuluvad samasse riskigrupi inimesed, kellel ühel on kõrgenenud normaalne vererõhk koos kolme kaasuva riskifaktoriga, organkahjustusega või diabeediga ning teisel hüpertensiooni III klass ilma muude riskifaktoriteta (2003).

Tabel 1. Hüpertensiooni klassifikatsioon vererõhu väärtuste alusel (ESH 2003)

Kategooria	Süstoolne (mmHg)	Diastoolne (mmHg)
Optimaalne	<120	<80
Normaalne	120-129	80-84
Kõrgenenud normaalne	130-139	85-89
Hüpertensioon klass I (kerge)	140-159	90-99
Hüpertensioon klass II (mõõdukas)	160-179	100-109
Hüpertensioon klass III (raske)	≥ 180	≥ 110

Vastavalt etioloogiale saab hüpertensiooni klassifitseerida järgnevalt (Kaplan 1993):

- a) **Essentsiaalne hüpertensioon (EH)**. Hüpertensioon, mida ei põhjusta mõni teine haigus ega vererõhku tõstev tegur. Arvesse tuleb võtta paljusid etioloogilisi faktoreid nagu näiteks liigne Na⁺ peetus neerudes, liigne keedusoola tarbimine, geneetilised riskifaktorid, jne. EH moodustab ligikaudu 95% kõigist arteriaalse hüpertensiooni juhtudest täiskasvanutel vanuses 18-65 a. EH-le on sageli iseloomulik perekondlik anamnees, aeglaselt progresseeruv kulg, asümptomaatiline kulg kuni organkahjustuste kujunemiseni (10-20 a. pärast) ning avaldumine vanuses 30-50 eluaastat.
- b) **Sekundaarne hüpertensioon**. Diagnoositakse juhul, kui kõrgeenenud vererõhk on põhjustatud mõnest teisest haigusest või teadaolevast vererõhku tõstvast tegurist (näiteks ravimi kõrvaltoime). Arteriaalse hüpertensiooni kõikidest juhtudest moodustavad sekundaarse hüpertensiooni vormid järgmise osa: parenhümatooossed neeruhaigused 3-4%, renovaskulaarne hüpertensioon 0,5-1%, endokriinsed haigused 0,1-0,3% ning muud põhjused 0,1-0,3%.

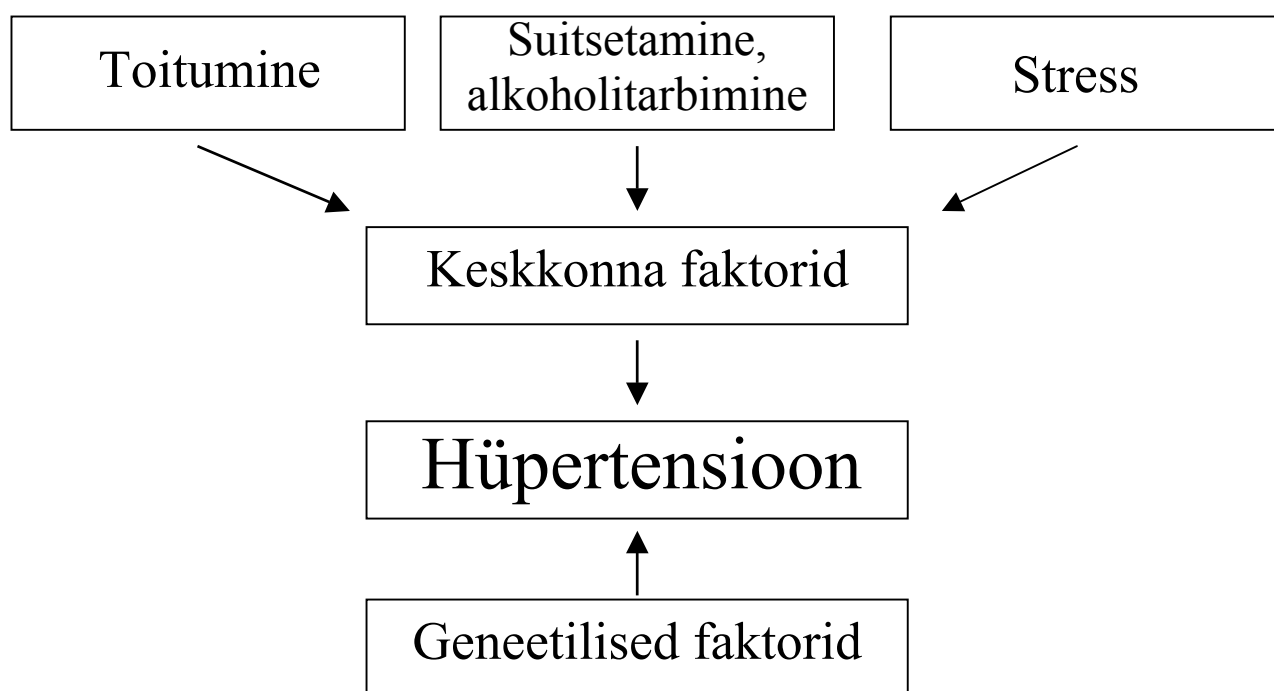
1.2.1 Essentsiaalne hüpertensioon kui komplekshaigus

Tänapäeval käsitletakse hüpertensiooni peamiselt kui kompleks- ehk multifaktoriaalset haigust. Püsivalt kõrgeenenud vererõhu tingib geneetiline eelsoodumus, inimese elustiil ja keskkonna faktorid. Viimaste all võib tähtsamate näidetena tuua urbaniseerumise, vähese füüsilise aktiivsuse, ülekaalulisuse, liigse loomsete rasvade ja soola tarbimise, suitsetamise ning liigse alkoholi tarbimise (joonis 1) (Snieder *et al.* 2003). Hüpoteesi kinnitab elustiili muutuse (soola tarbimise piiramine, kehakaalu alandamine) soodne mõju vererõhu alanemise näol. Teada on, et umbes pooled hüpertensiooni haigetest reageerivad positiivselt igapäevase soolakoguse vähendamisele (Poch *et al.* 2001). Perekonna uuringutest on selgunud,

et 30-60% indiviididevahelisest vererõhu varieeruvusest on geneetilise taustaga (Snieder *et al.* 2003).

Essentsiaalne hüpertensioon on kaasaja ühiskonnas kõige sagedamini esinev haigus, hõlmates enda alla 26,4% täiskasvanud elanikkonnast. Hinnanguline hüpertensioonihaigete arv 2000.a oli 972 miljonit ning aastaks 2025 prognoositakse selle tõusu 1,56 miljardini.

Eestis on südame- ja veresoonkonnahaigused, mille põhiliseks riskiteguriks peetakse kõrgeenenud vererõhku, peamisteks surma- ja haigestumispõhjusteks, moodustades viimase 15 aasta vältel üle 50% kõigist surmapõhjustest. Eesti Haigekassale tähendasid vereringeelundite haigused 2003.a. kogu ravikuludest 21,3% ehk 791,5 miljonit krooni (<http://www.haigekassa.ee/raviasutusele/tervisedendus/prioriteetid/> 20.04.2005.a).



Joonis 1. Hüpertensioon kui komplekshaigus.

1.2.2 Teadaolevad essentsiaalse hüpertensiooni monogeensed vormid

Vaatamata üldlevinud lähenemisele – hüpertensioon kui komplekshaigus, on kirjeldatud mitmeid mendeliaarselt päranduvaid hüpertensiooni vorme. Molekulaargeneetiliste uuringutega on tänaseks identifitseeritud mutatsioonid kaheksas geenis, mis osalevad vererõhu regulatsioonis (tabel 2) (Gong and Hubner 2006). Kõik need geenid osalevad ühtses metaboolses võrgustikus organismi soola – vedeliku tasakaalu kujundamisel, pakkudes nii üht võimalikku pidepunkti hüpertensiooni kui laiema fenotüübi tundmaõppimiseks. Siiski tuleb rõhutada, et need teadaolevad mutatsioonid pakuvad täna selgitust vaid 1:200 hüpertensiooni juhust (Kristjansson *et al.* 2002).

1.2.3 Neerud ja nende roll hüpertensiooni kujunemisel

Neerude ülesandeks on hoida konstantsetena kehavedelike maht, pH, koostis ja jaotus keha eri ruumide vahel; eritada toiduga organismi saanud vedelike ja soolade liiga ning kahjulikke ainevahetuse jääkprodukte. Keskmiselt filtreerivad neerud üle 170 liitri plasmat päevas, mis sisaldab umbes 23 mooli soolasid (Lifton *et al.* 2001).

Paljude teadlaste arvates on neerudel essentsiaalse hüpertensiooni kujunemisel tähtis roll (Bianchi *et al.* 1989), mida kinnitavad mitmed tänaseks tuvastatud neerus ekspresseeruvate geenide mutatsioonid, mis on seotud päriliku hüpertensiooniga (Tabel 2). Näiteks Liddle'i sündroom, mis on tingitud mutatsioonist distaalse nefroni epiteliaalse naatriumi-kanali kas beeta (kodeerib geen SCNN1B) või gamma subühikus (kodeerib geen SCNN1G). Mõlemal juhul on tulemuseks deletsioon valgu C-terminuses (45-75 ah), mille tagajärjeks on pidevalt funktsioneeriv kanal ning

Tabel 2. Mendeliiaarsed hüpertensiooni esinemise vormid inimesel (Gong and Hubner 2006)

Sündroom	Geen(id)	Geeni funktsioon, haiguse mehhanism
Mineraalkortikoidide liig (Mune <i>et al.</i> 1995)	<i>HSD11B2</i>	Geeni normaalse funktsiooni kaotus viib neeru mineraalkortikoidide retseptori võimendunud stimulatsioonile kortisooli toimet
Glükokortikoid- aldosteronism (Lifton <i>et al.</i> 1992)	<i>CYP11B2</i> ja <i>CYP11B1</i>	CYP11B2 ja CYP11B1 geenide ühinemise tagajärjel läheb aldosterooni sünteesi kontroll ACTH alla
Pseudohüpoaldosteronism. Tüüp I Liddle sündroom. (Shimkets <i>et al.</i> 1994)	<i>SCNN1B</i> ja <i>SCNN1G</i>	Mutatsioonid distaalse nefroni naatriumkanali beeta või gamma subühikus on põhjuseks suurenenud Na ⁺ ionide reabsorptsioonile
Pseudohüpoaldosteronism. Tüüp II. Gordons'i sündroom. (Wilson <i>et al.</i> 2001)	<i>WNK1</i> ja <i>WNK4</i>	Kodeerivad valke, mis mõjutavad ionide reabsorptsiooni distaalses nefronis
Raseduseaegne hüpertensioon (Geller <i>et al.</i> 2000)	<i>NR3C2</i>	Mutatsioon mineraalkortikoid- retseptori ligandiga seonduvas domeenis põhjustab hüpertensiooni. Täpsemad mehhanismid teadmata.

suurenenud naatriumsoolade resorptsioon, mis omakorda põhjustab kõrge vererõhu tingituna veremahu suurenemisest ja südame löögisageduse kasvust (Shimkets *et al.* 1994).

Neerud saavad toimida ühtlasi nii hüpertensiooni põhjustajatena kui ka “ohvritena” (Klahr 1989). Sõltuvalt hüpertensiooni raskusastmest ja kestvusest viib kõrge vererõhk varem või hiljem klassikaliste morfoloogiliste muutusteni neerudes, mis avaldub järkjärgulises päsmakeste hävimises, neeruparenhüümi skleroosis ja nefronite kaotuses, mille tulemuseks on progresseeruva kulgemisega neerukahjustus. Teisalt kaasneb neerupuudulikkusele peaaegu alati kõrge vererõhk, mis omakorda kiirendab neeru funktsiooni ja glomerulaalfiltratsiooni (GF) langust. Võimalike hüpertensiooni esilekutsuvate faktoritena käsitletakse ka neerude väikest massi, nefronite puudulikku embrüonaalset arengut või ebanormaalset funktsiooni (Brenner *et al.* 1988). Neeru massi osa hüpertensiooni kujunemises on tänaseni veel suhteliselt ebaselge, kuid siiski on mitmeid eksperimentaalseid ja kliinilisi töid, mis viitavad neeru massile kui täiesti iseseisvale riskitegurile. Väiksema neeru massi puhul on filtratsiooni pind väiksem ja võib tekkida hüperfiltratsioon, mis viib süsteemse arteriaalse vererõhu tõusule. Nii on kogunenud üha rohkem andmeid, mis viitavad asjaolule, et neeru mass ehk funktsioneerivate nefronite arv võib olla üheks iseseisvaks faktoriks hüpertensiooni patogeneesis (Brenner and Mackenzie 1997). Näiteks on mitmeid kliinilisi töid, eriti aga kliinilised transplantatsioonid, kus neerupuudulikkusega ja väljakujunenud kõrge vererõhuga loomadel suurendati neeru massi uue terve neeru siirdamisega. Tulemuseks oli vererõhu langemine normi piiridesse ning neeruhaiguse progresseerumise pidurdumine (Brenner and Anderson 1992).

1.2.4 Essentsiaalse hüpertensiooni kui komplekshaiguse päriliku komponendi kaardistamine

Hüpertensiooni kui multifaktoriaalse haiguse geneetilise tausta mõistmiseks (nagu ka teiste komplekshaiguste puhul) on peamisteks lähenemisviisideks:

- a) Aheldusanalüüs (*linkage analysis*)
- b) Assotsiatsiooniuuring (*association studies*)
- c) Mudelorganismide kasutamine (hüpertensiooni puhul enamasti rott)

1.2.4.1 Essentsiaalse hüpertensiooni uuringud aheldusanalüüsil

Aheldusanalüüsi meetod sai kõva tõuke 1990.a alguses kui hakkasid levima ülegenoomsed mikrosatelliitmarkerite abil koostatud geneetilised kaardid. Üks esimesi õnnestunud töid, kus tsüstilist fibroosi põhjustav geen leiti üles üksnes aheldusanalüüsiga, ilmus 1989 aastal. Tänapäevaks on sel meetodil kirjeldatud juba üle tuhande geeni (Botstein and Risch 2003). Aheldusanalüüs on olnud väga edukas meetod monogeensete haiguste kaardistamisel (Hirschhorn and Daly 2005). Aheldusanalüüsi põhimõtteks on jälgida geneetiliste markerite lahknemist sugupuus koos haigusmarkeriga (Kruglyak 1999). Tulenevalt kromosoomis lähestikku asuvate markerite aheldatusest ning jälgides haigetel indiviididel sama kromosoomipiirkonna või markeri esinemist, on võimalik prognoosida antud kromosomaalse regiooni seost vastava haiguse tekkeprotsessiga (Lander and Schork 1994). Komplekshaiguste haigusgeenide kaardistamisel on ülegenoomse aheldusanalüüsi suurimaks positiivseks küljeks hüpoteesivaba lähenemine, kuid siiski esineb antud meetodi kasutamisel ka mitmeid puuduseid (Collins *et al.* 1997):

1. mida kompleksem on haigus, seda raskem on aheldusanalüüsi kasutada;
2. komplekshaigustel esinevad vähese fenotüübilise toimega geenid, mille tuvastamiseks on vaja palju perekonnamaterjali;

3. madala penetrantsuse puhul raske defineerida haiguse avaldumise mudelit;
4. mudelivabal lähenemisel on väike võimsus;
5. komplekshaigused avalduvad reeglina hilja ning seetõttu ei ole võimalik uuringus kasutada suuri perekondasid;

Suurimateks hüpertensiooni kogu genoomi aheldusanalüüsi uuringuteks on olnud:

BRIGHT - läbiviijaks Briti Meditsiiniteaduste Nõukogu (*The Medical Research Council British Genetics of Hypertension*). Uuringu eesmärgiks oli teostada ülegenoomne aheldusanalüüs, et identifitseerida võimalikke hüpertensiooniga seotud kromosomaalseid regioone. Projektis osalesid 1599 Briti Euroopa päritolu hüpertensiivset perekonda, (3599 indiviidi 2010-st haigest sugulaspaarist) (Caulfield *et al.* 2003). Uuringu käigus saadi kõrgeimaks LOD skooriks 3,21, mis paigutus piirkonda 6q. Kuuendale kromosoomile on seega kokku viidatud viies uuringus (tabel 3. punkt 1, 2, 9, 14, 17). Tõenäoliste piirkondadena toodi välja veel 2q, mis oli varem tuvastatud kahes uuringus (tabel 3. punkt 7, 15,), 5q varem teada ühest uuringust (tabel 3. punkt 14,) ning 9q. Ükski eelpool nimetatud regioonidest ei sisaldanud varem teadaolevaid hüpertensiooni kandidaatgeene. Samuti suhtutakse skeptiliselt tuvastatud 6q piirkonnas saadud LOD skoori 3,21, mis klassikaliselt on olnud aktsepteeritav kui statistiliselt oluline tulemus, kuid Lander'i ja Kruglyak'-i poolt võrdsustatud statistilise olulisuse nivooaga 0,05 (Lander and Kruglyak 1995).

FBPP - läbiviija USA Rahvuslik Tervishoiu Instituut (*US National Institute of Health funded Family Blood Pressure Program*). Hõlmas selliseid projekte nagu GENOA (tabel 3, punkt 14), HyperGen (tabel 3, punkt 9), GenNet ja SAPHIR. Erinevate kriteeriumite alusel kaasati uuringutesse 6000 indiviidi, kes esindasid erinevaid etnilisi gruppe. Ühelgi neljast keskusest ei õnnestunud välja selgitada olulisi aheldusanalüüsi tulemusi ning lõpptulemusena kujunes projekt veelgi vähem edukaks kui BRIGHT uuring. Kõrgeimaks LOD skooriks saadi 2,96, mis lokaliseerus

Euroopa päritolu indiviididel piirkonda 1q, kuid lähemal vaatlusel antud regioonis siiski võimalikke hüpertensiooni kandidaatgeene ei tuvastatud (Mein *et al.* 2004).

FBPP ja BRIGHT uurimustöö tulemuste kõrvutamine kattuvaid lookuseid ei tuvastanud (Mein *et al.* 2004).

Hüpertensiooni aheldusuuringutest ühe suurima LOD skooriga (4,6) piirkond 18q (tabel 3. punkt 13) on leitud **DeCode** poolt läbi viidud Islandi hüpertensioonihaigete ja pereliikmete uuringus. See hõlmas kokku 490 perekonnaliiget 120-st perekonnast ning vaadeldi 904 mikrosatelliitmarkerit (Kristjansson *et al.* 2002). Kromosoomi 18q on varasemas uuringus tuvastanud ka (Pankow *et al.* 2000) (tabel 3. punkt 9).

Läbiviidud ülegenoomsed aheldusanalüüsi uuringud ennustavad paljude kromosomaalsete regioonide olulist rolli essentsiaalse hüpertensiooni kujunemises (Samani 2003). Siiski on lõplike järelduste tegemine problemaatiline, sest seniste uuringute tulemused kattuvad omavahel halvasti ning ei ole sageli geneetilise heterogeensuse või väikese võimsuse tõttu korratavad teistes populatsioonides (Gong and Hubner 2006). See omakorda tähendab, et tulemused on mitmeti tõlgendatavad ning reaalselt puudub võimalus kinnitada või välistada mingi piirkonna seost hüpertensiooni kujunemisega.

Tabelis 3 on ära toodud viimasel ajal teostatud suurimad inimese hüpertensiooniga seotud ülegenoomsed aheldusanalüüsi uuringud.

Tabel 3. Kokkuvõtte suurematest hüpertensiooni aheldusanalüüsi uuringutest (Binder 2006).

Jrk nr.	Viide uuringule	Tuvastatud piirkond ¹	Osalenud indiviide	Fenotüüp	Valimi päritolu
1.	(Caulfield <i>et al.</i> 2003)	2, 5, 6 , 8, 9	3599	HT	Euroopa
2.	(Hunt <i>et al.</i> 2002)	6	2959	SVR	Euroopa
3.	(Hunt <i>et al.</i> 2002)	1, 7, 12, 15	2959	HT	Euroopa
4.	(Levy <i>et al.</i> 2000)	17 , 18	1585	SVR	Euroopa
5.	(Xu <i>et al.</i> 1999)	3, 11, 15, 16, 17	1450	SVR, DVR	Hiina
6.	(Cooper <i>et al.</i> 2002)	2, 3, 5, 7, 10, 19	792	DVR	Nigeria
7.	(Hsueh <i>et al.</i> 2000)	2	694	DVR	Euroopa
8.	(Rice <i>et al.</i> 2000)	1, 2, 5, 7, 8, 19	679	SVR	Euroopa
9.	(Pankow <i>et al.</i> 2000)	6, 18	636	SVR	Euroopa
10.	(Harrap <i>et al.</i> 2002)	1, 4, 16, X	548	SVR	Euroopa
11.	(Rice <i>et al.</i> 2002)	3, 11	519	SVR	Euroopa
12.	(Atwood <i>et al.</i> 2001)	2 , 8, 18, 21	495	SVR	Mehhiko
13.	(Kristjansson <i>et al.</i> 2002)	2, 11, 17, 18	490	HT	Euroopa
14.	(Krushkal <i>et al.</i> 1999)	2, 5, 6, 9, 15 , 16, 18, 20, 21	427	SVR	Euroopa
15.	(Zhu <i>et al.</i> 2001)	2, 3, 8, 15	393	HT	Hiina
16.	(Cheng <i>et al.</i> 2001)	1, 2, 7	390	SVR	Euroopa
17.	(Allayee <i>et al.</i> 2001)	4, 6, 8, 19	388	SVR	Euroopa
18.	(Angius <i>et al.</i> 2002)	1, 2, 13, 15, 17, 19	77	HT	Euroopa

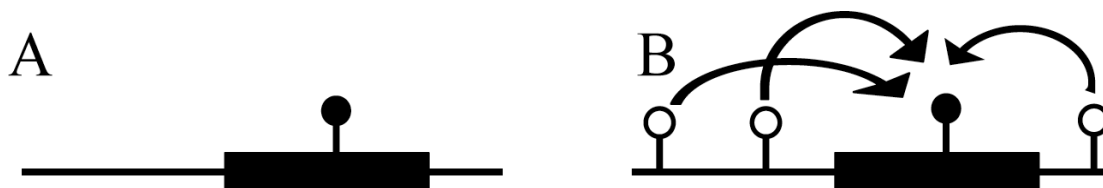
¹ Uuritud fenotüübiga seotud kromosoom; SVR – süstoolne vererõhk; DVR – diastoolne vererõhk; HT – hüpertensioon; **Tumedas kirjas** esile toodud statistiliselt olulised tulemused.

1.2.4.2. Essentsiaalse hüpertensiooni assotsiatsiooniuuringud

Üha rohkem mõistetakse, et aheldusanalüüs kui edukas meetod mendeliaarsete haiguste kaardistamiseks omab siiski suhteliselt vähest efektiivsust nende geenide lokaliseerimisel, mis osalevad multifaktoriaalsete haiguste tekkel (Collins *et al.* 1997). Komplekshaiguste geneetilise tausta mõistmiseks on sobivam rakendada assotsiatsiooniuuringuid, mille käigus võrreldakse geneetiliste markerite esinemissagedusi haigete indiviidide ja tervetest indiviididest moodustatud kontrollgrupi vahel (Cardon and Bell 2001). Saavutamaks assotsiatsiooniuuringul piisavat statistilist võimsust on uuringuks vajaminev indiviidide arv suur, kuid samas on neid võrreldes aheldusanalüüsi jaoks vajaminevate perekondade kogumisega lihtsam uuringusse kaasata (Binder 2006). Markeri suurem esinemissagedus patsientidel võrreldes kontrollgrupiga viitab alleeli assotsiatsioonile haigusfenotüübiga (Lander and Schork 1994). Klassikalised kandidaatgeenidel põhinevad assotsiatsiooniuuringud eeldavad eelnevat funktsioonipõhist lookuste ja testitavate alleelide valikut (info loomudelitest, monogeensetest sündroomidest, metaboolsete võrgustike uuringud, geeniekspressiooni uuringud), kuid positiivse küljena ei vaja pärilikkusmudeli defineerimist (Cambien *et al.* 1999). Kaheks traditsiooniliseks strateegiaks assotsiatsiooniuuringute läbiviimiseks on otsene strateegia ja kaudne strateegia. Otsese strateegia puhul määratakse ära funktsionaalsetes alades asuvad SNPd ehk cSNP-d (joonis 2 A). Enamus genoomis esinevatest SNPdest on neutraalsed, kuid teatud osa neist – eelkõige aminohapet muutvad SNPd – võivad haiguse avaldumist soodustada või vastupidi, seda välistada (Collins *et al.* 1999). Võrreldes omavahel haigete ja tervete indiviidide vastavaid positsioone, peaksid eeldatavalt riski kandvad SNP-d haigetel sagedamini esinema (Kruglyak 1999).

Kaudne strateegia kujutab endast mingite genoomi regioonide (või kogu genoomi) SNPde uurimist ja LD analüüsi. Kuna kõik funktsionaalsed muutused genoomis ei pruugi asuda kodeerivates alades, määratakse SNP-d nii kodeerivatest kui mittekodeerivatest aladest ja otsitakse juhuslikkuse printsiibil potentsiaalseid

haigusalleele (Collins et al. 1997). Kaudse strateegia eelduseks on alleelse aheldatuse ehk LD esinemine geneetiliste markerite ning haiguse tekkes osalevate geenide/alleelide vahel (joonis 2 B).



Joonis 2. Assotsiatsiooniuringul kasutatavad kaks strateegiat. A juhul eeldatakse, et uuritav SNP (täis ring) on otseses seoses haigusfenotüübiga, B juhul on uuritavad SNP-d (seest tühjad ringid) kaudselt seotud haigust põhjustava polümorfismiga ahelduse tasakaalustamatuse tõttu (Hirschhorn and Daly 2005).

Seoses uute ja võimsate genotüpseerimise meetodite väljatöötamisega on massiliselt kasvanud hüpertensiooni assotsiatsiooniuringud, mis baseeruvad kas mingil konkreetsel kandidaatgeenil, paljudel kandidaatgeenidel või siis üle kogu genoomi analüüsil.

Hüpertensiooni ja ka teiste kardiovaskulaarhaiguste enim uuritud kandidaatgeenid on seotud reniin-angiotensiin-aldosteroon (RAAS) süsteemiga, mille ühte võtmeensüümi (angiotensiin I-te konverteeriv ensüüm) kodeerivat geeni *ACE* on uuritud paljudes assotsiatsiooniuringutes nii iseseisvalt kui ka koos teiste kandidaatgeenidega. Euroopas ühe suurima kardiovaskulaarhaiguste kandidaatgeenide uurimisega tegeleva keskuse andmebaasis (GeneCanvas <http://ecgene.net/genecanvas/>) leidub kirjeid 18 projekti kohta, kus on uuritud *ACE* geeni. Vaatamata sellele ei ole tänaseks *ACE* geeni assotsiatsiooniuringute tulemused üheselt mõistetavad ning leidub nii positiivseid kui ka negatiivseid assotsiatsioone. Nii on RAAS süsteemi geenidega *ACE*, *AGT* ja *ATRI* sooritatud assotsiatsiooniuringus leitud *ACE* geeni 16 intronis paikneva 287 bp pikkuse Alu-insertsioon/deletsiooni kui ka SNP T3892C seos hüpertensiooniga (Barbalic *et al.*

2006). Samas on läbi viidud 13 kandidaatgeeni (s.h. *ACE* ja *AGT*) hõlmav assotsiatsiooniuuring, kus seos kõrge vererõhuga leiti ainult geenidel *APOC3* ja *LPL* (Sass *et al.* 2004). Euroopas läbiviidud suurimad hüpertensiooni assotsiatsiooniuuringud koos enim uuritud hüpertensiooni kandidaatgeenidega on ära toodud tabelis 4. Prantsusmaal läbi viidud "PEGASE" uuringus vaadeldi 16-st kandidaatgeenist 43 SNP-d, millest on publitseeritud tulemused *ADRB2* geeni kohta (Herrmann *et al.* 2002) ja *GNB3* geeni kohta (Brand *et al.* 1999), mille puhul assotsiatsiooni hüpertensiooniga ei tuvastatud. "PEGASE" uuringust kõige värskemad avaldatud andmed näitavad *SAH* geeni seost hüpertensiooniga (Telgmann *et al.* 2007) .

Tabel 4. Euroopas teostatud suuremad hüpertensiooni assotsiatsiooniuringud (<http://genecanvas.idf.inserm.fr/>) ja enim uuritud kandidaatgeenid (Binder 2006).

	Vaadeldud geenid	Uuritud SNP-de arv	Uuritud patsiente	Uuritud kontroll- indiviide
BENSOC uuring (Hüpertensioon)	<i>ACE, ADD1, AGT, AGTR1, CYP11B2, EDN1, EDNRA, EDNRB, NOS3, NPPA, PAII, TGFB1</i>	20	500	0
HYPERTENSION/ PWV uuring (Hüpertensioon)	<i>ACE, ADD1, AGT, AGTR1, CYP11B2, EDN1, EDNRA, EDNRB, NOS3, NPPA, PAII, TGFB1, CAV2, GINB3, SELP</i>	34	350	150
PEGASE uuring (Hüpertensioon)	<i>ACE, ADD1, ADDA2A, ADRA2A, ADRB2, AGT, AGTR1, CAV2, CYP11B2, ECE1, EDN1, EDNRA, EDNRB, GCGR, GNB3, IGF1, SAH</i>	43	800	800
Enim uuritud hüpertensiooni kandidaatgeenid (Binder 2006)	<i>ACE, ADD1, ADD2, ADR2A2, ADRA1A, ADRA2B, ADRB1, ADRB2, ADRB3, AGT, AGTR1, AMPD1, ANP, ATP1B1, ATP1A2, BDKRB2, CLCNKB, CYP11B2, EDN1, EDN2, EDNRA, EDNRB, ET1R, GCGR, GGT1, GNAS1, GNB3, GPD, GSK1, GSY, HSD11B1, HSD11B2, IGF1, INSR, LMNA, LPL, NKCC2, NOS1, NOS2A, NOS3, NPPA, NPR2, NPRA, NPRB, NR3C1, PDG, PLA2, PLA2G1B, PPARG, REN, ROMK, SAH, SCNN1A, SCNN1B, SCNN1G, SGK1, TGFB1, TNFRSF1B, VEGF, VIAR</i>			

1.2.4.3. Essentsiaalse hüpertensiooni uuringud loomudelitel abil

Haiguste loomudeleid on kasutatud juba kaua aega inimeste haiguste patoloogia mõistmiseks ja uute võimalike ravisuundade välja töötamiseks. Hüpertensiooni puhul on kujunenud enim kasutatavaks mudelorganismiks rott. Vaatamata roti ja inimese ~70 miljonit aastat tagasi toimunud evolutsioonilisele lahknemisele (Kumar and Hedges 1998), on roti peamised organismi soolasisaldust, neeru funktsiooni ja vererõhku reguleerivad mehhanismid esindatud samal keerukuse tasemel, mis inimeselgi (Bianchi *et al.* 2004). Rohkem kui 30 aastat tagasi hakkasid teadlased eksperimentaalselt rotte ristama, et luua kõrge ja madala vererõhuga liinid (Dahl *et al.* 1962). Esimeseks teadlaseks oli K. Dahl, kes näitas samanimelise rotiliini peal, et organismi reaktsioon kõrge soola tarbimisele on vahendatud neerudega (Dahl *et al.* 1974). Tänapäevaks on loodud erinevaid rotiliine nagu näiteks SHR (*spontaneously hypertensive rat*) – spontaanselt hüpertensiivsed rotid, SHRSP (*stroke-prone spontaneously hypertensive rat*) – insuldialtid spontaanselt hüpertensiivsed rotid ja Dahl soolatundlik liin, kus hüpertensiooni indutseerimiseks on vajalik kõrge soolasisaldusega dieet (Rapp 2000).

Hüpertensiooni kui kvantitatiivse tunnuse uurimiseks loomudeli abil on ristatud hüpertensiivseid ja normaalse vererõhuga rotiliine, et otsida aheldatust vererõhu ja geneetiliste markerite vahel. Eesmärgiks on vererõhku kontrolliva(te) lookus(t)e ehk QTL-ide (*quantitative trait loci*) identifitseerimine. QTL kaardistamine on meetod, mis lähtub fenotüübist ja seega ei vaja eelnevaid teadmisi kandidaatgeenidest või nende funktsioonist, küll aga võib selle meetodi kasutamine viia haigusega seotud uute kandidaatgeenide tuvastamiseni (McBride *et al.* 2003).

Esimene vererõhu regulatsiooniga seotud ülegenoomne QTL analüüs teostati 1991. a. ning identifitseeriti võimalikud hüpertensiooniga seotud lookused roti 10, 18 ja X kromosoomis (Hilbert *et al.* 1991); (Jacob *et al.* 1991). Tänapäevaks on kandidaat QTL-regioonid leitud praktiliselt igas roti kromosoomis, mis kinnitab jällegi fakti, et hüpertensioon on polügeenne haigus. Samas on probleemiks leitud piirkondade

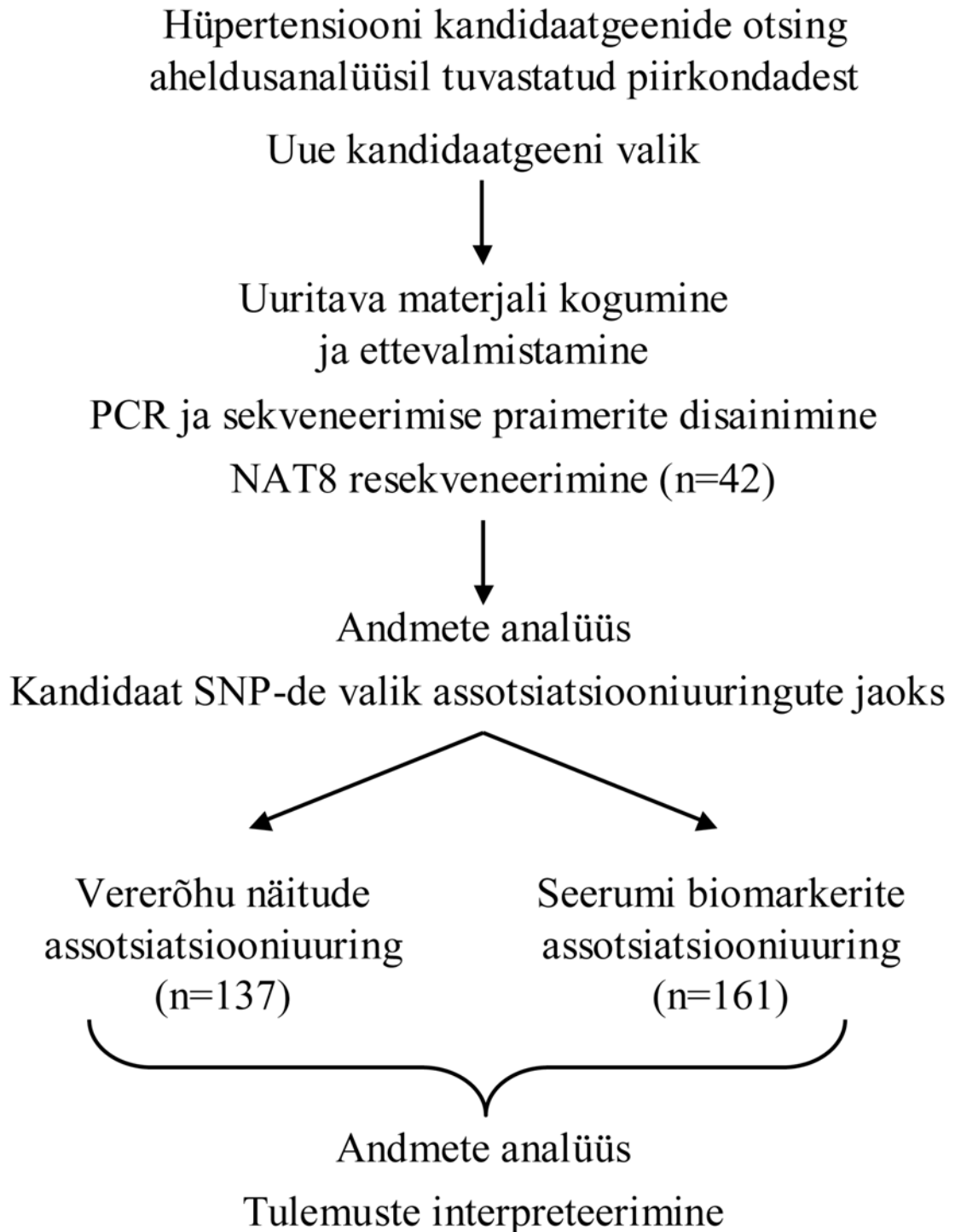
kümnete cM pikkune ulatus, muutes konkreetse kandidaatgeeni tuvastamise raskeks (Rapp 2000).

Ühe suurema uuringu tulemusena, kus otsiti uusi hüpertensiooniga seotud piirkondasid võrdleva genoomika abil, tuvastati 67 QTL piirkonda rotil. Vastavad segmendid kõrvutati inimese 26 kromosomaalse regiooniga 16 autosoomis. Leitud piirkondadest ühtis eelnevalt inimese ülegenoomsetel aheldusanalüüsidel tuvastatud piirkondadega 5 regiooni (1q, 2p13, 5q31, 15q22 ja 17q) (Stoll *et al.* 2000).

2. TÖÖ EESMÄRGID

1. Uute hüpertensiooni kandidaatgeenide tuvastamine lähtudes kirjanduses publitseeritud aheldusanalüüsi uuringute tulemustest.
2. Kromosoomis 2 paikneva uue hüpertensiooni kandidaatgeeni *NAT8* polümorfse varieeruvuse kirjeldamine Eesti populatsioonis.
3. Teostada *NAT8* geeni 5' ülesvoolu piirkonna SNP-dega assotsiatsiooniuringud vererõhu, vereseerumi biomarkerite ning glomerulaar-filtratsiooni väärtustega.
4. Teostada *in silico* analüüs hindamaks *NAT8* 5' ülesvoolu paiknevate SNP-de alleelide potentsiaalset mõju *NAT8* ekspressioonile.

3. TÖÖ SKEEM



4. MATERJAL JA METOODIKA

Käesoleva teadusuuringu läbiviimiseks on olemas Tartu Ülikooli Eetikakomitee luba nr. 122/13,22.12.2003; 137/20, 25.04.2005 ning kõik uuringus osalenud inimesed on allkirjastanud informeeritud nõusoleku vormi.

4.1 *NAT8* geeni amplifikatsioon inimese genoomist

Antud töö läbiviimiseks vajalik *NAT8* genoomse regiooni järjestus saadi NCBI andmebaasist (<http://www.ncbi.nlm.nih.gov>). *NAT8* geeni amplifitseerimiseks ja resekveneerimiseks vajalikud praimerid disainiti *Whitehead Institute for Biomedical Research* ja *Howard Hughes Medical Institute's* välja töötatud programmi Primer3 abil, kuhu oli juurdepääs veebipõhise kasutajaliidesega (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Kokku disainiti 4 PCR praimerit ja 8 resekveneerimise praimerit, mille unikaalsust kontrolliti BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) programmiga. Tulenevalt genoomse piirkonna iseärasustest ning vältimaks duplikaatgeenide üheaegset amplifitseerumist, teostati PCR kahes etapis kasutades kahte praimerite paari (tabel 5). Esmalt amplifitseeriti välimiste PCR praimeritega 10341 bp pikkune ala ning järgnevalt "sisemiste" PCR praimeritega soovitud 2322bp pikkune ala (vt. 5.4 joonis 6). Genoomse DNA (~100 ng) esmane PCR (10341 bp) teostati pikkade produktide amplifitseerimiseks välja töötatud DNA polümeraasiga (Long PCR Enzyme Mix, MBI Fermentas, Vilnius, Leedu) vastavalt tootja poolt soovitatud protokollile kasutades ABI GeneAmp PCR System 2700 (Applied Biosystems, Foster City, CA, USA) masinaid ning järgnevat PCR programmi:

Algne denaturatsioon	94°C	5 min	
Denaturatsioon	94°C	20 sek	} 10 tsükliit
Praimerite seondumine	68°C	30 sek	
Ekstensioon	68°C	7 min	
Denaturatsioon	94°C	20 sek	} 20 tsükliit (+5 sek/tsükkel)
Praimerite seondumine	68°C	30 sek	
Ekstensioon	68°C	7 min	
Inkubatsioon	68°C	10 min	
Säilitamine	4°C		

NAT8 geeni teine PCR, sisemiste praimeritega, sooritati 0,4 µl esmase (10341 bp) PCR produktiga kasutades Smart-Taq Hot DNA polümeraasi (Naxo, Tartu, Eesti) ning ABI GeneAmp PCR System 2700 PCR-i masinaid järgnevatel reaktsioonitingimustel:

Algne denaturatsioon	94°C	5 min	
Denaturatsioon	94°C	20 sek	} 10 tsükliit temp. langus tsükli kohta 1C°
Praimerite seondumine	68°C	30 sek	
Ekstensioon	68°C	2 min	
Denaturatsioon	94°C	20 sek	} 10 tsükliit
Praimerite seondumine	56°C	30s ek	
Ekstensioon	68°C	2 min	
Denaturatsioon	94°C	20 sek	} 10 tsükliit
Praimerite seondumine	54°C	30 sek	
Ekstensioon	68°C	2 min	
Denaturatsioon	94°C	20 sek	} 10 tsükliit
Praimerite seondumine	51°C	30 sek	
Ekstensioon	68°C	2 min	
Inkubatsioon	68°C	10 min	
Säilitamine	4°C		

Teise amplifikatsiooni lõpus saadud PCR produkt sisaldas kogu *NAT8* geeni (1569 bp), geeni 5' ülesvoolu järjestust (541 bp) ja geeni 3' allavoolu järjestust (75 bp), moodustades kokku 2185 bp pikkuse ala, mis resekveneeriti mõlemalt ahelalt Eesti populatsiooni juhuvalmis (n=42).

Tabel 5. *NAT8* PCR ja resekveneerimise praimerid

Praimeri nimetus	Praimeri järjestus 5'-3'
Välimised PCR praimerid	
NAT8_10kb_F	TGTAATCCCACACTTTTAGAGGCCAAG
NAT8_10kb_R	GAGTAGGTAAACAAAGCAGCCAGGAA
Sisemised PCR praimerid	
NAT8_2kb_F	GCAAATTGCTGACTTCAGACTC
NAT8_2kb_R	CAACAGCCAGAACACACACAAG
Sekveneerimise praimerid	
NAT8_seq_1F	TGCAACATTATGCTGGCACT
NAT8_seq_2F	CAGACTCAGCCACCCAGAAG
NAT8_seq_3F	AGGGTCAGAAAGACCTGCATT
NAT8_seq_4F	GCATCCAGGAGATACCGTCT
NAT8_seq_1R	ATTGCTGAAGCTGCCTCGAAC
NAT8_seq_2R	AGGTAGGTGCCCCCTCAAT
NAT8_seq_3R	CCTGAACAGGCAAGAGAGGT
NAT8_seq_4R	GTCAAGCCGAATGTCAACCT

4.2 PCR produktide analüüs ja töötlus

Soovitud PCR produkti olemasolu kontrolliti geelelektroforeesil. PCR produktis kasutamata jäänud praimerite ning desoksünukleotiidide eemaldamiseks kasutati eksonukleas I-te (Exo I) ja krevetil aluselist fosfataasi (sAP) (MBI Fermentas). Olemasolevatele PCR produktidele lisati 1U Exo I ja 1,5U sAP ensüümi ning ensüümtötluse reaktsioonid viidi läbi ABI GeneAmp PCR System 2700 masinatega järgmistel tingimustel:

Inkubeerimine	37°C	20min
Inaktiveerimine	80°C	15min
Säilitamine	4°C	

4.3. NAT8 geeni resekvenerimine

Puhastatud PCR proovidega (2185 bp) teostati sekveneerimisreaktsioonid kasutades selleks sekveneerimise primereid 1F, 2F, 3F, 4F, 1R, 2R, 3R ja 4R, mis olid disainitud ~80–100 bp ülekattuvusega, tagamaks korrektset kontiigide assambleerimist. Assotsiatsiooniuuringu jaoks 5'ülesvoolu piirkonna resekvenerimisel kasutati primereid 4F ja 4R (tabel 5).

Sekveneerimise reaktsioonide läbiviimiseks kasutati DYEnamic ET Terminator Cycle sekveneerimiskitti (Amersham Biosciences Inc., Milwaukee, WI, USA). Reaktsioonisegu kogumahus 10 µl sisaldas: 1 µl *Thermo Sequenase II* reaktsiooni "premix"-i, 3 µl lahjenduspuhvrit (100 µl 10X PCR puhvrit, 55 µl 25mM MgCl₂, 95 µl ddH₂O), 1 µl 2pM praimerit, 2 µl ddH₂O ja 3 µl vastavat puhastatud PCR produkti. Sekveneerimisreaktsioonid viidi läbi ABI PCR GeneAmp PCR System 2700 masinatega järgmistel tingimustel:

Denaturatsioon	95°C	35 sek	} 45 tsükli
Praimeri seondumine	50°C	15 sek	
Ekstensioon	60°C	1 min	
Säilitamine	4°C		

Järgnevalt proovid sadestati kasutades 10M ammoonium-atsetaati ja dekstraani. Igale proovile (10 µl) lisati 1,5 µl ammoonium-atsetaadi ja dekstraani segu:

- 1 µl ammoonium-atsetaat puhver (10M [pH>8])/EDTA (250mM)
- 1 µl dekstraani (10mg/ml)

Produktid segati hoolikalt ning igale lisati 30 µl 96% külma etanooli. Seejärel asetati 15 min sadenema –20°C juurde.

Proovid tseentrifuugiti 15 min. 16,1 rcf juures ning eemaldati supernatant. Igale proovile lisati 200 µl 70% külma etanooli ning tseentrifuugiti 5min. 16,1 rcf. Taas eemaldati supernatant ja seejärel proovid kuivatati 10-15min. 37°C juures. Viimaks lisati igale proovile 3 µl *ABI loading solution*'it (formamide loading dye, Amersham Biosciences Inc.) ning sekveneeriti ABI PRISM 377 sekvenaatoriga (Applied Biosystems).

4.4. DNA järjestuste esmane analüüs ja visualiseerimine

Esmaseks sekveneerimistulemuste kvaliteedi kontrolliks kasutati programmi BioEdit Sequence Alignment Editor (Hall T., *Department of Microbiology North Carolina State University*), mille abil vaadati läbi kõik sekveneeritud järjestused ning otsustati nende sobivus edasiseks analüüsiks.

Kontiigid assambleeriti programmiga Phrap (<http://www.phrap.org>), skaneeriti programmiga PolyPhred ja analüüsiti kasutades programmi Consed (Gordon *et al.*

2001). PolyPhred-i poolt identifitseeritud heterosügootsed positsioonid kontrolliti ning kinnitati edasiseks analüüsiks.

4.5 Statistiline analüüs

DNA järjestuse analüüsile ja polümorfsete positsioonide tuvastamisele järgnes statistiline analüüs kasutades programme HAPLOVIEW (version 3.32) (Barrett *et al.* 2005), PLINK (version 0.99), DNASP (version 4.10) (Rozas and Rozas 1999) ja GENEPOP (version 3.4) (Rousset and Raymond 1995). Kõigis kasutatud statistilistes testides loeti statistiliselt tähendusrikkaks $P < 0,05$.

1. Markerite üldine iseloomustus. Esmalt määrati identifitseeritud SNP-de alleelisagedused ning vastavus Hardy-Weinbergi tasakaalule (HWE) ($p^2 + 2pq + q^2 = 1$), kus p ja q on antud lookuse (näiteks SNP) alleelisagedused populatsioonis. Selleks kasutati programmi HAPLOVIEW, PLINK ja GENEPOP. HWE hindamiseks kasutatakse *Fisher's exact test*'i ning piiriks peeti P -väärtust $< 0,05$. Kõik tuvastatud SNP-d (v.a *singleton* positsioonid) vastasid HWE tasakaalule, seega võis resekveneerimise kvaliteeti pidada usaldusvääseks.

2. Assotsiatsioonianalüüsid viidi läbi programmi-paketiga PLINK (<http://pngu.mgh.harvard.edu/~purcell/plink/>), mis on paindlik käsuraprogramm ülegenoomsete genotüübi ja fenotüübi andmete analüüsiks. Hindamaks vererõhu, biomarkerite ja GF-i assotsiatsiooni 5' ülesvoolu paiknevate SNP-dega kasutati aditiivset alleelidoosi mudelit, mida loetakse sobivaks just komplekshaiguste päriliku komponendi hindamiseks. Kasutatud test võimaldas analüüsil arvesse võtta ka kaasuvaid parameetreid nagu näiteks vanus ja sugu. Paralleelselt kontrolliti genotüübi ja fenotüübi vahelist seost Mann-Whitney U testiga (<http://elegans.swmed.edu/~leon/stats/utest.cgi>), mis võrdleb andmepunktide jaotuvust ning ei vaja võrdset andmepunktide arvu kahes valimis.

3. Nukleotiidide mitmekesisusi ja neutraalsushüpoteesi kirjeldavate parameetrite arvutamiseks kasutati programmi DnaSP (Version 4.10):

- a. Nukleotiidide mitmekesisus (π) (*nucleotide diversity*) – keskmine nukleotiidide erinevus positsiooni kohta kahe järjestuse vahel.
- b. Theta (θ) on arvatud kogu polümorfismide arvust – $4N_e\mu$ (N_e – efektiivne populatsiooni suurus, μ – mutatsiooni määr nukleotiidi kohta generatsioonis). Theta (Watterson 1975) põhineb järjestuse olemasolevate polümorfsete positsioonide arvul, sõltub järjestuse pikkusest ja valimi suurusel.
- c. Tajima's D – testib hüpoteesi, et kõik mutatsioonid on selektiivselt neutraalsed. D test põhineb kõikide polümorfismide (Theta) ja kahe järjestuse vaheliste keskmiste erinevuste (π) erinevusel. Standardse neutraalse mudeli korral Theta ja π on võrdsed ja nende vahe on lähedane nullile. Positiivne Tajima D statistiline parameeter näitab keskmise sagedusega variatsioonide suuremat sagedust ning viidab selle geeni suhtes toimuvale positiivsele selektsioonile või sellele, et antud populatsioon koosneb mitmest erineva haplotüüpse mustriga alapopulatsioonidest, mis on omavahel segunenud (*admixture*). Negatiivne Tajima D parameeter näitab harvade variatsioonide suuremat sagedust, mille põhjusteks võivad olla puhastav valik ja hitchhiking'u efekt ja populatsiooni kasv

4. Geenisisese aheldatuse määramine. Alleelne aheldatus ehk ahelduse tasakaalustamatus (*linkage disequilibrium* ehk LD) on kahe (või enama) genoomse lookuse alleelne või genotüüpne mittejuhuslik, reeglipärane assotsiatsioon (Ardlie *et al.* 2002). Kasutades programmi GENEPOP (version 3.4) arvatud *Fisher's exact test*'i hinnati LD esinemist geeni sees paiknevate markerite vahel. *Fisher exact test*'i piiriks peeti P-väärtust $<0,05$. Lisaks hinnati LD tugevust markerite vahel ka programmiga Haploview (Versioon 3.2), mis kasutas r^2 testi, kus r^2 näitab statistilist korrelatsiooni kahe markeri vahel. Vastavad r^2 väärtused on 0 kuni 1 (0 – ei ole LD-d, 1 – täielik LD).

4.6 *In silico* analüüs

NAT8 genoomse piirkonna järjestus saadi NCBI (<http://www.ncbi.nlm.nih.gov/>) andmebaasist. *NAT8* ja *NAT8B* järjestuste omavaheliseks joondamiseks ning paraloogsete järjestuste variantide kirjeldamiseks kasutati veebipõhist ClustalW programmi (<http://www.ebi.ac.uk/clustalw/>).

In silico ennustatud transkriptsioonifaktorite seondumissaite kontrolliti programmi MatInspector 2.2 (<http://www.genomatix.de/products/MatInspector/>) veebipõhise variandiga, mis põhineb ekperimentaalselt tõestatud transkriptsioonifaktorite seondumissaitidel TRANSFAC-i andmebaasis (<http://www.gene-regulation.com/pub/databases.html>).

4.7 Assotsiatsiooniuuringus osalevad indiviidid

Uuritavaid indiviide koguti kahe assotsiatsiooniuuringu jaoks:

1. Uurimaks *NAT8* potentsiaalse promootorala genotüüpide seost vererõhuga.
2. Uurimaks *NAT8* potentsiaalse promootorala genotüüpide seost vere seerumi biomarkerite ning neeru funktsiooni iseloomustava glomerulaarse filtratsiooniga (GF).

Indiviidide kogumine *NAT8* assotsiatsiooniuuringu jaoks vererõhu näitajatega toimus koostöös Põhja-Eesti ja TÜ Kliinikumi verekeskustega, kust kaasati kokku 137 inimest, kelle detailsem iseloomustus on toodud tabelis 6. Kõik uuringus osalejad olid korduvad veredoonorid, kelle vererõhunäidud olid fikseeritud vere loovutamisele eelnenud meditsiinilise läbivaatuse käigus, kus hinnati ka indiviidi üldist tervislikku seisundit ja muuhulgas välistati ka ravimite (s.h. vererõhuravimite) tarbimine. Assotsiatsioonianalüüs viidi iga uuringus osaleja puhul läbi mitme aasta vererõhunäitudest arvatud mediaanväärtusega, mis analüüsi käigus kohandati vastavalt isiku soole ja vanusele vere loovutamise hetkel.

Teine grupp uuritavaid moodustati vere seerumi biomarkerite ja GF-i assotsiatsiooni analüüsiks ning koguti koostöös Tartu Ülikooli Kliinikumi (dots. Tiina Ristimäe, dr.

Gudrun Veldre, dots. Mai Rosenberg) ja Põhja-Eesti Regionaalhaiglaga (prof. Margus Viigimaa, dr. Aleksei Teor). Kokku kaasati uuringusse 161 indiviidi, kes vastasid alljärgnevatele kriteeriumitele:

- a) Diagnoositud essentsiaalne hüpertensioon (EH) ja/või hüpertensiivne nefropaatia (HN);
- b) EH-i pärilik taust;
- c) EH vastaste ravimite tarbimine;
- d) Sekundaarset hüpertensiooni põhjustavate haiguste puudumine;

Uuringurühma kuulusid 81 meest ja 80 naist, kelle keskmine vanus EH diagnoosimise hetkel oli $47 \pm 9,7$ aastat ning keskmine BMI uuringusse kaasamise hetkel $28 \pm 3,8$. Analüüsitavateks biomarkeriteks olid TÜ Kliinikumi Ühendlaboris ja PERH laboris vere seerumist määratud urea, kreatiniin ja albumiin, mida kasutati GF-i arvutusteks ning HDL/LDL kolesterool ja triglütseriidid, millega viidi läbi assotsiatsioonianalüüsid. Vereseerumi biokeemiliste näitajate keskmised väärtused on toodud tabelis 7. Hindamaks neerude üldist funktsiooni arvutati GF, mis põhines MDRD (*Modification in Diet in Renal Disease*) uuringus väljatöötatud valemil (Levey *et al.* 1999):

$$\text{GFR}(\text{ml}/\text{min}/1,73\text{m}^2) = 170 \times (\text{Kreatiniin})^{-0,999} \times (\text{Vanus})^{-0,176} \times (0,762 \text{ kui naine}) \times (\text{Urea})^{-0,17} \times (\text{Albumiin})^{+0,318}.$$

Valem võimaldas arvestada nii indiviidi demograafilisi andmeid (vanus, sugu, rass) kui ka vere seerumi biokeemiliste markerite väärtuseid (urea [mg/dl], kreatiniin [mg/dl] ja albumiin [g/dl]). Kõik biomarkeri assotsiatsioonianalüüsid, välja arvatud GF analüüs, kohandati samuti vastavalt indiviidi vanusele ja soole.

Tabel 6. Vererõhu assotsiatsioonianalüüsil osalenud indiviidide kirjeldus

Indiviidide arv (n)	137
mehed/naised (n)	76/61
Vanus ¹ (aastat)	40 ± 7,8
BMI	25 ± 2,6
Süstoolne vererõhk ² (mmHg)	134,2 ± 11,5
Min./Max. süstoolne vererõhk (mmHg)	113,5/171
Diastoolne vererõhk ² (mmHg)	84,6 ± 7,6
Min./Max. diastoolne vererõhk (mmHg)	65,5/110
Analüüsis kasutatud keskmine vererõhunäitude arv indiviidi kohta	4
Min/max vererõhunäitude arv ühe indiviidi kohta	2/14

¹ Keskmine vanus vererõhu mõõtmiste ajahetkel ± standardhälve ; ² keskmised väärtused ± standardhälve

Tabel 7. Uuritavate indiviidide keskmised vereseerumi biokeemilised väärtused

Biomarker	Keskmine ^A	Referents ^B	Ühik
Uurea (n=158)	7,8 ± 5,1	<8.3	mmol/l
Kreatiniin (n=158)	M: 133 ± 170	N:44 - 80	µmol/l
	N: 92 ± 114	M:62 - 106	
Albumiin (n=152)	43,4 ± 4,4	34 - 48	g/l
HDL kolesterool (n=149)	1,5 ± 0,4	>1,0	mmol/l
LDL kolesterool (n=149)	3,8 ± 1,0	< 3.0	mmol/l
Triglütseriidid (n=146)	2,0 ± 2,2	<1,7	mmol/l
GF (n=157)	72,6 ± 28,5	80 - 120	ml/min/1,73m ²

^A Keskmine ± standardhälve; ^BPERH labori koduleheküljelt

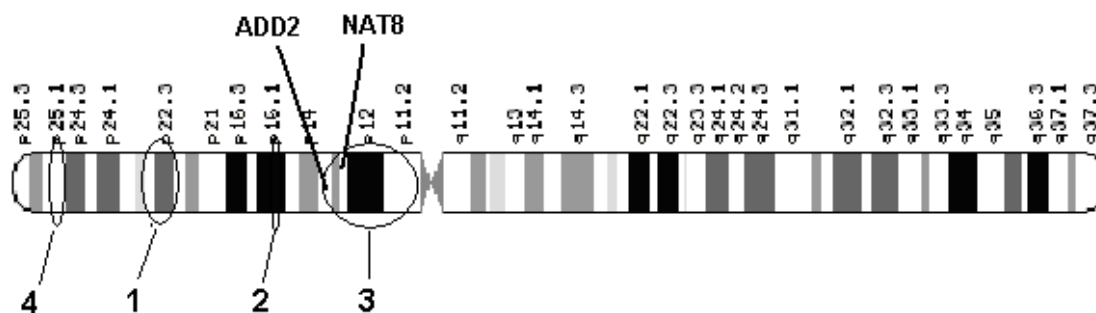
5. TULEMUSED

5.1 Uute hüpertensiooni kandidaatgeenide tuvastamine ahelduse analüüsil positiivseteks osutunud piirkondadest

Uute hüpertensiooni kandidaatgeenide otsingud andmebaasidest teostati lähtudes kirjanduses esitatud andmetest. Lähtepunktiks oli Samani (2003) poolt avaldatud artikkel, mis andis ülevaate 20-st üle kogu maailma läbiviidud aheldusanalüüsi uuringust ja nende tulemustest (Samani 2003). Detailsemaks analüüsiks valiti välja 25 regiooni 14-st kromosoomist (1q, 2p, 2q, 3p, 5p, 5q, 6q, 7q, 9q, 11q, 12q, 15q, 16q, 18q ja 19p) ning tähelepanu pöörati kandidaatgeenidele, mis on seotud just neerude arengu või metaboolsete radadega. Andmebaasi otsingu tulemusena tuvastati uuritavates genoomsetes piirkondades mitmeid juba varem teadaolevad hüpertensiooni kandidaatgeene nagu näiteks *SAH*, *SCNN1B*, *SCNN1G*, *NEDD4L*, *ADD2* (tabel 4, punkt 1). Kromosoomi 2 lühikeses õlas identifitseeriti seni varem vererõhu regulatsiooniga seostamata neerus ekspresseeruv *NAT8* (Ozaki *et al.* 1998).

5.2 Kromosoomi 2p piirkonna iseloomustus

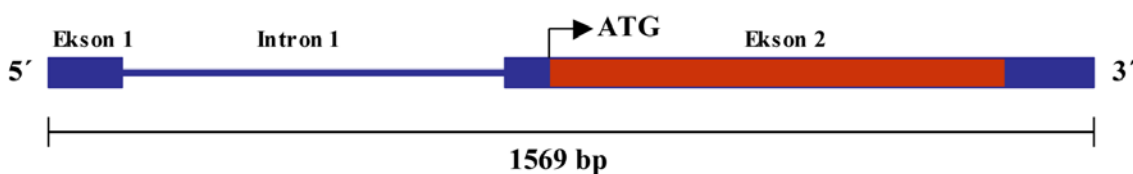
Kuna neljas uuringus on viidatud just teise kromosoomi lühikeses õlas paiknevate piirkondade olulisele rollile essentsiaalse hüpertensiooni kujunemisel, siis pöörati ka uue kandidaatgeeni valikul tähelepanu nendele regioonidele ning otsustati sealt edasiseks põhjalikumaks analüüsiks välja valida 2p13.1-p12 piirkonnas lokaliseeruv *NAT8*. Märkimisväärne on fakt, et koos *NAT8*-ga paikneb samas piirkonnas ka varem teadaolev hüpertensiooni kandidaatgeen *ADD2* (joonis 3). *NAT8* sobivus antud uuringusse uue hüpertensiooni kandidaatgeenina seisnes eelkõige tema võimalikul seosel neerude arengu ja talitlusega (Ozaki *et al.* 1998). Ühtlasi oli ajendavaks ka seni geeni kohta suhteliselt vähene teadaolev informatsioon ning resekveneerimiseks sobilik suurus (1569 bp).



Joonis 3. Kromosoomis 2p hüpertensiooniga seotud regioonid. 1. (Krushkal *et al.* 1999) – 2p (57-59cM); 2. (Rice *et al.* 2000) – 2p (86cM); 3. (Rice *et al.* 2002) - 2p (96-115cM); (Atwood *et al.* 2001) - (103 cM); 4. (Angius *et al.* 2002) – 2p (26,5-27,1cM).

5.3 NAT8 iseloomustus ja konserveeritus

NAT8 (N-atsetüültransferaas 8) geen lokaliseerub inimese 2. kromosoomis piirkonnas 2p13.1-p12. Geeni pikkuseks on 1569 bp, olemas on üks kodeeriv ekson ning transleeritava valgu pikkuseks on 227 aa molekulmassiga 25617 Da (joonis 4). Kuigi *NAT8* poolt kodeeritud valgu funktsioon on täpselt teadmata, lubab selle sarnasus atsetüültransferaasidega ning kõrge ekspressioon terves neerus ja maksas oletada selle tähtsust vastavate organite arengul ja normaalsel talitlusel (Ozaki *et al.* 1998).



Joonis 4. *NAT8* geeni intron/ekson struktuur

Sarnaste domeenide olemasolu viitab geeni evolutsiooniliselt konserveerunud funktsioonile. *NAT8* valgu identsus šimpansi *NAT8*-ga 99%, hiire *Nat8*-ga (*Cml4*) 57%, roti *Nat8*-ga 64% ja kannuskonna *NAT8* homoloogi *XCML*-iga 33% (joonis 5).

XCML geenil arvatakse olevat kannuskonna embrüonaalses arengus oluline funktsioon (Popsueva *et al.* 2001).

```

14
1. MAPCHIRKYQESDRQWVGLLSRGMAEHAPATFRQLLKLPRTLILLGGPLALLVSGSW 60
2. MAPCHIRKYQESDRQWVGLLSRGMAEHAPATFRLLKLPRTLILLGGPLALLVSGSW 60
3. MASFRIROFQERDYKQVVDVFSRGMEHPTAFRHLITLPRTLILLAVVPLAIVLVSGSW 60
4. MASFHIRQFQERDYEQVDMFSRGMKEHPTAFRHLITLPRTLILLGVPALVIVLVSGSW 60
5. MANVSIKRYKTSDYETARFLFAEGTKEHLPACMYTLITPRFYFITFVAFTSVFMGTGSY 60

73
1. LLALVFSISLFPALWFLAKKPWTEYVDMTLCIDMSDITKSYLSERGSFCFWVAESEKVVG 120
2. LLALVFSISLFPALWFLAKKPWTEYVDMTLCIDMSDITKSYLSEHGSCFWVAESEKVVG 120
3. FLAVVCIFFLFLFLWFLASKPWKNYVSKCLHTDMADITKSYLSVRGSCFWVAESGGQVVG 120
4. LLAVVCIFFLFPFLWFLAGQPWKNYVSKCLHTDMADITKSYLSDRGSCFWVAESGGQIVG 120
5. VLALTSLVALLAAGWYGLYSEFHGLASRFLRKDMLDIEKSYMMSENACFWVAEIDGKVVG 120

121      130      143
1. MVGALPVD DPTLREKRLQLFHLFVDSEHRRQGI AKALVRTVLQFARDQGYSEVILDTGTI 180
2. MVGALPVD DPTLREKRLQLFHLFVDSEHRRQGI AKALVRTVLQFARDQGYSEVILDTGTI 180
3. TVAARPVKDPPLGRKQLQLFRLSVSSQHRGQGI AKALVRTVLQFARDQGYSDVILVTGLL 180
4. TVGALPVKDPSPGRKQLQLFRLSVSSQHRGQGI AKALVRTVLQFARDQGYTDVILVTGLL 180
5. TVGAQPSTD---ADDELLLQRISVARDYRQLRI GTKLCQTVILDFARQRCFNAVCLLETANI 177

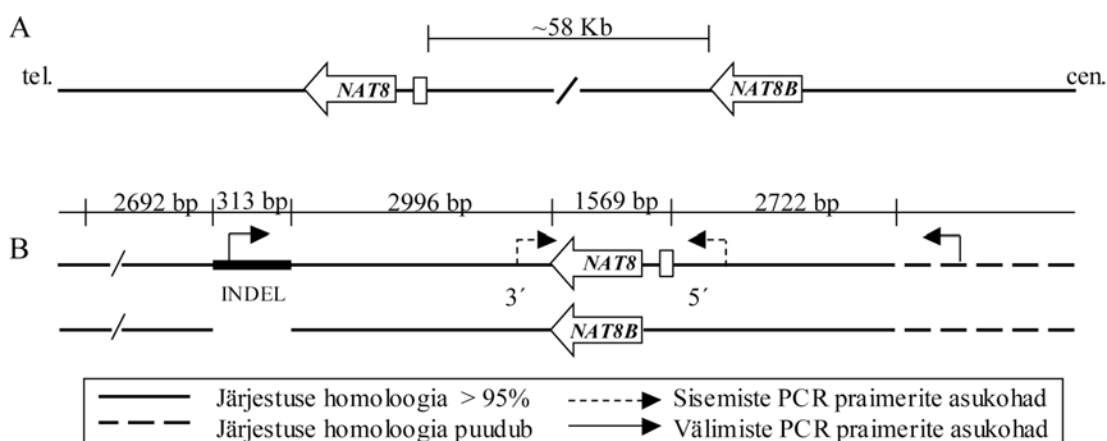
1. QLSAMALYQSMGFKKT-GQSFFCVWARLVALHTVHF IYHLPSSKVGSL 227
2. QLSAMALYQSMGFKKT-GQSFFCVWARLVALHTVHF IYHLPSSKVGSL 227
3. QQGAVTLYYSMGFQKT-GESEFVDILTWLVDVSLIHF IYPLPSAQKYEL 227
4. QQGAVTLYYSMGFQKT-GESEFMDILTWLVDVSLIHF IYPLPSS----- 222
5. QRAATNLYYERVGFKKSRVEILPSLVHQYTSFTVAYRYN IKS----- 219

```

Joonis 5. NAT8 geeni aminohappeline konserveeritus inimesel (1), šimpansil (2), hiirel (3), rotil (4) ja kannuskonnal (5). Mustade kastidega on esile toodud identsed aminohapped NAT8 suhtes. Numbritega on joondamise kohal tähistatud resekveneerimisel tuvastatud aminohappe muutuseid põhjustavad SNP-d (14 Arg-His, 73 Ala-Val, 121 Met-Thr, 130 Pro-Pro, 143 Phe-Ser; vt. ka tabel 9)

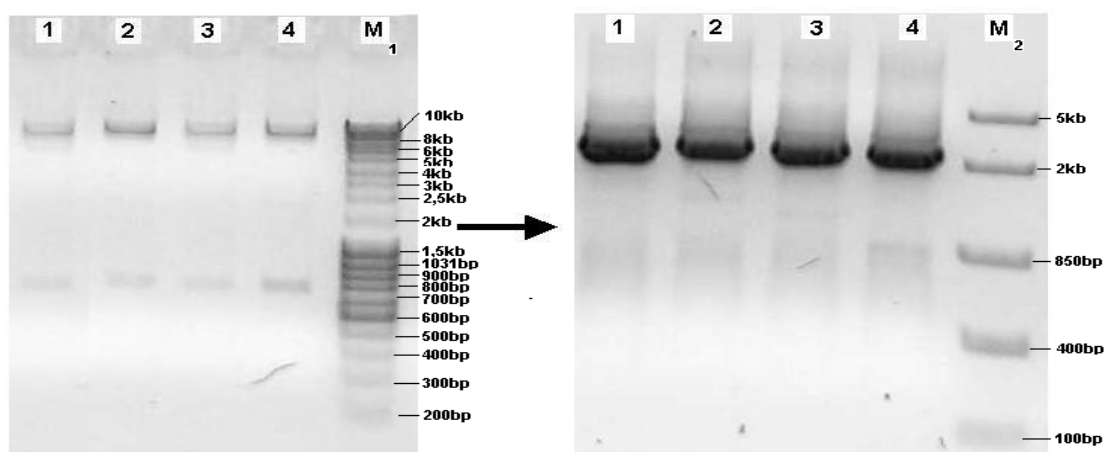
5.4 NAT8 genoomse piirkonna detailsem iseloomustus ja spetsiifiline amplifikatsioon

Järjestuste joondamised *NAT8* ja tema läheduses paikneva duplikaatgeeni *NAT8B* (joonis 6 A) ning nende külgnevate alade vahel näitasid kõrget sarnasust. Rohkem kui 90% järjestuse homoloogia esines nii geenide kodeerivate järjestuste vahel kui ka 2722 bp ulatuses geenidest ülesvoolu ning 2996 bp ulatuses geenidest allavoolu. Seetõttu puudus võimalus paigutada *NAT8* spetsiifilised PCR praimerid geeni vahetusse lähedusse ning need disainiti geenist eemale. Välimine “*forward*” PCR praimer paigutati geenist 3’ allavoolu suunas tuvastatud 313 bp pikkuse indel-i sisse ning välimine “*reverse*” PCR praimer paigutati geeni 5’ ülesvoolu piirkonda, kus *NAT8* ja *NAT8B* järjestuste homoloogia lõppes (joonis 6 B). *NAT8* ja *NAT8B* ning nende külgnevate genoomsete piirkondade joondamine on toodud lisas 1.



Joonis 6. A. *NAT8* ja duplikaatgeeni *NAT8B* paigutus inimese genoomis. B. *NAT8*, *NAT8B* ning nende külgnevate alade nukleotiidsete järjestuste joondamised ja PCR praimerite suhtelised asukohad.

Saadud *NAT8* spetsiifilisest 10341 bp pikkusest produktist amplifitseeriti sisemiste PCR praimeritega resekvenerimiseks soovitud pikkusega ala. Olles läbi viinud PCR reaktsioonid (punkt 4.1.), kontrolliti produktide olemasolu ja pikkust geelelektrofooresil, millest tehtud foto (joonis 7) kinnitas soovitud (10341 bp ja 2322 bp) suurustega produktide olemasolu. Pikemat fragmenti amplifitseerides oli tulemuseks ka 800bp suurune kõrvalprodukt, kuid selle olemasolu töö edasist käiku ei häirinud.

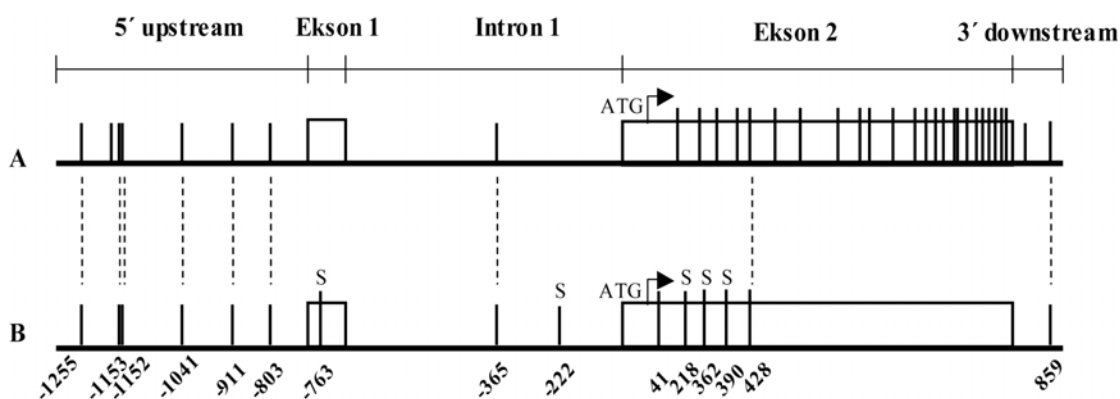


Joonis 7. PCR produktide olemasolu ja pikkuse kontrollimine geelelektrofooresil. 1-4 DNA proovid, M₁– MBI Fermentase marker #SM0403, M₂ – MBI Fermentase marker #SM1118.

5.5 NAT8 geeni resekveneerimine ja polümorfne varieeruvus

Huvipakkuva geeni pikkuseks oli 1569bp, millele lisati 541 bp mRNA alguspunktist ülesvoolu ja 75 bp mRNA lõpust allavoolu ning kogu piirkond (2186 bp) resekveneeriti täielikult mõlemalt ahelalt Eesti populatsiooni esindaval juhuvalimil (n=42). Kokku tuvastati 15 ühenukleotiidsset polümorfismi (joonis 8), mis teeb 2185 bp pikkusel alal keskmiselt 1 SNP 146 bp kohta. Leitud SNP-dest 10 olid harvad (MAF <10%), 5 sagedased ning 8 SNP-d olid varasemalt kirjeldatud dbSNP andmebaasis (dbSNP - <http://www.ncbi.nlm.nih.gov>). Viiest sagedast SNP-st 4 paikneb potentsiaalses promootorpiirkonnas, mille arvatav päritolu on seotud geenikonversiooniga NAT8 ja NAT8B geeni vahel. Eksonis paiknevatest SNP-dest 4 põhjustavad mittesünonüümse muutuse aminohappelises järjestuses (tabel 8). Kuna kõik aminohappelised muutused paigutuvad piirkondades, mis ei ole evolutsiooniliselt konserveerunud, siis seetõttu neid edasisse uuringusse ei kaasatud (joonis 5).

NAT8 resekveneerimisel tuvastatud SNP-de hulk (15) erines oluliselt andmebaasis sama piirkonna kohta olemasolevast informatsioonist, pakkudes sekveneeritud ala kohta 34 SNP-d. Geeni tegelikult palju väiksem varieeruvus viitab suure tõenäosusega andmebaasis olevatele paraloogsetele järjestuste variantidele (PSV – *paralogous sequence variants*), mis ei ole SNP-d vaid kajastavad NAT8 ja NAT8B järjestuste erinevusi (joonis 8).



Joonis 8. A. NAT8 SNP-de paigutus vastavalt dbSNP andmebaasile. B. Resekveneerimisel tuvastatud SNP-d. Punktiiriga on ühendatud kattuvad SNP-d ja S – üksikud “singleton” positsioonid.

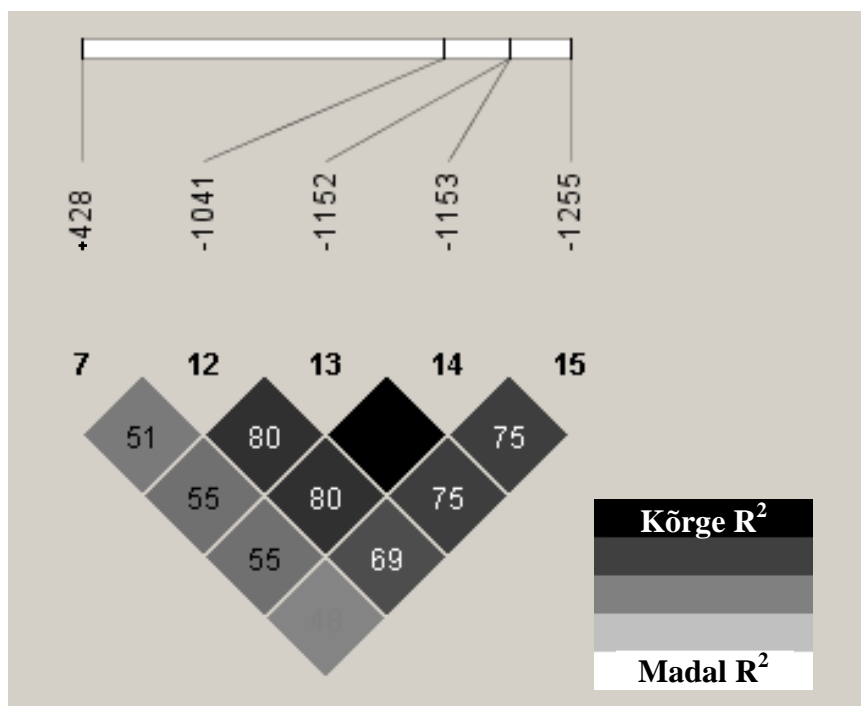
Tabel 8. *NAT8* genoomse piirkonna resekvenerimisel leitud SNP-d

Region	Asukoht ¹	dbSNP rs#	SNP ²	Koodoni muutus ⁴	MAF ³ n=42
5' ülesvoolu	-1255	rs4852954	T/C	-	0.25
	-1153	rs4852953	A/G	-	0.226
	-1152	rs4852952	C/T	-	0.226
	-1041	rs4852951	G/C	-	0.238
	-911	rs2280506	T/C	-	0.083
	-803	rs2280507	C/T	-	0.071
Ekson1	-763	-	T/C	-	0.012 (S)
Intron1	-365	rs6744273	A/G	-	0.071
	-222	-	T/G	-	0.012 (S)
Ekson2	41	-	C/T	Arg-His 14	0.024
	218	-	G/A	Ala-Val 73	0.012 (S)
	362	-	A/G	Met-Thr 121	0.012 (S)
	390	-	G/T	Pro-Pro 130	0.012 (S)
	428	rs13538	A/G	Phe-Ser 143	0.167
3' allavoolu	859	rs1879662	G/A	-	0.071

¹Asukoht ATG suhtes; ²Mazoorne/minoorne alleel; ³Minoorse alleeli sagedus; ⁴Aminohappe muutus ja positsioon (vt. ka joonis 5); S – “*singleton*” positsioon.

5.6 NAT8 alleelse assotsiatsiooni muster

Leitud viie sagedase SNP-ga (minoorse alleeli sagedus >10%) viidi läbi alleelse assotsiatsiooni analüüs, mille nullhüpoteesiks on, et kahe uuritava markeri alleelid päranduvad sõltumatult (Weiss and Terwilliger 2000). Potentsiaalses promootorpiirkonnas paiknevad SNP-d on omavahel aheldunud ($r^2 \geq 0,5$), markerite -1152 ja -1153 vaheline assotsiatsioon osutus täielikuks ($r^2=1$). Geeni sees LD-d ei tuvastatud, sest valdavalt on tegemist harvade SNP-dega, ainsa geenis paikneva ja sagedase (MAF=0,167) SNP-na on nõrgalt aheldunud +428 ($r^2 \geq 0,5$) (joonis 9).



Joonis 9. Alleelse aheldatuse muster NAT8 geenis leitud viie sagedase SNP-ga (MAF > 10%). R^2 väärtused alates "0"(valge) kuni "1"(must).

5.7 NAT8 nukleotiidne varieeruvus

NAT8 geenil teostati nukleotiidse varieeruvuse analüüsid (tabel 9), mis näitasid kogu geenis (1569 bp; $\pi=0,00037$) ja mRNA-s (953 bp; $\pi=0,00045$) madalamat varieeruvust võrreldes 74 kirjanduses viidatud geeni keskmise ($\pi=0,0008$) varieeruvusega (Crawford *et al.* 2004). Negatiivsed Tajima D väärtused geeni (-1,58) ja mRNA (-1,49) kohta kinnitavad harvade SNP-de ülekaalu ning viitavad puhastavale valikule. Erinevalt vähestele SNP-dele ja madalale nukleotiidide mitmekesisusele geenis oli potentsiaalses promootoralas (-500 bp kuni +100 bp alates mRNA alguspunktist) näha tavalisest kõrgem varieeruvus, ulatudes umbes 10 korda ($\pi=0,0032$) kõrgemale kui NAT8 geeni piirkonnas.

Tabel 9. Nukleotiidide mitmekesisus NAT8-l ja teistel uritud hüpertensiooni kandidaatgeenidel

Parameeter	Geen	mRNA	Promootori piirkond ²	
		NAT8	Hüp. geenid ¹	
Regiooni pikkus (bp)	1568	953	600	600
π	0,00037	0,00045	0,00320	0,00040
θ	0.00102	0.00127	0.00221	0,00028
D ^T	-1,58	-1,49	1,05	-0,74...2,02

¹29 eelnevalt resekveneeritud hüpertensiooni kandidaatgeenide keskmine

²Region -500 bp kuni +100 bp mRNA alguspunkti suhtes; π - keskmine nukleotiidide erinevus aluspaari kohta.; θ - polümorfsete positsioonide arvul baseeruv oodatav heterosügootsus aluspaari kohta; D^T - Tajima D väärtus.

5.8 Kahekümne üheksa (29) hüpertensiooni kandidaatgeeni promootorala resekveneerimine ja nukleotiidiline mitmekesisus*

Uurimaks kuivõrd tüüpiline on vererõhu regulatsioonis osalevate geenide promootoralade kõrge varieeruvus, teostati võrdlev resekveneerimine 29 varem teadaoleva hüpertensiooni kandidaatgeeni promootoraladele (-500 bp kuni +100 bp alates mRNA alguspunktist) ning võrreldi saadud tulemusi *NAT8* geeni vastavate andmetega. Promootorpiirkonnad resekveneeriti geenidel: *ACE*, *ADD1*, *ADD2*, *AGT*, *AGTR1*, *AGTR2*, *KCNJ1*, *CYP17A1*, *NPR1*, *CLCNKA*, *HSD11B1*, *CYP11B2*, *CLCNKB*, *KLK1*, *NPR2*, *BSND*, *NR3C2*, *CYP21A2*, *SAH*, *SGK*, *SLC12A3*, *SCNNIA*, *IL1A*, *SLC14A2*, *SLC22A2*, *REN*, *ATP1A*, *SCNN1B* ja *SLC8A1*. Kokku tuvastati uuritavate geenide promootoralas 23 SNP-d, mis jagunesid 13 geeni vahel. 16 geeni promootorpiirkonnas SNP-d puudusid, mis oli ka uuritud 29 hüpertensiooni kandidaatgeeni madala keskmise nukleotiidse varieeruvuse ($\pi=0,0004$; $\min=0$, $\max=0,0017$) põhjuseks. Suurema varieeruva promootoralaga geenideks olid *IL1A* ($\pi=0,0017$) ja *CYP11B2* ($\pi=0,0016$), mis on samuti duplitseerunud geenid. Saadud tulemus näitab *NAT8* ebatavaliselt kõrget (~ 8 X kõrgem) varieeruvust promootorpiirkonnas võrreldes teiste hüpertensiooni kandidaatgeenidega ($\pi=0,00320$ vs. $\pi=0,0004$; tabel 9), kuid sarnast mustrit kahe teise uuritud duplitseerunud geeniga.

*Ühistööna (Katrín Kepp ja Peeter Juhanson, publitseerimata andmed).

5.9 NAT8 SNP-de assotsiatsioonianalüüs vererõhu, vere seerumi biomarkerite ja GF-iga

Kuna komplekshaiguste geneetilist tausta ei moodusta mitte ainult üksikud aminohappe muutused, vaid lisaks ka potentsiaalsetes reguleerivates alades paiknevad geeni ekspressiooni mõjutavad allelid, siis teostati vererõhu ja vere seerumi biomarkerite assotsiatsioonianalüüs NAT8 geeni promootorala SNP-dega (-803, -911, -1041, -1152, -1153 ja -1255).

Vererõhu assotsiatsioonianalüüsis osalenud inimesed (n=137) olid pikaajalised, vererõhuravimeid mittetarbivad terved veredoonorid, kelle vererõhunäitudest (keskmiselt 4 näitu) arvutatud mediaanväärtustega analüüs läbi viidi. Assotsiatsioonianalüüsi käigus arvestati nii indiviidi sugu kui ka vanust vererõhunäitude fikseerimisel. Kasutades lineaarse regressiooni meetodit ja aditiivset geneetilist mudelit tuvastati süstoolse vererõhu seos omavahel täielikult aheldunud NAT8 SNP-dega -803 (C/T) ja -911 (T/C) ($p=0,015$) (tabel 11). Antud testi kohta programmi PLINK poolt näidatud väljundfailis oli mõlema SNP regressiooni koefitsient negatiivne, viidates SNP-de -803 ja -911 minoorse alleeli potentsiaalsele protektiivsele efektile vererõhu kõrgenemise suhtes. Statistiliselt olulist seost süstoolse vererõhu ning SNP-de -803 ja -911 vahel kinnitas ka LRT (*likelihood ratio test*) test ($p=0,028$). Assotsiatsiooni andnud SNP-de minoorse alleeli efekti kontrolliti Mann-Whitney U testiga ($p=0,04$), kus võrreldi SNP-de -803 ja -911 mazoorse alleelide homosügootide (keskmine 135,5 mmHg) ja heterosügootide (mediaan 130,5 mmHg) süstoolsete vererõhkude väärtuste jaotuvust (joonis 10 A). Assotsiatsioonianalüüsid ülejäänud 5' ülesvoolu piirkonna SNP-dega statistiliselt olulisi tulemusi ei andnud, samuti ei nähtud mingit seost diastoolse vererõhuga (tabel 10).

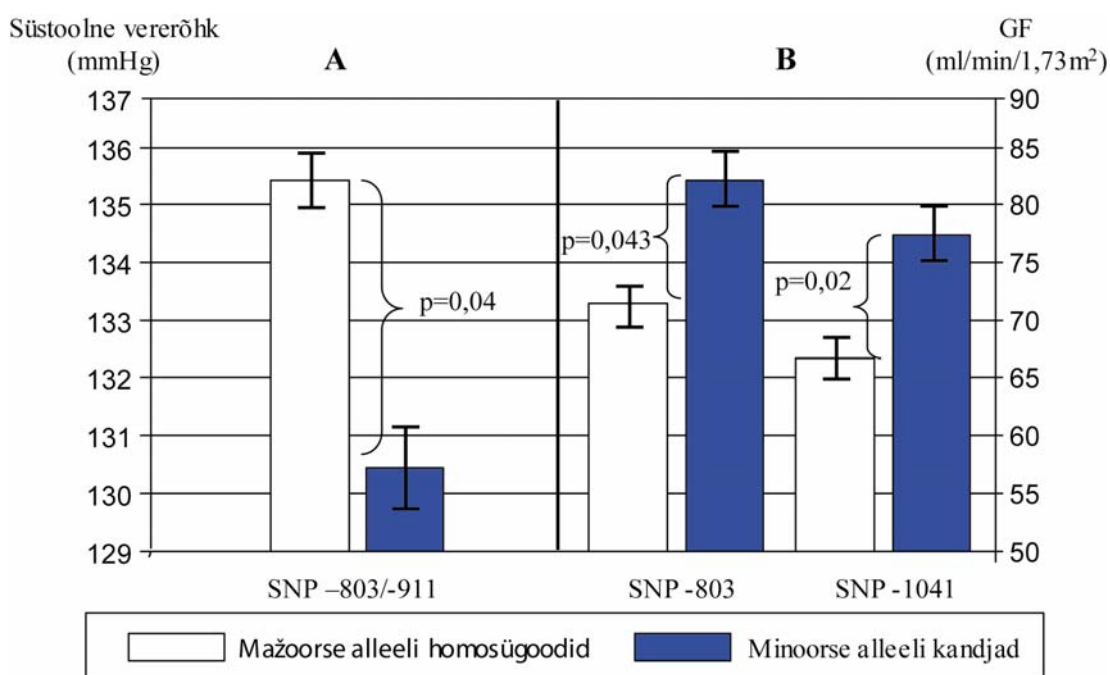
Vere seerumi biomarkerite ja GF-i assotsiatsioonianalüüsis NAT8 5' ülesvoolu paiknevate SNP-dega osalesid patsiendid, kellel oli diagnoositud essentsiaalne hüpertensioon ning patsiendid, kellele oli diagnoositud lisaks essentsiaalsele

hüpertensioonile ka hüpertensiivne nefropaatia. Sarnaselt vererõhu assotsiatsioonianalüüsile sooritati ka biomarkerite ja GF-i analüüsil samad statistilised testid, mis näitasid SNP-de -803 ($p=0,042$) ja -1041 ($p=0,05$) assotsiatsiooni GF-iga (tabel 10) ning nende SNP-de minoorsete alleelide protektiivset efekti neerufunktsiooni häirumisele kõrgenenud vererõhu tagajärjel. SNP -803 (C/T) mazoorse alleeli homosügootide GF-i arvatud keskmine oli 72 ning minoorse alleeli kandjatel 82 (Mann-Whitney $p=0,043$). SNP-1041 (G/C) mazoorse alleeli homosügootide GF-i keskmine oli 66 ning minoorse alleeli kandjatel vastavalt 77 (Mann-Whitney $p=0,02$) (joonis 10 B). Seerumi biomarkerite (HDL/LDL-kolesterool ja triglütseriidid) ja *NAT8* 5' ülesvoolu piirkonnas paiknevate SNP-de vahel assotsiatsiooni ei tuvastatud.

Tabel 10. Assotsiatsioonitestide p väärtused (aditiivne geneetiline mudel)

SNP ID	Vererõhk		Seerumi biomarkerid			GF
	SYS (n=137)	DIA (n=137)	HDL-Chol (n=149)	LDL-Chol (n=149)	Trigl (n=146)	
-803 (C/T)	0,015	ns	ns	ns	ns	0,042
-911 (T/C)	0,015	ns	ns	ns	ns	<i>0,060</i>
-1041 (G/C)	ns	ns	ns	ns	ns	0,050
-1152 (C/T)	ns	ns	ns	ns	ns	<i>0,081</i>
-1153 (A/G)	ns	ns	ns	ns	ns	<i>0,081</i>
-1255 (T/C)	ns	ns	ns	ns	ns	<i>0,070</i>

SYS – süstoolne vererõhk; DIA – diastoolne vererõhk; HDL-Chol – HDL-kolesterool; LDL-Chol – LDL kolesterool; Trigl – triglütseriidid; GF – glomerulaarne filtratsioon; rasvases kirjas on esitatud statistiliselt olulised tulemused $p<0,05$; kursiivis statistiliselt piiripealsed tulemused $0,05<p<0,1$; ns – statistiliselt mitte oluline $p>0,1$.



Joonis 10. A. Süstoolse vererõhu keskmised väärtused (koos standardveaga) aheldunud SNP-de -803 ja -911 mažoorse alleeli homosügootide ja minoorse alleeli kandjatel. B. GF-i keskmised väärtused (koos standardveaga) SNP-de -803 ja -1041 mažoorse alleeli homosügootide ja minoorse alleeli kandjatel. Joonisel toodud Mann-Whitney U testi p väärtused.

4.10 *NAT8* promootorala SNP alleelide potentsiaalne mõju geeniekspressioonile

Selgitamaks *NAT8* 5' ülesvoolu piirkonnas paiknevate SNP-de potentsiaalset mõju geeniekspressioonile, ennustati programmiga MatInspector potentsiaalseid transkriptsioonifaktorite seondumissaite (TFSS) analüüsitud 5' regioonis. TFSS-ide ennustamiseks sooritati kaks päringut kahe alternatiivse 5' ülesvoolu järjestusega, kus esimeses päringus polümorfed positsioonid DNA järjestuses olid asendatud vastavalt mažoorsete alleelidega ning teises päringus minoorsete alleelidega. SNP-de -911, -1152, ja -1153 minoorsete alleelide olemasolu *NAT8* 5' ülesvoolu regiooni nukleotiidses järjestuses põhjustas mitmete TFSS-ide modifitseerumise. Tabelis 11 on esitatud muutused promootorala potentsiaalsetes TFSS-ides vastavalt eelpool nimetatud SNP alleelidele. Valdavalt jäi SNP-de minoorsete alleelide esinemise

korral TFSS-e vähemaks, juurde tekkis vaid PSE (*Proximal sequence element of RNA polymerase II-transcribed snRNA genes*) (tabel 11). Muutused TFSS-ides võivad viidata NAT8 mažoorse ja minoorse promootorvariandi funktsionaalsusele transkriptsiooni initsiatsioonil ja regulatsioonil.

Tabel 11. NAT8 5'ülesvoolu piirkonna SNP alleelide mõju potentsiaalsetele transkriptsioonifaktorite seondumissaitidele

SNP ID	Lühend	Olemasolu mažoorse alleeli korral	Olemasolu minoorse alleeli korral	Faktori nimetus
-1152 ja -1153	2 X SRF	+	-	<i>Serum response factor</i>
	PSE	-	+	<i>Proximal sequence element of RNA polymerase II-transcribed snRNA genes</i>
	MZF1	+	-	<i>Myeloid zinc finger protein</i>
-911	ERR	+	-	<i>Estrogen related receptor</i>
	RORA	+	-	<i>RAR-related orphan receptor alpha1</i>
	c-Rel	+	-	<i>c-Rel</i>
	HBP1	+	-	<i>HMG box-containing protein 1</i>
	NF-kB	+	-	<i>Nuclear factor kappa B</i>
	MOK-2	+	-	<i>Ribonucleoprotein associated zinc finger protein</i>

6. ARUTELU

6.1 Andmebaasides olemasolevate SNP-de usaldusväärsus

Antud töös resekveneeritud piirkonna kohta leidis andmebaasis dbSNP 34 SNP-d, neist ainult 15 tuvastati resekveneerimise teel Eesti populatsiooni juhuvalimis (n=42). Andmebaasi informatsiooni kohaselt peaks *NAT8* geenis olema väga suur SNP-de tihedus just teises eksonis, mida resekveneeritud indiviidile ei täheldatud (joonis 8). Seega vaadates probleemi laiemalt tekib küsimus, kui võrd usaldusväärsed on andmebaasides olevad duplitseerunud piirkondades paiknevad SNP-d ja mida nad tegelikult tähendavad? Selgub, et selliseid SNP-sid on võimalik tõlgendada mitmeti:

1. Tegemist võib olla konkreetse geeni polümorfse positsiooniga ehk reaalse SNP-ga.
2. Tegemist võib olla PSV-ga, mis kajastab lihtsalt kõrgelt sarnaste genoomsete piirkondade omavahelist erinevust. Seda kinnitavad ka andmed, et kui võrd umbes 90% kaardistatud refSNP-dest omavad inimese genoomis unikaalset positsiooni, siis umbes 5% on kaardistatud kahte asukohta ja teine 5% esinevad rohkem kui kolmes kohas (Estivill *et al.* 2002). PSV-dele viitab ka andmebaasis olemasolev rohkem kui kaks korda tihedam SNP-de hulk duplitseerunud piirkondades (Fredman *et al.* 2004)
3. Tegemist on reaalse SNP-ga, mis paigutub positsioonis, kus kaks kõrgelt sarnast järjestust omavahel erinevad (MSV – *multisite variation*).

Andmebaasis kirjeldatud ning resekveneerimise tulemustel leitud SNP-de arvu suure erinevuse võib põhjustada võimalus, et andmebaasis võivad olla kirjeldatud väga madala sagedusega esinevad ja/või populatsiooni-spetsiifilised SNP-d. Kuna antud töö raames resekveneeriti suhteliselt väike hulk indiviide ning ainult Eesti populatsioonist, siis ei tohiks seda võimalust täielikult välistada.

Tõenäolisema põhjenduse paljudele andmebaasis olevatele *NAT8* SNP-dele saab aga välja tuua *NAT8* ja tema sugulasgeeni *NAT8B* järjestuste omavahelisest joondamisest (lisa 1). Ilmneb, et 27 andmebaasis olevat SNP-d ehk 79% selle geeni kohta esitatud SNP-dest langeb kokku erinevustega kahe duplitseerunud piirkonna järjestuses, mis

võib tähendada, et andmebaasis kirjeldatud SNP-de puhul võib osaliselt olla tegu hoopis paraloogsete järjestuste variantidega (PSV-d), mis on sisestatud andmebaasi tänu ebakorrektselt järjestuste assambleerimisele. Resekveneerimisel tuvastati ka 7 MSV-d, mille puhul oli tegemist tegelike SNP-dega, mis paiknesid positsioonides, kus *NAT8* ja *NAT8B* järjestused omavahel erinesid. Selliste SNP-de päritolu on tõenäoliselt tingitud geenikonversioonist *NAT8* (aktseptor) ja *NAT8B* (donor) geenide vahel. Geenikonversiooni toimumise eelduseks on DNA järjestuste kõrge sarnasus, mis iseloomustab just duplitseerunud genoomseid regioone (Bailey *et al.* 2002).

Fülogeneetiliste analüüside põhjal arvatakse, et *NAT8/NAT8B* duplikatsioon leidis aset pärast Vana Maailma ahvide lahknemist. Arvatakse, et geeni mõlemad koopiad on endiselt aktiivsed näiteks šimpansil, kuid inimese *NAT8B* geenis on leitud kaks enneaegset stopkoodoneid põhjustavat polümorfset muutust, mis tõenäoliselt viivad nende geenivariantide puhul mitte-funktsionaalse valgutekkimisele (Hahn and Lee 2006).

6.2 *NAT8* kui võimalik hüpertensiooni kandidaatgeen

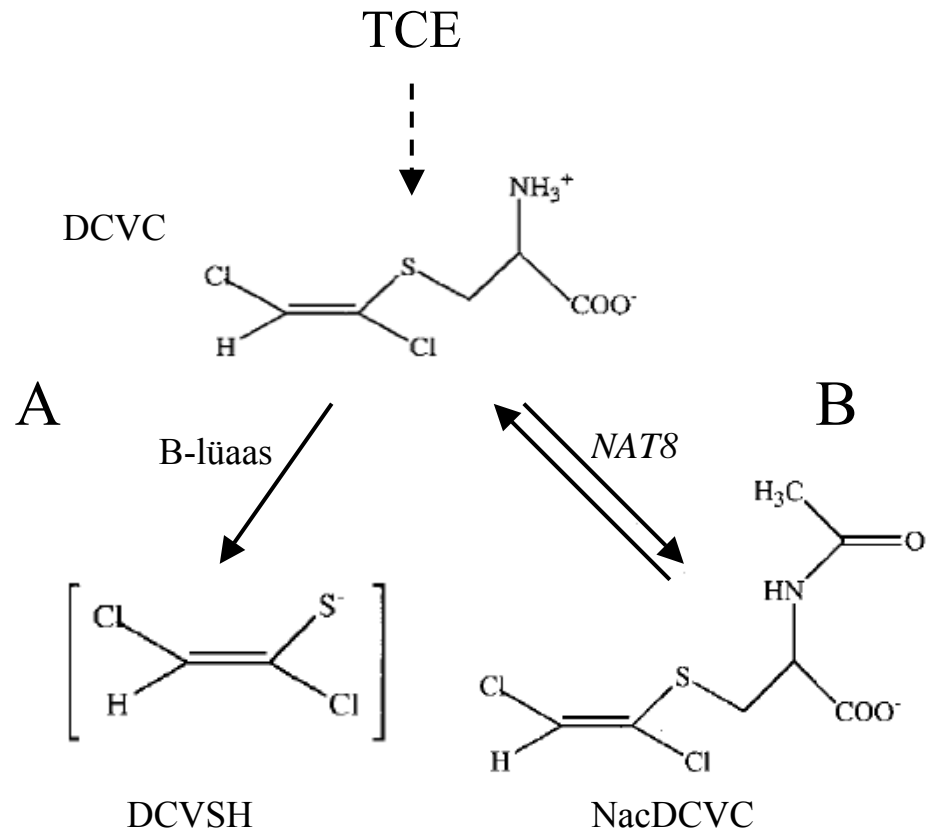
NAT8 valiti uuringusse kui hüpertensiooni kandidaatgeen lähtudes tema ekspressioonist nii areneva kui ka täiskasvanud inimese neerus ning asukohast “kuumas” 2p piirkonnas, mida varasema aheldusanalüüsi uuringu tulemusena on seostatud hüpertensiooniga (2p 96-115cM) (Rice *et al.* 2000). Samuti on piirkonnale 2p12 (*NAT8* lokalisatsioon 2p12-p13.1) viidanud võrdleval genoomikal põhinenud uuring, kus hüpertensiooniga seotud roti QTL regioonid kõrvutati inimese kromosomaalsete regioonidega (Stoll *et al.* 2000).

Tänapäevani on informatsioon *NAT8* täpse funktsiooni kohta puudulik, kuid siiski on näidatud, et *NAT8* poolt kodeeritud valk omab märkimisväärset sarnasust mitmete bakteriaalsete atsetüültransferaasidega, mis on seotud ravimiresistentsusega (Ozaki *et al.* 1998). *NAT8* spetsiifiline ja tugev ekspressioon inimese terves neerus ja maksas

viitab tema tähtsusele tagamaks normaalset neerude ja maksa embrüonaalset arengut ning hilisemat talitlust (Ozaki *et al.* 1998). *NAT8* homoloog kannuskonnal (geen *XCML*) on seotud embrüonaalse arenguga (Popsueva *et al.* 2001) ning tema võrdlus *NAT8*-ga (joonis 5) näitas konserveerunud domäänide olemasolu, mis jällegi kinnitab *NAT8* võimalikku osalust inimese embrüonaalses arengus. Evolutsiooniliselt konserveerunud funktsioonile viitab ka resekveneerimisel leitud vähene SNP-de arv eksonis.

Tänaseks on teada mitmeid vererõhu regulatsiooniga seotud mendeliaarseid sündroome (Gong and Hubner 2006), mis on seotud neeru filtreeriva osakese nefroniga, mille ülesandeks on organismi soola – vedeliku tasakaalu tagamine. Sellest lähtuvalt on tõenäoline, et ka paljud essentsiaalse hüpertensiooni kui komplekshaiguse kujunemisega seotud geenid toimivad läbi neerude metaboolsete radade.

NAT8 seost neerudes toimivate detoksifikatsiooni protsessidega on näidatud uuringus, kus vaadeldi kemikaali TCE (trikloroetüleen) metaboliidi S-(1,2-diklorovinüül)-L-tsüsteiini (DCVC) mõju koekultuuris kasvatatud neeru rakkudele. Vaatamata sellele, kus TCE lagundamise esmased protsessid aset leiavad, transporditakse vaheühend DCVC edasiseks metabolismiks neerudesse, kus on kaks võimalikku alternatiivset rada (joonis 11). A) DCVC aktivatsioon DCVSH-ks, mida loetakse nefrotoksiliseks ühendiks. B) DCVC atsetüleeritakse ning moodustub ühend (Nac-DCVC), mis väljutatakse organismist uriiniga. Rakkude 10 päevane eksponeerimine DCVG-le põhjusas *NAT8* geeni 2 kordse ekspressiooni langemise (Lock *et al.* 2006). Ühe hüpoteesi kohaselt võivad *NAT8* promootor piirkonna mazoorsed SNP alleelid olla tundlikud toksiliste ühendite suhtes (TCE metaboliidid), geeni ekspressiooni pärsitakse ning eelistatud on metaboolse raja kasutamine (joonis 11, A), mille tulemusena tekib nefrotoksiline vaheühend. *NAT8* promootor piirkonna SNP-de minoorsete alleelide korral säilitatakse geeni aktiivsus ning tagtakse TCE metaboliidi väljutamine uriiniga (joonis 11 B).



Joonis 11. TCE (trikloroetüleen) metaboliidi DCVC (S-(1,2-diklorovinül)-L-tsüsteiin) kaks võimalikku rada neerus (Cummings *et al.* 2000). A – nefrotoksilise vaheühendi süntees; B – atsetüleerimine N-atsetüültransferaasi toimel.

6.3 NAT8 potentsiaalse promootorpiirkonna SNP-d

Seoses komplekshaigustega on hakatud tähelepanu pöörama just nõ. “vaiksetele” genoomsetele muutustele, mis ei lokaliseeru geeni kodeerivas alas, kuid samas on võimalised mõjutama geeni ekspressiooni, splaissingut või normaalset funktsiooni (Pagani and Baralle 2004). Inimese ja hiire genoomsete järjestuste joondamisest on

selgunud, et umbes 80% konserveerunud elementidest paigutuvad mittekodeerivates piirkondades (Pennacchio 2003), mis viitab nende alade olulisele funktsioonile ning seal aset leidvate muutuste tähtsusele geeniekspressiooni seisukohalt.

Antud töös pöörati tähelepanu *NAT8* potentsiaalsele promootorpiirkonnale, kus tuvastati keskmisest 8 korda suurem varieeruvus võrreldes teiste hüpertensiooni kandidaatgeenidega. Kõrge SNP-de arv promootorpiirkonnas võib aga omada regulatoorseid tähtsust näiteks transkriptsioonifaktorite seondumissaitide modifitseerimises, mis lõppkokkuvõttes võib mõjutada geeni ekspressiooni.

Sooritatud transkriptsioonifaktorite seondumissaitide ennustus tuvastas mitmete potentsiaalsete seondumissaitide kadumise promootorpiirkonnas juhul kui sisestatud järjestuses esinesid polümorfsete positsioonide minoorseid alleelid (tabel 11). Kõrvuti paiknevate SNP-de –1152 (C/T) ja –1153 (A/G) minoorsete alleelide esinemise korral kadus ära kaks *serum response* faktori (SRF) seondumissaiti. SRF on laialdaselt ekspresseeritud TF, mis on seotud paljude raku kasvu ja differentseerumisega seotud geenide regulatsioonis (Chai and Tarnawski 2002). SRF-i seondumissaitide kadumise protektiivset efekti kõrge vererõhu suhtes on näidatud hiirtel veresoonte seinte silelihasrakkudes paikneva müosiini kerget ahelat kodeeriva geeni (*smMLCK*) promootoriga. Võrreldes *smMLCK* geeni promootoreid tavalisel ja SHR rotiliinil, tuvastati viimasel 12 bp pikkune insertioon, mis soodustas SRF-i seondumist ning ühtlasi ka geeni ekspressiooni. SRF-i seondumissaidi blokeerimisel vähenes *smMLCK* geeni ekspressioon ja ühtlasi ka SHR rotiliini vererõhk (Han *et al.* 2006).

SNP –911 (T/C) minoorse alleeli esinemine potentsiaalses promootorala järjestuses põhjustas kokku mitmete ennustatud TF-i seondumissaidi kadumise (tabel 11). Muuhulgas kadusid sellised potentsiaalsed TFSS-id nagu ERR (*estrogen related* retseptor), RORA (*RAR-related orphan* retseptor alfa) ja NF-κB (tuuma faktor kappa B). TF RORA on seotud imetajate tsirkaadse rütmi regulatsiooniga ning tema ekspressiooni on täheldatud ajus, tüümuses, skeletilihastes, nahas, südames, maksas ja ka neerudes. RORA on seotud ka lipiidide metabolismiga reguleerides kolesterooli taset läbi apoA-I geeni transkriptsiooni ja NF-κB aktiivsuse (Boukhtouche *et al.*

2004). NF-kB omakorda on tähtis TF, mille kontrolli alla kuuluvad mitmed põletiku ja immuunvastusega seotud geenid. Häired NF-kB aktivatsioonis on näidatud artriidi, astma, arteroskleroosi, glomerulonefriidi ja AIDS-i puhul (Chen *et al.* 1999).. Transkriptsioonifaktorite näol on tegemist äärmiselt oluliste geeni ekpressiooni mõjutavate molekulidega, mis sageli läbi keeruliste toimemehanismide on haigusfenotüübi põhjustajateks. Kuivõrd hüpertensiooni kui multifaktoriaalse haiguse kujunemises osalevad palju geenid, siis on tõenäoline, et tähtis roll on ka väga erinevatel transkriptsioonifaktoritel.

KOKKUVÕTE

Käesoleva töö kirjanduse osas anti ülevaade komplekshaigustest essentsiaalse hüpertensiooni näitel. Tutvustati komplekshaiguste kaardistamise strateegiaid ning toodi välja suuremad seni teostatud essentsiaalse hüpertensiooni aheldus- ja assotsiatsiooniuuringud ning nende tulemused.

Käesoleva töö eksperimentaalses osas otsiti uusi kandidaatgeene aheldusanalüüsi tulemusel lokaliseeritud genoomsetest piirkondadest ning valiti välja uus võimalik hüpertensiooni kandidaatgeen *NAT8*, mis resekveneeriti 42-l Eesti populatsiooni indiviidil. Resekvneerimise tulemusena saadi uut informatsiooni *NAT8* geeni polümorfse ja nukleotiidse varieeruvuse kohta. Erinevalt andmebaasides olemasolevast informatsioonist tuvastati *NAT8* geenis vähe SNP-sid, mis võib viidata andmebaasis esinevatele PSV-dele. Vastupidiselt *NAT8* geenis esinevale vähesele varieeruvusele tuvastati geeni potentsiaalses promootoralaas mitmeid sagedasi SNP-sid, millega viidi läbi assotsiatsiooniuuringud kasutades vererõhunäite ja vere seerumi biomarkereid. Assotsiatsiooniuuringute tulemusena tuvastati *NAT8* geeni potentsiaalse promootorala SNP-de seos süstoolse vererõhu ja neeru funktsiooni iseloomustva glomerulaarfiltratsiooniga.

Töö viimases osas vaadeldi *NAT8* promootorala SNP-de alleelide potentsiaalset mõju geeniekspressioonile ning leiti mitmeid muutuseid transkriptsioonifaktorite seondumissaitides.

SUMMARY

Essential hypertension (EH), or steadily high blood pressure (BP), is a complex disease diagnosed when blood pressure systolic and diastolic readings constantly exceed normal values (>140 and/or >90 mmHg). Kidneys have an important function in BP regulation. It is suggested that defective embryonic kidney development and function is a predisposing factor for EH. Novel potential BP regulating renal genes were screened from human genomic regions identified as responsible for hypertension by linkage analyses. Chromosomal region 2p13.1-p12 harbors kidney-expressed *NAT8* (*N-acetyltransferase 8*) and its duplicate pseudogene *NAT8B*. *NAT8* is highly expressed in normal human embryonic kidney and liver, and suggested to play an important role in development and maintenance of kidney and liver structure and function.

We resequenced the entire human *NAT8* gene for random Estonian samples to obtain exhaustive polymorphism coverage of the gene and to estimate the population genetic parameters. The identified *NAT8* polymorphism pattern is different from public database records. The lower actual number of SNPs indicates that most of the *NAT8* SNPs present in databases are actually paralogous sequence variants with its duplicate gene *NAT8B*.

Based on the identified *NAT8* SNP distribution we focused on the variation of the 5'upstream region (putative promoter) and conducted an association study for *NAT8* polymorphisms with blood pressure measurements and serum biomarkers. Significant associations ($p < 0.05$) were detected between *NAT8* SNPs for systolic BP and glomerular filtration rate (GFR). There was no association for diastolic BP or lipid metabolism (LDL/HDL-cholesterol, triglycerides).

Finally, *in silico* analysis was performed to predict the possible functional impact of alternative alleles of 5'upstream SNPs on *NAT8* expression.

TÄNUAVALDUSED

Täna oma juhendajat Maris Laant suurepärase ideede ning abi eest käesoleva magistr töö valmimisel. Samuti täna kaasjuhendajat Mai Rosenbergi koos teiste TÜ Kliinikumi, Põhja-Eesti Regionaalhaigla ning Põhja-Eesti ja TÜ Verekeskuse arstidega, kes olid abiks uuritavate indiviidide valikul ning kõiki uuringus osalenud inimesi. Meeldiva töökeskonna ning abi eest soovin tänada kõiki labori inimesi.

KIRJANDUSE VIITED

- Allayee, H., T. W. de Bruin, et al. (2001). "Genome scan for blood pressure in Dutch dyslipidemic families reveals linkage to a locus on chromosome 4p." Hypertension **38**(4): 773-8.
- Angius, A., E. Petretto, et al. (2002). "A new essential hypertension susceptibility locus on chromosome 2p24-p25, detected by genomewide search." Am J Hum Genet **71**(4): 893-905.
- Ardlie, K. G., L. Kruglyak, et al. (2002). "Patterns of linkage disequilibrium in the human genome." Nat Rev Genet **3**(4): 299-309.
- Atwood, L. D., P. B. Samollow, et al. (2001). "Genome-wide linkage analysis of blood pressure in Mexican Americans." Genet Epidemiol **20**(3): 373-82.
- Bailey, J. A., Z. Gu, et al. (2002). "Recent segmental duplications in the human genome." Science **297**(5583): 1003-7.
- Barbalic, M., T. Skaric-Juric, et al. (2006). "Gene polymorphisms of the renin-angiotensin system and early development of hypertension." Am J Hypertens **19**(8): 837-42.
- Barrett, J. C., B. Fry, et al. (2005). "Haploview: analysis and visualization of LD and haplotype maps." Bioinformatics **21**(2): 263-5.
- Bianchi, G., E. Niutta, et al. (1989). "A possible primary role for the kidney in essential hypertension." Am J Hypertens **2**(2 Pt 2): 2S-6S.
- Bianchi, G., G. Tripodi, et al. (2004). "Na⁺, kidney, hypertension and genes: lessons from rats." J Hypertens **22**(8): 1461-4.
- Binder, A. (2006). "Identification of genes for a complex trait: examples from hypertension." Curr Pharm Biotechnol **7**(1): 1-13.
- Botstein, D. and N. Risch (2003). "Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease." Nat Genet **33** Suppl: 228-37.
- Boukhtouche, F., J. Mariani, et al. (2004). "The "CholesteROR" protective pathway in the vascular system." Arterioscler Thromb Vasc Biol **24**(4): 637-43.
- Brand, E., L. Bankir, et al. (1999). "Glucagon receptor gene mutation (Gly40Ser) in human essential hypertension: the PEGASE study." Hypertension **34**(1): 15-7.
- Brenner, B. M. and S. Anderson (1992). "The interrelationships among filtration surface area, blood pressure, and chronic renal disease." J Cardiovasc Pharmacol **19** Suppl 6: S1-7.
- Brenner, B. M., D. L. Garcia, et al. (1988). "Glomeruli and blood pressure. Less of one, more the other?" Am J Hypertens **1**(4 Pt 1): 335-47.
- Brenner, B. M. and H. S. Mackenzie (1997). "Nephron mass as a risk factor for progression of renal disease." Kidney Int Suppl **63**: S124-7.
- Cambien, F., O. Poirier, et al. (1999). "Sequence diversity in 36 candidate genes for cardiovascular disorders." Am J Hum Genet **65**(1): 183-91.
- Cardon, L. R. and J. I. Bell (2001). "Association study designs for complex diseases." Nat Rev Genet **2**(2): 91-9.

- Caulfield, M., P. Munroe, et al. (2003). "Genome-wide mapping of human loci for essential hypertension." Lancet **361**(9375): 2118-23.
- Chai, J. and A. S. Tarnawski (2002). "Serum response factor: discovery, biochemistry, biological roles and implications for tissue injury healing." J Physiol Pharmacol **53**(2): 147-57.
- Chen, F., V. Castranova, et al. (1999). "New insights into the role of nuclear factor-kappaB, a ubiquitous transcription factor in the initiation of diseases." Clin Chem **45**(1): 7-17.
- Cheng, L. S., R. C. Davis, et al. (2001). "Coincident linkage of fasting plasma insulin and blood pressure to chromosome 7q in hypertensive hispanic families." Circulation **104**(11): 1255-60.
- Collins, A., C. Lonjou, et al. (1999). "Genetic epidemiology of single-nucleotide polymorphisms." Proc Natl Acad Sci U S A **96**(26): 15173-7.
- Collins, F. S., M. S. Guyer, et al. (1997). "Variations on a theme: cataloging human DNA sequence variation." Science **278**(5343): 1580-1.
- Cooper, R. S., A. Luke, et al. (2002). "Genome scan among Nigerians linking blood pressure to chromosomes 2, 3, and 19." Hypertension **40**(5): 629-33.
- Crawford, D. C., T. Bhangale, et al. (2004). "Evidence for substantial fine-scale variation in recombination rates across the human genome." Nat Genet **36**(7): 700-6.
- Cummings, B. S., R. C. Zangar, et al. (2000). "Cytotoxicity of trichloroethylene and S-(1, 2-dichlorovinyl)-L-cysteine in primary cultures of rat renal proximal tubular and distal tubular cells." Toxicology **150**(1-3): 83-98.
- Dahl, L. K., M. Heine, et al. (1962). "Role of genetic factors in susceptibility to experimental hypertension due to chronic excess salt ingestion." Nature **194**: 480-2.
- Dahl, L. K., M. Heine, et al. (1974). "Genetic influence of the kidneys on blood pressure. Evidence from chronic renal homografts in rats with opposite predispositions to hypertension." Circ Res **40**(4): 94-101.
- ESH (2003) "2003 European Society of Hypertension-European Society of Cardiology guidelines for the management of arterial hypertension." J Hypertens **21**(6): 1011-53.
- Estivill, X., J. Cheung, et al. (2002). "Chromosomal regions containing high-density and ambiguously mapped putative single nucleotide polymorphisms (SNPs) correlate with segmental duplications in the human genome." Hum Mol Genet **11**(17): 1987-95.
- Fredman, D., S. J. White, et al. (2004). "Complex SNP-related sequence variation in segmental genome duplications." Nat Genet **36**(8): 861-6.
- Geller, D. S., A. Farhi, et al. (2000). "Activating mineralocorticoid receptor mutation in hypertension exacerbated by pregnancy." Science **289**(5476): 119-23.
- Gong, M. and N. Hubner (2006). "Molecular genetics of human hypertension." Clin Sci (Lond) **110**(3): 315-26.
- Gordon, D., C. Desmarais, et al. (2001). "Automated finishing with autofinish." Genome Res **11**(4): 614-25.

- Hahn, Y. and B. Lee (2006). "Human-specific nonsense mutations identified by genome sequence comparisons." *Hum Genet* **119**(1-2): 169-78.
- Han, Y. J., W. Y. Hu, et al. (2006). "Increased myosin light chain kinase expression in hypertension: Regulation by serum response factor via an insertion mutation in the promoter." *Mol Biol Cell* **17**(9): 4039-50.
- Harrap, S. B., Z. Y. Wong, et al. (2002). "Blood pressure QTLs identified by genome-wide linkage analysis and dependence on associated phenotypes." *Physiol Genomics* **8**(2): 99-105.
- Herrmann, S. M., V. Nicaud, et al. (2002). "Polymorphisms of the beta2 - adrenoceptor (ADRB2) gene and essential hypertension: the ECTIM and PEGASE studies." *J Hypertens* **20**(2): 229-35.
- Hilbert, P., K. Lindpaintner, et al. (1991). "Chromosomal mapping of two genetic loci associated with blood-pressure regulation in hereditary hypertensive rats." *Nature* **353**(6344): 521-9.
- Hirschhorn, J. N. and M. J. Daly (2005). "Genome-wide association studies for common diseases and complex traits." *Nat Rev Genet* **6**(2): 95-108.
- Hsueh, W. C., B. D. Mitchell, et al. (2000). "QTL influencing blood pressure maps to the region of PPH1 on chromosome 2q31-34 in Old Order Amish." *Circulation* **101**(24): 2810-6.
- Hunt, S. C., R. C. Ellison, et al. (2002). "Genome scans for blood pressure and hypertension: the National Heart, Lung, and Blood Institute Family Heart Study." *Hypertension* **40**(1): 1-6.
- Jacob, H. J., K. Lindpaintner, et al. (1991). "Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat." *Cell* **67**(1): 213-24.
- Kaplan, N. M. (1993). "Essential workup of the hypertensive patient." *Heart Dis Stroke* **2**(2): 104-8.
- Klahr, S. (1989). "The kidney in hypertension--villain and victim." *N Engl J Med* **320**(11): 731-3.
- Kristjansson, K., A. Manolescu, et al. (2002). "Linkage of essential hypertension to chromosome 18q." *Hypertension* **39**(6): 1044-9.
- Kruglyak, L. (1999). "Prospects for whole-genome linkage disequilibrium mapping of common disease genes." *Nat Genet* **22**(2): 139-44.
- Krushkal, J., R. Ferrell, et al. (1999). "Genome-wide linkage analyses of systolic blood pressure using highly discordant siblings." *Circulation* **99**(11): 1407-10.
- Kumar, S. and S. B. Hedges (1998). "A molecular timescale for vertebrate evolution." *Nature* **392**(6679): 917-20.
- Lander, E. and L. Kruglyak (1995). "Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results." *Nat Genet* **11**(3): 241-7.
- Lander, E. S. and N. J. Schork (1994). "Genetic dissection of complex traits." *Science* **265**(5181): 2037-48.
- Levey, A. S., J. P. Bosch, et al. (1999). "A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation.

- Modification of Diet in Renal Disease Study Group." Ann Intern Med **130**(6): 461-70.
- Levy, D., A. L. DeStefano, et al. (2000). "Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the framingham heart study." Hypertension **36**(4): 477-83.
- Lifton, R. P., R. G. Dluhy, et al. (1992). "A chimaeric 11 beta-hydroxylase/aldosterone synthase gene causes glucocorticoid-remediable aldosteronism and human hypertension." Nature **355**(6357): 262-5.
- Lifton, R. P., A. G. Gharavi, et al. (2001). "Molecular mechanisms of human hypertension." Cell **104**(4): 545-56.
- Lock, E. A., J. L. Barth, et al. (2006). "Changes in gene expression in human renal proximal tubule cells exposed to low concentrations of S-(1,2-dichlorovinyl)-l-cysteine, a metabolite of trichloroethylene." Toxicol Appl Pharmacol **216**(2): 319-30.
- McBride, W., C. Ferrario, et al. (2003). "Hypertension and medical informatics." J Natl Med Assoc **95**(11): 1048-56.
- Mein, C. A., M. J. Caulfield, et al. (2004). "Genetics of essential hypertension." Hum Mol Genet **13 Spec No 1**: R169-75.
- Mune, T., F. M. Rogerson, et al. (1995). "Human hypertension caused by mutations in the kidney isozyme of 11 beta-hydroxysteroid dehydrogenase." Nat Genet **10**(4): 394-9.
- Ozaki, K., T. Fujiwara, et al. (1998). "Isolation and mapping of a novel human kidney- and liver-specific gene homologous to the bacterial acetyltransferases." J Hum Genet **43**(4): 255-8.
- Pagani, F. and F. E. Baralle (2004). "Genomic variants in exons and introns: identifying the splicing spoilers." Nat Rev Genet **5**(5): 389-96.
- Pankow, J. S., K. M. Rose, et al. (2000). "Possible locus on chromosome 18q influencing postural systolic blood pressure changes." Hypertension **36**(4): 471-6.
- Pennacchio, L. A. (2003). "Insights from human/mouse genome comparisons." Mamm Genome **14**(7): 429-36.
- Poch, E., D. Gonzalez, et al. (2001). "Molecular basis of salt sensitivity in human hypertension. Evaluation of renin-angiotensin-aldosterone system gene polymorphisms." Hypertension **38**(5): 1204-9.
- Popsueva, A. E., N. N. Luchinskaya, et al. (2001). "Overexpression of camello, a member of a novel protein family, reduces blastomere adhesion and inhibits gastrulation in *Xenopus laevis*." Dev Biol **234**(2): 483-96.
- Rapp, J. P. (2000). "Genetic analysis of inherited hypertension in the rat." Physiol Rev **80**(1): 135-72.
- Rice, T., T. Rankinen, et al. (2002). "Genomewide linkage scan of resting blood pressure: HERITAGE Family Study. Health, Risk Factors, Exercise Training, and Genetics." Hypertension **39**(6): 1037-43.

- Rice, T., T. Rankinen, et al. (2000). "Genome-wide linkage analysis of systolic and diastolic blood pressure: the Quebec Family Study." Circulation **102**(16): 1956-63.
- Rozas, J. and R. Rozas (1999). "DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis." Bioinformatics **15**(2): 174-5.
- Rousset, F. and M. Raymond (1995). "Testing heterozygote excess and deficiency." Genetics **140**(4): 1413-9.
- Samani, N. J. (2003). "Genome scans for hypertension and blood pressure regulation." Am J Hypertens **16**(2): 167-71.
- Sass, C., S. Cheng, et al. (2004). "Genetic influences on blood pressure within the Stanislas Cohort." J Hypertens **22**(2): 297-304.
- Shimkets, R. A., D. G. Warnock, et al. (1994). "Liddle's syndrome: heritable human hypertension caused by mutations in the beta subunit of the epithelial sodium channel." Cell **79**(3): 407-14.
- Snieder, H., G. A. Harshfield, et al. (2003). "Heritability of blood pressure and hemodynamics in African- and European-American youth." Hypertension **41**(6): 1196-201.
- Stoll, M., A. E. Kwitek-Black, et al. (2000). "New target regions for human hypertension via comparative genomics." Genome Res **10**(4): 473-82.
- Zhu, D. L., H. Y. Wang, et al. (2001). "Linkage of hypertension to chromosome 2q14-q23 in Chinese families." J Hypertens **19**(1): 55-61.
- Telgmann, R., E. Brand, et al. (2007). "SAH gene variants are associated with obesity-related hypertension in Caucasians: the PEGASE Study." J Hypertens **25**(3): 557-64.
- Watterson, G. A. (1975). "On the number of segregating sites in genetical models without recombination." Theor Popul Biol **7**(2): 256-76.
- Weiss, K. M. and J. D. Terwilliger (2000). "How many diseases does it take to map a gene with SNPs?" Nat Genet **26**(2): 151-7.
- Wilson, F. H., S. Disse-Nicodeme, et al. (2001). "Human hypertension caused by mutations in WNK kinases." Science **293**(5532): 1107-12.
- Xu, X., J. J. Rogus, et al. (1999). "An extreme-sib-pair genome scan for genes regulating blood pressure." Am J Hum Genet **64**(6): 1694-701.

LISA 1. NAT8 ja NAT8B järjestuste joondus

PCR praimer 2kb F

NAT8_29_10_06 CAGAGGAAGTCCCCAGCCAAGCTCAGG**C**AAAATTGCTGACTT**C**AGACT**C**ATGAACTAAAT 60
 NAT8B_20_10_06 CAGAGGAAGTCCCCAGCCAAGCTCAGG**C**AAAATTGCTGACTT**C**AGACT**C**ATGAGCTAAAT 60

NAT8_29_10_06 AAATATTTGTTTCATTTGGTGGTTGTTCTTATGCAACATTATGCTGGCACTAGAAAAATG 120
 NAT8B_20_10_06 AAATATTGATTTTCATTTTATGATTGTTCTTATGCAACATTATCCTGGCACTAGAAAAATG 120
 ***** ** *****

NAT8_29_10_06 AAACAGATGGGAATCAAAAAGACTGAATGAAGCAGAAACCCAGAAGCTACCCACACAAGCA 180
 NAT8B_20_10_06 AAACAGATGGGAATCAAAAAGACTGAATGAAGCAGAAACCCAGAAGCTACCCACAGGCA 180

NAT8_29_10_06 ATCACATGGGACTCAGATGTGAATGGACAATAAGCTTTTATTGCA**T**TGAAAGGTCATTGC 240
 NAT8 B20_10_06 ATCGCATGGGACTCAGATGTTAATGGACAATAAGCTTTTATTGCA**C**AGAAAGGTCATTGC 240
 *** *****

NAT8_29_10_06 AGT**G**AAGGTTGGGGATTGCTTGCTGCTACAG**C**TGAATGGATTCTATTCTGACCAATACA 300
 NAT8B 20_10_06 AGT**C**AGAGGTTGGGGATTGCTTGCTGCTACAG**C**CGAATGGATTCTATTCTGACCAATACA 300
 *** *****

NAT8_29_10_06 CACAGAAAG**G**ATCA**C**AGACT**T**CCCT**A**CC**T**TAG**A**AGAAGGGAGG**T**C**G**TAGATGAAATGAAC 360
 NAT8B 20_10_06 GAAGGAAAG**A**ATC**A**TAGAG**C**CC**T**GC**C**TGAG**C**AGAAGGGAGG**T**C**A**TAGATGAAATGAAC 360
 * *****

NAT8_29_10_06 TGTATGAAG**G**CCACT**T**AGCCTGGCCACAC**C**AGAAGAAGGACTGGCCCGTCTTCTTGAA 420
 NAT8B 20_10_06 TGTATGAAG**A**CC**C**AGCCTGGCCACAC**C**TGAAGAAGGACTGGCCCGTCTTCTTGAA 420

NAT8_29_10_06 GCCC**T**CTCTGGTAGAG**G**CCATAGCAGAGAGCTGGAT**C**GT**C**CGGTGTCCAGGATAAC 480
 NAT8B 20_10_06 GCCC**A**CTCTGGTAGAG**C**CCATGGCAGAGAGCTGGAT**C**TTG**C**TGGTGTCCAGGACAAC 480

NAT8_29_10_06 TTCACTGTAGCCCTGGTCCCGGGCAA**A**CTGGAGGACAGTCCTGACCAGGGCTTTTGTCTAT 540
 NAT8B 20_10_06 TTCACTGTAGCCCTAGTCCCGGGCAA**A**CTGGAGGACAGTCCTGACCAGGGCTTTTGTCTAT 540

NAT8_29_10_06 CCCCTGACGACGGTGCTCACTGTCCAC**A**AGAGATGAAACAGCTGCAACCGCTTCTCCCT 600
 NAT8B_20_10_06 CCCCTGACCAGGTGCTCA**T**TGTCCAC**A**GAGAGATGAAACAGCTGCAACCGCTTCTCCCT 600

NAT8_29_10_06 CAAGGTGGGATCATCAACAGGCAGAGCTCCTACCATGCCACCACCTTCTCTTCAGACTC 660
 NAT8B_20_10_06 CAAGGTGGGATCATCAACGGGCAGAGCTCCTACTGTGCCACCACCTTCTCTTCAGATTCC 660

NAT8_29_10_06 AGCCACCCAGAAGCAGGAGCCACGCTCACTCAGGTAGGATTTGGTAATGTCAGACATGTC 720
 NAT8B_20_10_06 ACCCACCCAGAAGCAGGAGCCACACTCACTCAGGTAGGATTTGGTATGTCAGACATGTC 720
 * *****

NAT8_29_10_06 TGTGCACAATGTCATGTCCACATACTCCGTCCAGGGTTTTTTGGCAAGGAACACAGGGC 780
 NAT8B_20_10_06 TGTGCGCAATGCTATGTCTACATAACCGCTCCAGGGTTTTTTGGCAAGGAACACAGGGC 780

NAT8_29_10_06 AGGGAAGAGGCTGATGCTGAACACGAGGGCTAGAAGCCAGGATCCAGAGACCAGGAGTAG 840
 NAT8B_20_10_06 AGGAAGGAGGCTGAGGCTGAACACGAGGGCCAGAATCCAGGAGCCAGAGACCAGGAGTAG 840
 *** * *****

NAT8_29_10_06 GCGAGGGGCCCCCAAGTAAGAGTATGAGGGTTCGAGGCAGCTTCAGCAATTGCCGGAA 900
 NAT8B_20_10_06 GGCAAGGGCCCCCAAGTAAGAGTATGAGGGTTCGAGGCAGCTTCAGTAATCGCCGGAA 900
 *** **

NAT8_29_10_06 GGTGGCTGGGGCATGCTCGCCATCCCCGGGAGAGCAAGCCCACAACCCACTGGCGGTC 960
 NAT8B_20_10_06 GGTGGCTGGGGCGTGTTCGGCCATCCCCGGGAGAGCAAGCCCACGACCTACTTGCGGTC 960
 ***** **

NAT8_29_10_06 GCTCTCCTGGTATTTGCGGATGTGACAAGGAGCCATGGACAGACTTCTGTGTCTGAAATC 1020
 NAT8B_20_10_06 GCTCTCCTGGTATTTGCGGATGTGATAAGGAGCCATGGACAGACTTCTGTGTCTGAAATC 1020
 ***** **

NAT8_29_10_06 TCTAAGTCCAGGAGACAGAGCTAGGCCTCCCTGCACGCCTTTACGTCAGAC-TGGGGAAAG 1079
 NAT8B_20_10_06 ACTGAGTCCAGGAAACAGAGCTAGGCCTCCCTGCACGCCTTTACGCCAGGCCTGGGGAAAG 1080
 ** *****

NAT8_29_10_06 AGAGAGAAAATCATGTGAGTGCCTGCATGGCTCCCAGCCTTGCAGGAACAGGGAGACCTG 1139
 NAT8B_20_10_06 AGAGAGGAAACCACGTGAGTGCCTGCATGGCTCCCAGCCTTGCAGGAACAGGGAGACCTG 1140
 ***** **

NAT8_29_10_06 CTCAAGGGTGTGTCTTTCCGCGTTTGCCCATGAGGAAGCAGGCCTCTGCCACTGGGAGAA 1199
 NAT8B_20_10_06 CTTAAGGGTGTGTCTTTCCGCGTTTGCCCATGAGGAAGCAGGCCTCTGCCACTGGGAGAA 1200
 ** *****

NAT8_29_10_06 GGTAGGGTCAGAAAGACCTGCATTGGAATCTGGTTCATCCCTTTCTAGCTGTGTGACCT 1259
 NAT8B_20_10_06 GGTAGCATGAGAAAGACCTGCATTGGAATCTGGTTCATGCCTTTCTAGCTGTGTGACCT 1260
 ***** *

NAT8_29_10_06 TAGGCATGTCACCTAACCTCTCTGAGCCTCCATTTCTTCAACTATAGAACAGGGGAGAAG 1319
 NAT8B_20_10_06 TGGGCATGTCACCTAACCTCTCTGAGCCTCCACTTCTTATATAGAACAGGGGAGAAG 1320
 * *****

NAT8_29_10_06 AAAACCTGCCTTCTAGAGTTGTTTTGAAGATTGAATAAGATG C C C C T T C T A T A A A G C C C T G G 1379
 NAT8B_20_10_06 AAAACCTGCCTTCTAGAGTTGTTTTGAAGATTGAATAAAAATAC C C C T T C T A T A A A G C C C T G A 1380
 ***** **

NAT8_29_10_06 TAGATGAACTACCCACCGCTCTTCCCTTTTCTGTTTCCTTTTCTGGCTCTCTCAGCTTC 1439
 NAT8B_20_10_06 TAGATGAACTACCCACCACTCTTCC---CTCTGTTTCCTTTTCTGGCTCTCTCAGCTTC 1437

NAT8_29_10_06 TCCGTGCATTGAGGGGCACCTACCTCAGGCAGGCCACACGTGCAACCCTTTTGTCTTTC 1499
 NAT8B_20_10_06 TTGGTGCATTGAGAGAGTACCTACCTCAGGCAGGCCACATGCCAACCCTTTGGTGCTTC 1497
 * *****

NAT8_29_10_06 TCCAGGGACCATAACATCGCAAAAATTCTCTGTCCAA-CACAAGTTTCAGGAAAGGCTCCC 1558
 NAT8B_20_10_06 TCCAGGGACCATAACATCGCAGAAATCCTGTGACCAAGCACAAGTTTCAGGAAAGGCTCCC 1557

NAT8_29_10_06 TCCAGCTTTATTGTAAGTTTTTAGGTTGCTGAGGTGGCATAAGAGCCAAGCATAAAAAAC 1618
 NAT8B_20_10_06 TCCAGCTTTATTGTAAGTTTTTAGGTTGCTGAGGTGGCATAAGTGCCAAGCATAAAAAAC 1617

NAT8_29_10_06 AGCCCAGCCTCCAGATCCCAAGGAAGGTGCATGCTCAGAGCACTGTCTTCCACCCAGTC 1678
 NAT8B_20_10_06 AGCCCAGCCTCCAGATCCCAAGGAAGGTGCATGCTCAGAGCACTGTCTTCCACCCAGTC 1677

NAT8_29_10_06 CCCTCACCTGTCCACCACAAGAGCCTTGAGTGTCCAGGGAGCTTCAGCTCTCAGCCGCTCA 1738
 NAT8B_20_10_06 CCCTCACCTGTCCACCACAAGAGCCTTGAGTGTCCAGGGAGCTTCAGCTGTCCAGCCGCTCA 1737

```

NAT8_29_10_06      CTGGCATCCAGGAGATACCGTCTGGGGCTTGCTCACCCCTGGAAAGTAGACTGCCTGCCT 1798
NAT8B_20_10_06     CTGGCATCCAGGAGATACCGTCTGGGGCTTGCTCACCCCTGGAAAGCAGGCTGCCTGCCT 1797
*****

NAT8_29_10_06      TATTGCTTGGCTCCTGAATGTTTGGTAATCATCAACCCTAGGGTCAAAGAGCTGGGCTGT 1858
NAT8B_20_10_06     CATTGCTTGGCTCCTGAATGTTTGGTAATCATCAACCCTAGGGTCAAAGAGCTGGGCTGT 1857
*****

NAT8_29_10_06      AACACAGCAGGGAGCCCAGTGTGCATTCCCTCTTTGTGGTCAAAGGGAACTGCACCTGGAGGG 1918
NAT8B_20_10_06     AACACAGCAGGGAGCCCAGTGTGCATTCCCTCTTTATAGTCAAAGGGAACTGGCCTGGAGGG 1917
*****

NAT8_29_10_06      GGAGACTCAGGGTCTGGAACCGCCTGGGTTCCCTTGCCCTTCTGGAAGGCCTGCCACCTCT 1978
NAT8B_20_10_06     GGAAACTCAGGGTCTGGAACCGCCTGGGTTCCCTTGCCCTTCTGGAAGGCCTGCCACCTCT 1977
***

NAT8_29_10_06      CTTGCCTGTTTCAGGCCCGTGCATCTCTTGCCACCTGGCACACTTAGTACTCTCGGTCC 2038
NAT8B_20_10_06     CTTGCCTGTTTCAGGCCCTGGGCATCTTTCTGCCACCTGGCACACTTGATACTCTC----- 2032
*****

NAT8_29_10_06      CTGTCCACATATACGATCCACCCTGGTCCCTTGTACACATATACGACGAGCCTGAGAACTG 2098
NAT8B_20_10_06     -----TGAGATCCACCTTGGTCCCTGTACACATATACCACGAGCCTGAGAACTG 2081
*

NAT8_29_10_06      TACCATTACCCTGCAATCAGTCACCATTGGCCAGAAGCCGAGTCCCCACTGCTTGGCAAC 2158
NAT8B_20_10_06     TGCCATTACCCTGCACTCAGTCACCATTGGCCAGAAGCCGAGTCCCCACTGCTTGGCAAC 2141
*

NAT8_29_10_06      AAGGCCCTTCCCAGCTAGGCTTCTCCTCGCTGAGGTTTCCCTGGTCCACGAAGCATCTGT 2218
NAT8B_20_10_06     AAGGCCCTTCCCAGCTAGGCTTCTCCATGCTGAGGTTTCCCTGGTCCACAAAGCATCGAT 2201
*****

NAT8_29_10_06      TGTGAGCAGTTCCCTTGCACCTCACCTACCAACCCCCATCTCCACCCATTGATTGATTGAG 2278
NAT8B_20_10_06     TGTGAGCAGTTCTTTGCACCTCACCTACCAACCCCCACCTCCACCCATTGATTGATTGAG 2261
*****

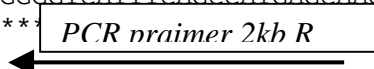
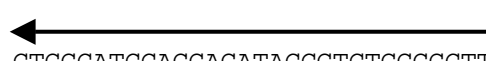
NAT8_29_10_06      TCAACAGTTGTGTGTCATGCACTTCCCTGTGTCTGGGGTCATTTTCCAGCCATGAGCAAGGT 2338
NAT8B_20_10_06     TCAACAGGTGTGTGCGATGCACTTCCCATGTCTGGGGTCATTTTCCAGCCATGAGCAAGGT 2321
*****

NAT8_29_10_06      TGACATTCGGCTTGACAGGCATGGCTGTGTGTGGCTTGTGTGTGTTCTGGCTGTTGGTCA 2398
NAT8B_20_10_06     TGACATTCGGCT-GACAGGCATGGCTGTGTGTGGCTTGTGTGTGTTCTGGCTGTTGGTCA 2380
*****

NAT8_29_10_06      GAGGCTGTGCCACACCCGTTGTAAATATTTTGCACCTGCAGATGTTGGGGAGGCCCAAAT 2458
NAT8B_20_10_06     GAGGCTGTGCCACACCCGTTGTAAATATTTTGCACCTGCAGATGTTGGAGAGGCCCAAAT 2440
*****

NAT8_29_10_06      TCATGCTCTT
NAT8B_20_10_06     TCAAGCTCTT
***

```



mRNA on skeemil märgitud peenema joonega ja transleeritav piirkond jämedama joonega; kastidesse on paigutatud andmebaasis olemasolevad SNP-d, mis langevad kokku NAT8 ja NAT8B vaheliste järjestuste erinevustega.