Dizajn panela za analizu gena BRCA1 i BRCA2 i priprema biblioteke za sekvenciranje

Sajfert, Martina

Master's thesis / Diplomski rad

2018

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:478643

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-29**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Biološki odsjek

Martina Sajfert

Dizajn panela za analizu gena *BRCA1* i *BRCA2* i priprema biblioteke za sekvenciranje

Diplomski rad

Zagreb, 2018.

Ovaj rad je izrađen u Laboratoriju za naprednu genomiku Instituta Ruđer Bošković, pod vodstvom dr. sc. Olivera Vugreka. Rad je predan na ocjenu Biološkom odsjeku Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu radi stjecanja zvanja magistra molekularne biologije.

Zahvaljujem se dr. sc. Oliveru Vugreku na ukazanoj prilici za izradu diplomskog rada i stjecanje novog znanja kao i svim djelatnicima laboratorija na ugodnom društvu i poticajnoj radnoj atmosferi. Posebna zahvala mag. chem. Filipu Rokiću na stručnim savjetima, trudu i strpljenju tijekom izrade i pisanja diplomskog rada.

Također, velika hvala izv. prof. dr. sc. Petri Korać na suvoditeljstvu, savjetima i konstruktivnim kritikama.

Naročito se zahvaljujem svojim roditeljima i bratu, suprugu Filipu i bliskim prijateljima na potpori, razumijevanju i poticanju kroz sve godine obrazovanja.

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Biološki odsjek Diplomski rad

DIZAJN PANELA ZA ANALIZU GENA *BRCA1* I *BRCA2* I PRIPREMA BIBLIOTEKE ZA SEKVENCIRANJE

Martina Sajfert

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Nasljedni sindrom karcinoma dojke i jajnika uzrokovan je nasljednim mutacijama u genima *BRCA1* i *BRCA2*. Nositelji mutacija imaju 80% veći rizik razvoja karcinoma dojke te 50% veći rizik razvoja karcinoma jajnika. Geni *BRCA1/2* su tumor-supresorski geni uključeni u stanične putove važne za kontrolu oštećenja DNA i ključni su faktor pri izbjegavanju nastanka mutacija. U svrhu kliničkog genskog testiranja, mutacije u genima *BRCA1/2* uglavnom se određuju Sangerovim sekvenciranjem. Metodama sekvenciranja sljedeće generacije (NGS) značajno se smanjuje cijena i vrijeme potrebno za sekvenciranje DNA te se omogućuje visoka osjetljivost i specifičnost uz upotrebu znatno manje količine početnog uzorka. Cilj ovog istraživanja bila je samostalna konstrukcija panela za umnažanje fragmenata gena *BRCA1/2* te optimizacija uvjeta umnažanja kodirajućih regija za navedene gene u svrhu pripreme biblioteke za NGS. Konstruirane su početnice za umnažanje kodirajućih regija navedenih gena i optimizirani su uvjeti njihova umnažanja višestrukim PCR-om. Na temelju odabranih uvjeta umnažanja, na uzorku genomske DNA iz krvi i tkiva uklopljenog u parafin pripremljene su i sekvencirane biblioteke. Rezultati su obrađeni bioinformatičkim programskim paketima i uspoređeni s komercijalnim testom za analizu gena *BRCA1/2*. Analizom dobivenih rezultata, potvrđena je uspješna konstrukcija genskog panela uz potrebu dodatne optimizacije u svrhu unaprjeđenja detekcije nasljednih mutacija.

(67 stranica, 41 slika, 9 tablica, 64 literaturnih navoda, jezik izvornika: hrvatski jezik)

Rad je pohranjen u Središnjoj biološkoj knjižnici
Ključne riječi: karcinom dojke, *BRCA1*, *BRCA2*, višestruki PCR, genski panel, NGS, ciljano sekvenciranje

Voditelj: Dr. sc. Oliver Vugrek, v. zn. sur.
Suvoditelj: Dr. sc. Petra Korać, izv. prof.
Ocjenitelji: Dr. sc. Petra Korać, izv. prof.
Dr. sc. Ana Galov, izv. prof.
Dr. sc. Željka Vidaković-Cifrek, izv. prof.

Rad prihvaćen: 14.2.2018.

BASIC DOCUMENTATION CARD

University of Zagreb

Faculty of Science

Division of Biology

Graduation Thesis

DESIGN OF *BRCA1* AND *BRCA2* PANEL AND OPTIMIZATION OF LIBRARY PREPARATION FOR NEXT-GENERATION SEQUENCING

Martina Sajfert

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Hereditary breast and ovarian cancer syndrome is caused by germline mutations in *BRCA1* and *BRCA2* genes. Carriers of this mutations have 80% higher risk of developing breast cancer and 50% higher risk of developing ovarian cancer. *BRCA1/2* genes are tumor-suppressor genes involved in cellular pathways which play an important role in repairing DNA damage. In today's clinical genetic testing, mutations in *BRCA1/2* genes are mainly identified by Sanger's sequencing. Next-generation sequencing methods (NGS) significantly reduce the price and time of sequencing and offer much higher sensitivity and specificity for determination of single nucleotide variants (SNVs). The aim of this study was design of *BRCA1*/2 panel for amplification of *BRCA1/2* gene fragments and optimization of their amplification conditions for the purpose of library preparation for NGS. We designed primers for the amplification of both coding regions and performed optimization of multiplex PCR reactions. Based on the selected amplification conditions, DNA samples isolated from blood and paraffin-embedded tissue were used for library preparation and sequencing. The results were analyzed using various bioinformatics tools and compared with the commercial *BRCA1/2* gene test. The analysis of the obtained results confirmed the successful genetic panel construction with the need for further optimization in order to improve detection of mutations.

(67 pages, 41 figures, 9 tables, 64 references, original in: Croatian)

Thesis deposited in the Central Biological Library
Key words: breast cancer, *BRCA1*, *BRCA2*, multiplex PCR, gene panel, NGS, targeted sequencing

Supervisor: Dr. sc. Oliver Vugrek, SRA
Cosupervisor: Dr. sc. Petra Korać, Assoc. Prof.
Reviewers: Dr. sc. Petra Korać, Assoc. Prof.
Dr. sc. Ana Galov, Assoc. Prof.
Dr. sc. Željka Vidaković-Cifrek, Assoc. Prof.

Thesis accepted: 14.2.2018.

Sadržaj

1. Uvod……...1

1.1. Tumori……...1

1.2. Karcinom dojke……..……..3

1.3. Geni *BRCA1* i *BRCA2*……..…...6

1.3.1. Uloga BRCA1 i BRCA2 u popravku molekule DNA…….....8

1.3.2. Uloga BRCA1 i BRCA2 u transkripcijskoj regulaciji…….……..10

1.3.3. Uloga BRCA1 i BRCA2 u kontrolnim točkama staničnog ciklusa……...12

1.3.4. Utjecaj hormona na razvoj karcinoma dojke…….....……...….13

1.4. Sekvenciranje……..……15

1.4.1. Razvoj metoda sekvenciranja…….15

1.4.2. Sekvenciranje sljedeće generacije……..16

1.4.2.1. Sekvenciranje kratkih sekvenci – druga generacija sekvenciranja……17

1.4.2.2. Sekvenciranje dugih sekvenci – treća generacija sekvenciranja……19

1.4.3. Sekvenciranje sparenih krajeva……...…20

1.4.4. Obrada podataka…….…21

1.4.5. Primjena metoda sekvenciranja sljedeće generacije……..21

1.4.6. Ciljano sekvenciranje…….……23

1.5. Gensko testiranje i liječenje karcinoma dojke……...……24

2. Cilj istraživanja…….……26

3. Materijali i metode…….……..27

3.1. Materijali……..……...27

3.2. Metode……...…..27

3.2.1. Konstrukcija početnica……...….27

3.2.2. Izolacija genomske DNA……...28

3.2.3. Određivanje koncentracije DNA u otopinama……...……29

3.2.4. Pojedinačne PCR-reakcije...…….…….…..30

3.2.5. Elektroforeza u agaroznom gelu….……...31

3.2.6. Optimizacija reakcija višestrukog PCR-a……..31

3.2.7. Reakcije višestrukog PCR-a za sekvenciranje sljedeće generacije……33

3.2.8. Pročišćavanje produkata višestrukog PCR-a…….34

3.2.9. Priprema biblioteke za sekvenciranje…….…34

3.2.10. Određivanje veličine fragmenata dobivenih produkata…….…36

3.2.11. Sekvenciranje uzoraka……...…36

3.2.12. Obrada rezultata…….…36

3.2.13. Usporedba s komercijalno dostupnim testovima za analizu mutacija gena *BRCA1* i *BRCA2*……..……..37

4. Rezultati……...…...38

4.1. Izolacija genomske DNA……..38

4.2. Pojedinačne PCR-reakcije.…….……...…..……..…38

4.3. Reakcije višestrukog PCR-a……...……..41

4.3.1. Reakcije višestrukog PCR-a na uzorku genomske DNA iz krvi……41

4.3.2. Reakcije višestrukog PCR-a na uzorku genomske DNA iz FFPE-a…….44

4.4. Rezultati DNA-čipa visoke osjetljivosti pripremljenih biblioteka……..47

4.5. Rezultati sekvenciranja……50

5. Rasprava……..……58

6. Zaključci……...…63

7. Literatura……...…….64

8. Prilozi……...67

9. Životopis

Popis kratica

ATM – kinaza ataksije teleangiektazije, od eng. *ataxia-telangiectasia mutated kinase*

ATR – kinaza povezana s ATM, od eng. *ATM-related kinase*

BARD1 –protein 1 prstenaste domene povezan s BRCA1, od eng. *BRCA1-associated RING domain protein 1*

BRCA1 – gen 1 predispozicije za karcinom dojke, od eng. *breast cancer susceptibility gene 1*

BRCA2 - gen 2 predispozicije za karcinom dojke, od eng. *breast cancer susceptibility gene 2*

ChIP – kromatinska imunoprecipitacija, od eng. *chromatin immunoprecipitation*

CHK2 – kontrolna kinaza G2/M , od eng. *G2/M control kinase*

CRT – ciklička reverzibilna terminacija, od eng. *cyclic reversible termination*

DDR – kinaze uključene u prepoznavanje oštećenja molekule DNA, od eng. *DNA damage response kinases*

DSB – dvolančani lom, od eng. *double-strand breaks*

ER – estrogenski receptor, od eng. *estrogen receptor*

ERE – estrogenski element, od eng. *estrogen response element*

ExAC - konzorcij sakupljenih eksoma, od eng. *The Exome Aggregation Consortium*

FFPE – tumorsko tkivo uklopljeno u parafin, od eng. *formalin-fixed paraffin embedded tissue*

gnomAD – baza sakupljenih genoma, od eng. *The Genome Aggregation Database*

HBOCS – nasljedni sindrom karcinoma dojke i jajnika, od eng. *hereditary breast and ovarian cancer syndrome*

HER2 – receptor ljudskog epidermalnog faktora rasta 2, od eng. *human epidermal growth factor receptor 2*

HGP – projekt ljudskog genoma, od eng. *Human Genome Project*

HR – homologna rekombinacija, od eng. *homologous recombination*

ICGC – Međunarodni projekt genomskog konzorcija, od eng. *International Genome Consortium*

INDEL – insercije i delecije, od eng. *INsertion/DELetion*

LOH – gubitak heterozigotnosti, od eng. *loss of heterozygosity*

NES – jezgrin izlazni signal, od eng. *nuclear export signal*

NGS – sekvenciranje sljedeće generacije, od eng. *next-generation sequencing*

NHEJ – nehomologno spajanje krajeva, od eng. *non-homologous end-joining*

NLS – jezgrin lokalizacijski signal, od eng. *nuclear localization signal*

PE – spareni krajevi, od eng. *paired-end*

PCNA – jezgrin antigen proliferirajućih stanica, od eng. *proliferating cell nuclear antigen*

PCR – lančana reakcija polimerazom, od eng. *polymerase chain reaction*

PR – progesteronski receptor, od eng. *progesteron receptor*

P/CAF – faktor povezan s p300/CBP, od eng. *p300/CBP-associated factor*

SBL – sekvenciranje vezanjem, od eng. *sequencing by ligation*

SBS – sekvenciranje sintezom, od eng. *sequencing by synthesis*

SMRT – sekvenciranje pojedinačne molekule u stvarnom vremenu, od eng. *single-molecule real-time sequencing*

SNA – dodatak jednog nukleotida, od eng. *single-nucleotide addition*

SNV – promjene u jednom nukleotidu, od eng. *single nucleotide variant*

SSA – sparivanje jednolančanih krajeva, od eng. *single-strand annealing*

TCGA – projekt Atlas genoma ljudskih malignih tumora, od eng. *The Cancer Genome Atlas*

WGS – sekvenciranje čitavog genoma, od eng. *whole genome sequencing*

ZMW – valovodi, od eng. *zero-mode waveguides*

1. UVOD

1.1. Tumori

Tumor je naziv za nenormalnu nakupinu tkiva koja može biti benigna i ne širi se ili maligna i ima mogućnost invazije u okolna tkiva i metastaziranja. Maligni tumor predstavlja zloćudni oblik novotvorine i opisuje se kao grupa bolesti s nereguliranim staničnim rastom, invazijom i širenjem stanica od mjesta nastanka malignog tumora, tj. primarnog mjesta do drugih područja u tijelu (Pecorino, 2012). Zbog nastanka sekundarnih tumora, tj. metastaza, teško ih je iskorijeniti operacijama ili lokalnim iradijacijama (Alberts, 2013). Danas je klasificirano više od 100 tipova malignih tumora koji se razlikuju ovisno o tkivu u kojem su nastali: maligni tumor epitelnih stanica – karcinom (u 85% slučajeva malignih tumora), maligni tumor mezodermalnih stanica - sarkom, maligni tumor žljezdanog tkiva - adenokarcinom (Pecorino, 2012), maligni tumor bijelih krvnih stanica i njihovih prekursora (hematopoetskih stanica) - leukemija ili limfom te maligni tumor stanica živčanog sustava (Alberts, 2013).

Protoonkogeni su važni regulacijski stanični geni koji kodiraju proteine uključene u proces staničnog rasta i proliferacije. Onkogeni su nenormalno eksprimirani ili mutirani oblici odgovarajućih protoonkogena koji dovode do nenormalne proliferacije stanica i nastanka tumora. Tumor-supresorski geni kodiraju za proteine koji imaju ulogu inhibicije staničnog rasta i stvaranja tumora. Za inicijaciju i progresiju tumora nužne su mutacije koje dovode do aktivacije onkogena i inaktivacije tumor-supresorskih gena (Cooper i Hausman, 2004). Onkogeni djeluju kao dominantni geni što znači da je dovoljna mutacija u samo jednom alelu za inicijaciju stvaranja tumora. S druge strane, tumor-supresorski geni djeluju kao recesivni geni što znači da su potrebne mutacije u oba alela kako bi došlo do gubitka njegove funkcije. Recesivne mutacije podržavaju Knudsonovu hipotezu dvostrukog hica (eng. *Knudson's two-hit hypothesis*) koja govori kako su potrebne mutacije u oba alela za okidanje kancerogeneze. Pacijenti koji naslijede mutaciju jednog tumor-supresorskog alela mogu kroz vrijeme steći drugu somatsku mutaciju pa stoga imaju predispoziciju i povećan rizik za razvoj malignog tumora (Pecorino, 2012).

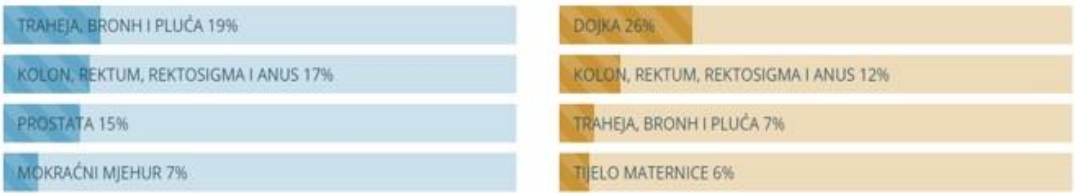
Stanice malignog tumora nastaju klonalno iz prve mutirane stanice koja sadrži inicijalnu mutaciju. Većina agenasa koji uzrokuju maligni tumor utječu na promjene u sekvenci molekule DNA, odnosno stvaraju mutacije. U tumorskim stanicama, molekula DNA sadrži točkaste mutacije (promjene u jednom paru baza) i/ili velike kromosomske aberacije (delecije, insercije ili kromosomske translokacije). Osnova kancerogeneze je nakupljanje više različitih mutacija u stanicama kroz vrijeme pa s godinama života raste i rizik za razvoj malignog tumora. Do

akumulacije mutacija dolazi tek nakon što je izbjegnut stanični obrambeni mehanizam, npr. mehanizam za popravak molekule DNA. Svaka promjena u molekuli DNA koja nije popravljena prenosi se na stanicu kćer (Pecorino, 2012). Mutacije mogu nastati u proteinima uključenima u popravak molekule DNA, staničnо signaliziranje, stanični ciklus i stanični rast, programiranu staničnu smrt i arhitekturu tkiva što dovodi do promjene u ukupnom broju stanica. Jedan od problema u razumijevanju malignih tumora je što nasljedna promjena može biti zbog genetičke promjene, kao što je promjena u sekvenci molekule DNA ili zbog epigenetičke promjene, kao što je trajna promjena u obrascu genske ekspresije, ali bez promjene u samoj sekvenci molekule DNA. Zbog mnoštva mutacija i epigenetičkih promjena, gotovo sve stanice malignog tumora su genetički nestabilne, a promjene mogu uzrokovati razni genetički faktori, stil života, infekcije i izlaganje različitim okolišnim faktorima (Alberts, 2013).

Tumorogeneza je proces od više koraka i svaki od tih koraka odražava genetičke promjene koje dovode do progresivne transformacije normalnih stanica u visoko maligne derivate. Definirano je 6 obilježja gotovo svih vrsta malignih tumora koja su esencijalna za tumorogenezu: stjecanje mogućnosti za autonomno staničnо signaliziranje; izbjegavanje signala supresora staničnog rasta; izbjegavanje apoptoze stanica; neograničeni potencijal za replikaciju; mogućnost angiogeneze te mogućnost invazije i metastaziranja. U stjecanju tih obilježja nužna je genomska nestabilnost i brojne mutacije, upala koja pogoduje nastanku tumora, nepravilna regulacija staničnog metabolizma te izbjegavanje imunosnog odgovora organizma (Hanahan i Weinberg, 2011).

Gotovo sve otkrivene mutacije u tumorskim stanicama su somatske mutacije gdje je molekula DNA somatske stanice oštećena. Takve mutacije se ne prenose na potomstvo i nisu nasljedne, ali se prenose na stanice kćeri dijeljenjem oštećene stanice pa se maligni tumor smatra bolešću genoma na staničnoj razini. Mutacije u molekuli DNA spermalne ili jajne stanice prenose se na potomstvo i mogu povećati rizik za razvoj malignog tumora, ali rijetko odmah uzrokuju njegov nastanak (Pecorino, 2012).

U Hrvatskoj, muškarci najčešće obolijevaju od malignog tumora traheja, bronha i pluća, a žene od karcinoma dojke (Slika 1).



Slika 1. Najčešći oblici malignih tumora kod muškaraca (lijevo) i žena (desno) u Hrvatskoj (preuzeto i prilagođeno prema http://www.onkologija.hr).

1.2. Karcinom dojke

Karcinom dojke je najčešća zloćudna (maligna) bolest među ženama u svijetu, može se javiti i kod muškaraca, ali je 100 puta češća pojavnost kod žena (Makki, 2015). Karcinom dojke je heterogen i hormonski ovisan maligni tumor te predstavlja 22,9% od ukupnih malignih tumora među ženama te je glavni uzrok smrti među ženama (Donepudi i sur. 2014). Kod žena u Aziji, najveći rizik za razvoj karcinoma dojke je u 40-tim godinama, dok je kod žena u Americi i Europi u 60-tim godinama (Mehrgou i Akouchekian, 2016). Na razvoj karcinoma dojke utječu brojni faktori kao što su: godine, genetički faktori, rana pojava prve menstruacije, kasni ulazak u menopauzu, kasno rođenje djeteta, izbjegavanje dojenja, izlaganje zračenju, korištenje hormonske terapije, neprimjerena tjelesna težina, način života, korištenje alkohola, pušenje i mnogi drugi (Gelband i sur. 2015). Kao rizični faktor za razvoj karcinoma dojke smatra se produljeno izlaganje organizma ženskim spolnim hormonima – estrogenima (Pecorino, 2012). Među svim rizičnim faktorima, najvažnija je obiteljska povijest karcinoma dojke (Mehrgou i Akouchekian, 2016).

Dojka se sastoji od dva osnovna tkiva – žljezdanog tkiva i potpornog tkivo. Žljezdano tkivo se sastoji od mliječnih žlijezda, koje proizvode mlijeko i podijeljene su u 15-20 režnjeva (lobula) koji se dijele na još manje režnjiće, i od mliječnih kanalića (duktusa) kojima se mlijeko iz mliječnih žlijezda odvodi do bradavica. Karcinom dojke je najčešća maligna lezija i nastaje iz epitelnih stanica unutarnjeg sloja režnjeva ili kanalića, a može se naići i na različite tipove sarkoma i limfoma. Općenito, morfološki se dijeli na karcinom *in situ* i invazivni karcinom, a ovisno o tome od kuda potječe klasificira se kao duktalni ili lobularni. Invazivni duktalni karcinom je najčešći oblik i pojavljuje se u 55% slučajeva dijagnosticiranog karcinoma dojke. Također, napravljena je i molekularna klasifikacija karcinoma dojke na temelju profila eksprimiranih gena, takozvanih tkivnih markera. Tako se razlikuju različiti podtipovi: luminalni A, luminalni B, tip nalik bazalnim stanicama i HER2 tip (Tablica 1) (Makki, 2015). Luminalni

A tip ima najbolju prognozu, najvišu stopu preživljenja i nisku stopu vraćanja karcinoma. Luminalni B tip se dijagnosticira u ranijoj životnoj dobi, ali ima lošiju prognozu od luminalnog A tipa. Oko 15-20% karcinoma dojke je okarakteriziran kao tip nalik bazalnim stanicama i pojavljuje se kod mlađih žena. Taj tip se najčešće razvija kod žena koje imaju mutacije u genu *BRCA1*, vrlo je agresivan te su loše prognoze. HER2 tip se pojavljuje kod 10-15% karcinoma dojke, ima lošu prognozu, sklon je vraćanju i stvaranju metastaza (Mehrgou i Akouchekian, 2016).

Tablica 1. Molekularni podtipovi karcinoma dojke. ER – estrogenski receptor (eng. *estrogen receptor*); PR – progesteronski receptor (eng. *progesteron receptor*); HER2 – receptor ljudskog epidermalnog faktora rasta 2 (eng. *human epidermal growth factor receptor 2*). (+) eksprimiran receptor; (-) - nema ekspresije receptora).

podtip karcinoma dojke	**ER**	**PR**	**HER2**
luminalni A	ER ili PR +		-
luminalni B	ER ili PR +		+
tip nalik bazalnim stanicama	-	-	-
HER-2 tip	-	-	+

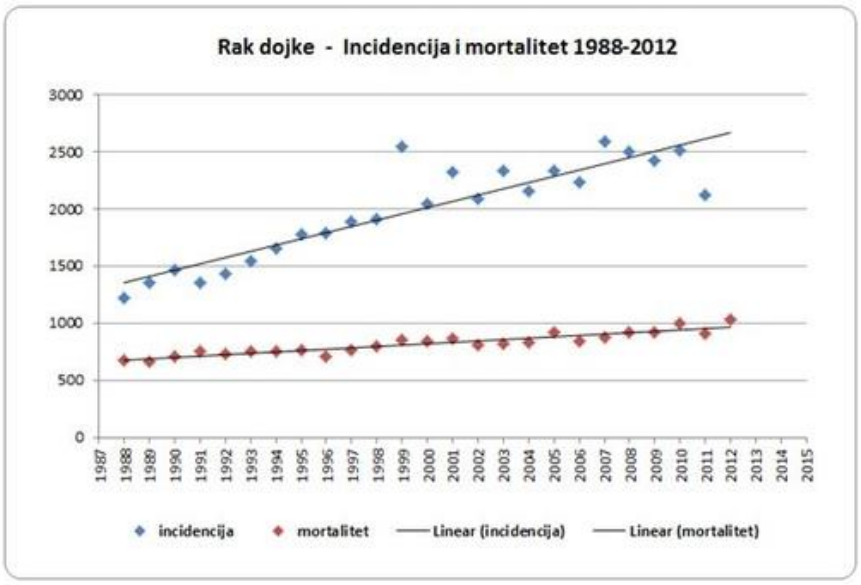
Veliki napredak u istraživanju karcinoma dojke omogućen je otkrićem gena *BRCA1* (eng. *breast cancer susceptibility gene 1*) i *BRCA2* (eng. *breast cancer susceptibility gene 2*) (Cavanagh i Rogers, 2015). Nasljedni sindrom karcinoma dojke i jajnika (HBOCS, od eng. *hereditary breast and ovarian cancer syndrome*) uzrokovan je nasljednim mutacijama u genima *BRCA1* i *BRCA2* koje se prenose kroz generacije muškim i ženskim gametama (Ahmadloo i sur. 2017). Nositelji tih mutacija su heterozigoti, što znači da imaju jednu normalnu i jednu mutiranu kopiju gena (Powell i Kachnic, 2003). Sindrom pokazuje autosomalno dominantno nasljeđivanje što znači da gubitkom drugog alela, odnosno gubitkom heterozigotnosti (LOH, od eng. *loss of heterozygosity*) dolazi do razvoja karcinoma (Roy i sur. 2011). Geni *BRCA1/2* kodiraju za tumor-supresorske proteine i slijede Knudsonovu hipotezu što znači da jedna naslijeđena mutacija daje predispoziciju za razvoj karcinoma dojke ili jajnika (Pecorino, 2012). Nositelji mutacija imaju 50-80% veći rizik razvoja karcinoma dojke te 30-50% veći rizik razvoja karcinoma jajnika. HBOC sindrom čini 5-7% svih uzroka karcinoma dojke, dok ostatak čini sporadični karcinom dojke koji nije uzorkovan nasljeđivanjem mutacija (Roy i sur. 2011).

Još uvijek je nejasno zašto mutacije u *BRCA1/2* uzrokuju predispozicija za karcinom dojke ili jajnike unatoč tome što je funkcija proteina BRCA1/2 esencijalna za sve stanične tipove. Pretpostavlja se da je epitelno tkivo dojke i jajnika osjetljivije na transformaciju u slučaju mutacija u genima *BRCA1/2* zbog specifičnog tkiva koje odgovara na utjecaj hormona (Yoshida i Miki, 2004). Također, pokazano je da nositelji mutacija u *BRCA1* pretežno razvijaju ER-negativne tumore, dok nositelji mutacija u *BRCA2* ne razvijaju preferencijalno određeni podtip karcinoma dojke (Roy i sur. 2011).

Također, drugi oblici karcinoma, kao što su karcinom prostate, gušterače i želuca imaju veći potencijal razvoja kod pojedinaca koji su nositelji mutacija u tim genima (D'Argenio i sur. 2015). Uz *BRCA1/2* gene, otkriveno je da su geni *TP53, FGFR2, TNRC9, MAP3K1, LSP1, CASP8, TGFB1* (Polyak, 2014), *PALB2, CHEK2, ATM, RAD51, MSH2, PMS2, MRE11, RAD50, NSB1, CDH1, BARD1* i mnogi drugi također povezani s rizikom za razvoj karcinoma dojke (Kamps i sur. 2017). Mehanizam kojim djeluju na tumorogenezu karcinoma dojke je uglavnom neistražen, ali se također mogu koristiti pri identifikaciji osoba s visokim rizikom za razvoj karcinoma dojke (Polyak, 2014).

Dosada je otkriveno oko 2000 mutacija u genima *BRCA1 i BRCA2* koje uglavnom dovode do nastanka skraćenih oblika proteina i time gubitka njihove funkcije. Najčešće mutacije su male insercije i male delecije (indel, od eng. INsertion/DELetion); promjene u jednom nukleotidu (SNV, od eng. *single nucleotide variant*) koje mogu biti besmislene (eng. *nonsense*) mutacije čime dolazi do promjene kodona u stop kodon i preranog završetka translacije proteina ili pogrešne (eng. *missense*) mutacije čime dolazi do zamjene aminokiselina i time promjene strukture i funkcije proteina; veliko gensko preslagivanje; prerani završetak transkripcije i problemi prilikom izrezivanja (eng. *splicing*) (Mehrgou i Akouchekian, 2016). Poznate patogene promjene u jednom nukleotidu su široko rasprostranjene duž cjelokupne kodirajuće regije gena *BRCA1/2* kao i unutar očuvanih regija introna oba gena (Wallace, 2016).

Na svjetskoj razini, 2012. godine je dijagnosticirano 1,7 milijuna novih slučajeva karcinoma dojke (http://www.wcrf.org), dok je 2011. godine preko 508 000 žena umrlo od posljedica karcinoma dojke (http://www.who.int). Prema podacima Registra za rak, Zavoda za javno zdravstvo Republike Hrvatske, godišnje od karcinoma dojke u Hrvatskoj oboli preko 2500 žena te je treći uzrok smrti u ženskoj populaciji (Slika 2) (http://www.onkologija.hr).



Slika 2. Kretanje incidencije i mortaliteta od karcinoma dojke u Hrvatskoj u periodu od 1988. do 2012. godine (preuzeto i prilagođeno prema http://www.onkologija.hr).

1.3. Geni *BRCA1* i *BRCA2*

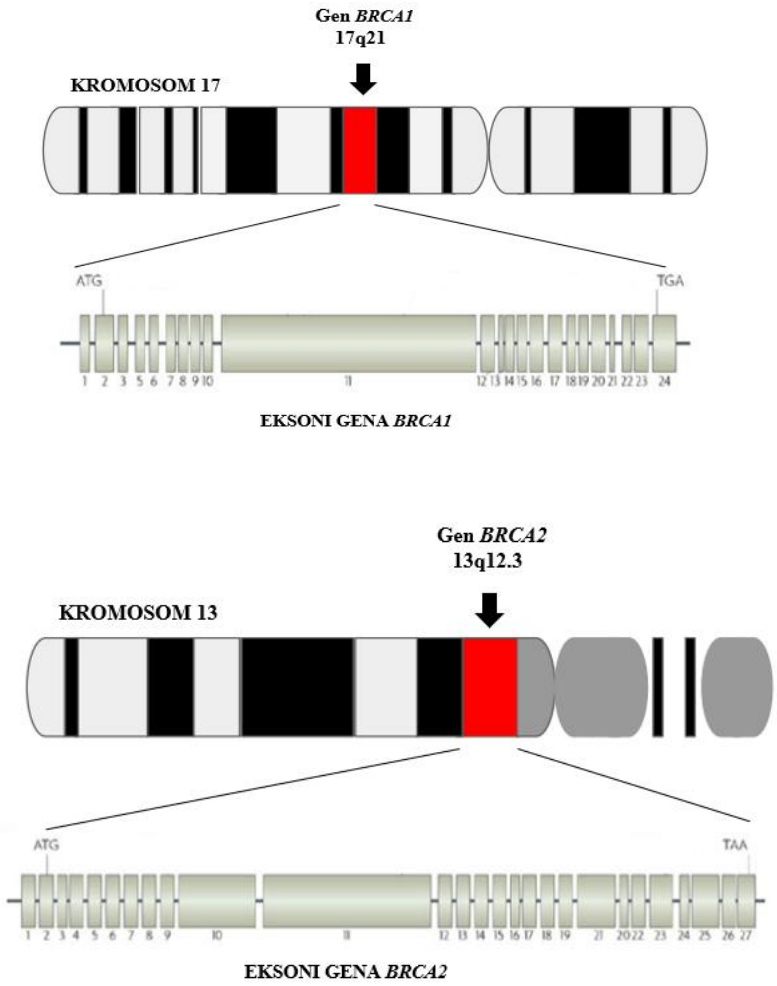
Genetičke studije iz 1990. godine pokazale su da je rizik za razvoj karcinoma dojke u nekim obiteljima povezan s kromosomom 17q21 (Hall i sur. 1990). Gubitak heterozigotnosti na 17q je pronađen u većini obiteljskog karcinoma dojke i jajnika, što je sugeriralo da se radi o tumor-supresorskim genima. Gen predispozicije na karcinom dojke, *BRCA1* (NG_005905.2), identificiran je 1994. godine i sastoji se od 24 kodirajuća egzona raširena na preko 100 kb genomske DNA te kodira za protein od 1863 aminokiseline (Slika 3) (Miki i sur. 1994). BRCA1 ostvaruje interakcije s proteinima uključenima u napredovanje staničnog ciklusa, regulaciju genske transkripcije, odgovoru na oštećenje molekule DNA i ubikvitinaciju. Gubitak njegove funkcije rezultira pogreškama u popravku molekule DNA i transkripciji, nenormalnom duplikacijom centrosoma, neispravnom regulacijom kontrolne točke G2/M staničnog ciklusa, oštećenjem kromosoma i neispravnom regulacijom diobenog vretena (Petrucelli i sur. 2010).

S obzirom da je 45% obiteljskog karcinoma dojke pokazivalo vezu s genom *BRCA1*, nastavljena je potraga za drugim genom predispozicije na karcinom dojke te je tako 1995. godine otkriven gen *BRCA2* (NG_012772) na kromosomu 13q12.3 (Yoshida i Miki, 2004). Gen sadrži 27 kodirajuća egzona i kodira za protein koji se sastoji od 3418 aminokiseline (Slika 3) (Powell i Kachnic, 2003). Gen *BRCA2* nema prepoznatljivog proteinskog motiva i očite povezanosti s genom *BRCA1*, ali je također uključen u procese popravka molekule DNA (Petrucelli i sur. 2010). Kod nositelja mutacija u genu *BRCA2* kasnije se razvija karcinom dojke

u odnosu na nositelje mutacija u genu *BRCA1* te su mutacije u *BRCA2* povezanije s razvojem karcinoma prostate, želuca, melanoma i karcinoma dojke kod muškaraca (Powell i Kachnic, 2003).

S obzirom da su oba proteina BRCA uključena u homolognu rekombinaciju i popravak dvolančanog loma molekule DNA, mutacije u genima *BRCA1/BRCA2* uzrokuju pogreške prilikom rekombinacije što destabilizira genom i dovodi do kromosomskih rearanžmana i mutacija (Pecorino, 2012).

Mutacije u genima *BRCA1/2* pojavljuju se u 1 na 400 do 1 na 800 osoba, ovisno o etnicitetu te spadaju među najviše istražene gene (Petrucelli i sur. 2010). Detekcija mutacija u genima *BRCA1/2* važno je zbog različitih terapeutskih pristupa i potencijalno specifične ciljane terapije (Castéra i sur. 2014).



Slika 3. Shema pozicije gena *BRCA1/2* i njihovih eksona na kromosomima (preuzeto i prilagođeno prema Fackenthal i Olopade, 2007).

1.3.1. Uloga BRCA1 i BRCA2 u popravku molekule DNA

BRCA1 i BRCA2 su uključeni u komplekse koji aktiviraju popravak dvolančanog loma (DSB, od eng. *double-strand breaks*) i iniciraju homolognu rekombinaciju (HR, od eng. *homologous recombination*) (Yoshida i Miki, 2004). DSB se smatra najopasnijim oblikom oštećenja molekule DNA budući da je istodobno ugrožen integritet oba lanca molekule DNA. Nastaje kao nusprodukt replikacije molekule DNA ili za vrijeme izlaganja ionizirajućem zračenju i drugim genotoksičnim spojevima. DSB se popravlja homolognom rekombinacijom ili nehomolognim spajanjem krajeva. HR je glavni mehanizam zaštite integriteta genoma proliferirajućih stanica budući da su drugi putevi popravka dvolančanog loma skloni pogreškama i stvaraju kromosomske delecije i translokacije. Prepoznavanje oštećenja molekule DNA i popravak dvolančanog loma uključuje senzore koji detektiraju slomljene krajeve molekule DNA, efektore koji izvršavaju popravak i medijatore koji olakšavaju interakciju između senzora i efektora (Roy i sur. 2011).

Kao odgovor na oštećenje molekule DNA, razne kinaze uključene u prepoznavanje oštećenja molekule DNA (DDR, od eng. *DNA damage response kinases*) kao što su kinaza ataksije teleangiektazije (ATM, od eng. *ataxia-telangiectasia mutated kinase*), kinaza povezana s ATM (ATR, od eng. *ATM-related kinase*) i kontrolna kinaza G2/M (CHK2, od eng. *G2/M control kinase*), fosforiliraju različite aminokiselinske ostatke proteina BRCA1. BRCA1 se hiperfosforilira i premješta u replikacijsku viljušku označenu jezgrinim antigenom proliferirajućih stanica (PCNA, od eng. *proliferating cell nuclear antigen*) (Yoshida i Miki, 2004).

Dvolančani lom aktivira ATM kojoj je supstrat kompleks MRN (MRE11-RAD50-NSB1) koji svojom 5'-3' egzonukleaznom aktivnošću stvara jednolančane 3' krajeve (Pecorino, 2012). BRCA1/2 stvaraju kompleks s proteinom RAD51 koji uz pomoć RAD52 oblaže jednolančanu molekulu DNA i stvara nukleoproteinski filament koji prodire u dvolančanu molekulu DNA, tamo se sparuje s homolognom regijom i tako omogućuje izmjenu lanaca molekule DNA (Slika 4) (Yoshida i Miki, 2004). BRCA1 dovodi RAD51 na mjesto oštećenja molekule DNA preko interakcije s PALB2 (eng. *partner and localizer of BRCA2*) i BRCA2 (Roy i sur. 2011). Rezolvaza raspetljava čvor tj. Hollidayevu strukturu koja je rezultat homologne rekombinacije (Pecorino, 2012). Kolokalizacija proteina BRCA1/2 s RAD51 na mjestima rekombinacije i popravka molekule DNA pokazuje kako proteini BRCA1/2 imaju ulogu i u detekciji i u popravku dvolančanog loma (Yoshida i Miki, 2004).

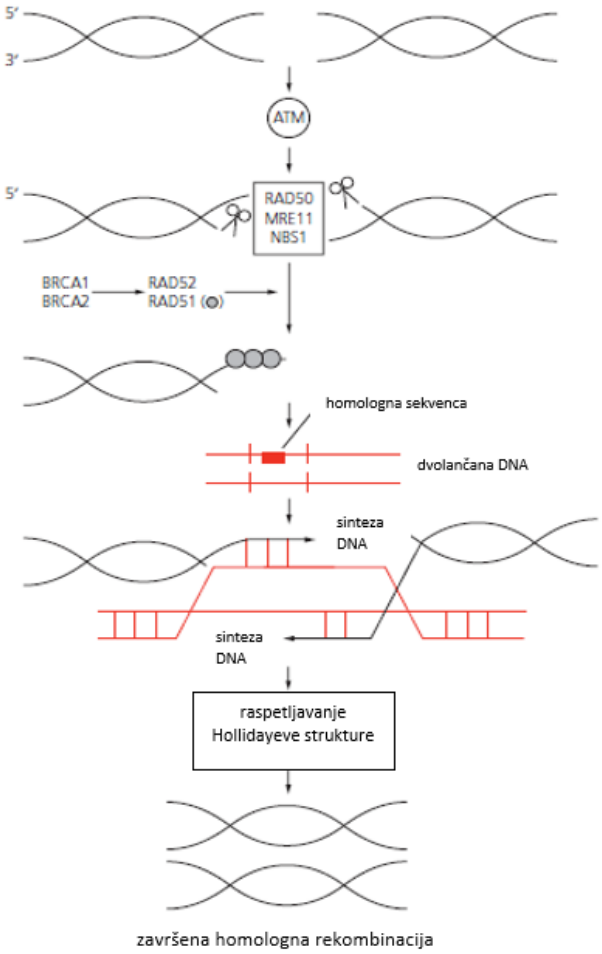
BRCA2 sadrži DNA-vezujuću domenu koja veže jednolančanu i dvolančanu molekulu DNA te 8 ponavljanja aminokiselinskog motiva - ponavljanja BRC, kodirana velikim egzonom

11, koja vežu RAD51 (Slika 5) (Roy i sur. 2011). Pokazano je da protein BRCA2 ima direktniju ulogu u popravku homolognom rekombinacijom u odnosu na protein BRCA1. Interakcije između BRCA2 i RAD51 su neophodne za održavanje stanične diobe i kromosomske strukture. BRCA2 regulira unutarstaničnu lokalizaciju i funkciju RAD51 tako što premješta RAD51 od mjesta sinteze na mjesto obrade oštećenja molekule DNA (Yoshida i Miki, 2004).

BRCA1 možda ima ulogu i u drugim putevima popravka molekule DNA kao što je popravak nehomolognim spajanjem krajeva (NHEJ, od eng. *non-homologous end-joining*) ili sparivanjem jednolančanih krajeva (SSA, od eng. *single-strand annealing*). Pretpostavlja se da je BRCA1 uključena u HR, NHEJ i SSA preko interakcija s kompleksom MRN koji je važan u inicijaciji popravka dvolančanog loma (Roy i sur. 2011). Također, pokazana je kolokalizacija BRCA1 s H2AX (γ-H2AX) koji se fosforilira nakon dvolančanog loma što upućuje na to da BRCA1 možda regrutira kinaze odgovorne za fosforilaciju H2AX (Yoshida i Miki, 2004).

Istraživanja su pokazala da BRCA1 doprinosi i regulaciji aktivnosti c-Abl. c-Abl tirozinska kinaza je eksprimirana i lokalizirana u citoplazmi i jezgri te ima ulogu u indukciji apoptoze, regulaciji transkripcije i popravku molekule DNA. Time se pretpostavlja da BRCA1, uz direktnu ulogu u popravku molekule DNA ili regulaciju staničnog ciklusa, ima ulogu i u staničnom odgovoru na oštećenje molekule DNA (Yoshida i Miki, 2004).

Osobe koje imaju naslijeđenu heterozigotnu mutaciju u genu *BRCA1* ili *BRCA2* imaju neispravan popravak homolognom rekombinacijom što ukazuje da je BRCA1-BRCA2 put homologne rekombinacije važan za supresiju nasljednog karcinoma dojke (Roy i sur. 2011).



Slika 4. Shema homologne rekombinacije (preuzeto i prilagođeno prema Pecorino, 2012).

1.3.2. Uloga BRCA1 i BRCA2 u transkripcijskoj regulaciji

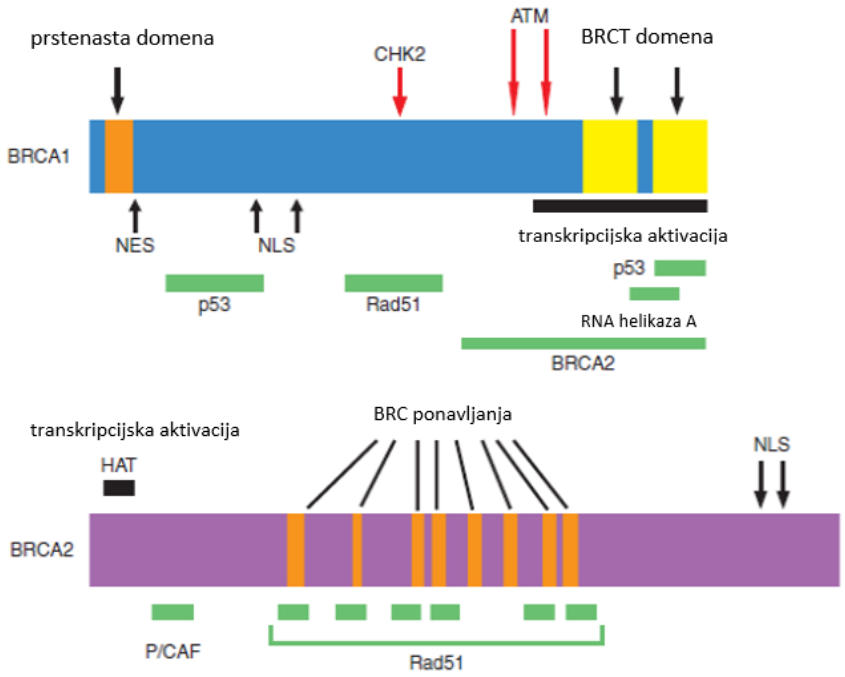
BRCA1 ima ulogu i u popravku oksidativno oštećene molekule DNA za vrijeme transkripcije, najčešće procesom izrezivanja baza ili nukleotida. Unutar stanice, BRCA1 je dio velikog proteinskog kompleksa veličine 3 MDa (Powell i Kachnic, 2003). BRCA1 ima N-terminalnu cink-vezujuću prstenastu domenu koja ima E3 aktivnost ubikvitinske ligaze te C-terminalnu domenu BRCT koju prepoznaju mnogi proteini popravka molekule DNA (Slika 5) (Roy i sur. 2011). C-terminus proteina BRCA1 veže i aktivira bazalnu transkripcijsku mašineriju tako što stvara kompleks s RNA polimerazom II preko RNA helikaze A, a njegova N-terminalna prstenasta domena omogućuje interakciju s proteinom 1 prstenaste domene povezanim s BRCA1 (BARD1, od eng. *BRCA1-associated RING domain protein 1*) pomoću kojeg se BRCA1 veže kao heterodimer na holoenzim RNA polimerazu II. Nadalje, BRCA1 veže veliki broj transkripcijskih faktora pomoću kojih se prenose signali RNA polimerazi II (Yoshida i Miki, 2004). Mnoge nasljedne mutacije u BRCA1 su pronađene u prstenastoj

domeni i domeni BRCT što pokazuje da su obje domene uključene u suprimiranje karcinoma dojke i jajnika (Roy i sur. 2011).

Istraživanja su pokazala da se prekomjernom ekspresijom BRCA1 stimulira inhibitor straničnog ciklusa p53 kao i faktor odgovora na stres p21. BRCA1 ostvaruje interakcije s p53, služi kao njegov koaktivator i tako stimulira njegov put (Yoshida i Miki, 2004). p53 ima ulogu u održavanju genomskog integriteta regulacijom transkripcije ciljnih gena koji su uključeni u zaustavljanje staničnog ciklusa, apoptozu ili popravak molekule DNA. Istraživanja pokazuju da su mutacije u *TP53* prisutne u 30-50% malignih tumora te da se često događaju istovremeno s mutacijama u genima *BRCA1* ili *BRCA2* (Roy i sur. 2011). To dovodi do gubitka regulacije kontrolnih točaka staničnog ciklusa, induciranja nekontrolirane proliferacije i invazivnog rasta (Yoshida i Miki, 2004).

Nadalje, BRCA1 veže transkripcijski faktor ZBRK1 koji specifično prepoznaje određeni motiv u dijelu molekule DNA koja kodira za promotorsku regiju proteina p21. BRCA1 i STAT1 sudjeluju u regulaciji p21 tako što se BRCA1 veže na transkripcijsku aktivacijsku domenu STAT1 što dovodi do indukcije seta IFN- γ reguliranih gena. Time se pretpostavlja da BRCA1 ima ulogu u povezivanju oštećenja molekule DNA s putem odgovora na stres kako bi se izvršio specifični stanični odgovor kao što je zaustavljanje staničnog ciklusa ili apoptoza (Yoshida i Miki, 2004).

S druge strane, uloga BRCA2 u transkripcijskoj regulaciji je daleko manja. Smatra se da bi mogao imati utjecaj na aktivaciju transkripcije tako što modulira acetilaciju histona. Stvara interakcije s transkripcijskim koaktivatorskim proteinima, kao što su faktor povezan s p300/CBP (P/CAF, od eng. *p300/CBP-associated factor*) i p300/CBP koji posjeduju histonsku acetilaznu aktivnost, i regrutira ih na transkripcijske komplekse (Yoshida i Miki, 2004). Također, BRCA2 ima ulogu i u zaštiti lanca molekule DNA od degradacije prilikom zastoja replikacijske viljuške (Roy i sur. 2011).



Slika 5. Značajke proteina BRCA1/2: N-terminalna prstenasta domena, jezgrin lokalizacijski signal (NLS), jezgrin izlazni signal (NES), dvije C-terminalne domene BRCT (žuto), mjesta fosforilacije kinazama (CHK2 i ATM) (crvene strelice), interakcijski proteini (zeleni štapići), 8 ponavljanja motiva BRC (narančasto) (preuzeto i prilagođeno prema Yoshida i Miki, 2004).

1.3.3. Uloga BRCA1 i BRCA2 u kontrolnim točkama staničnog ciklusa

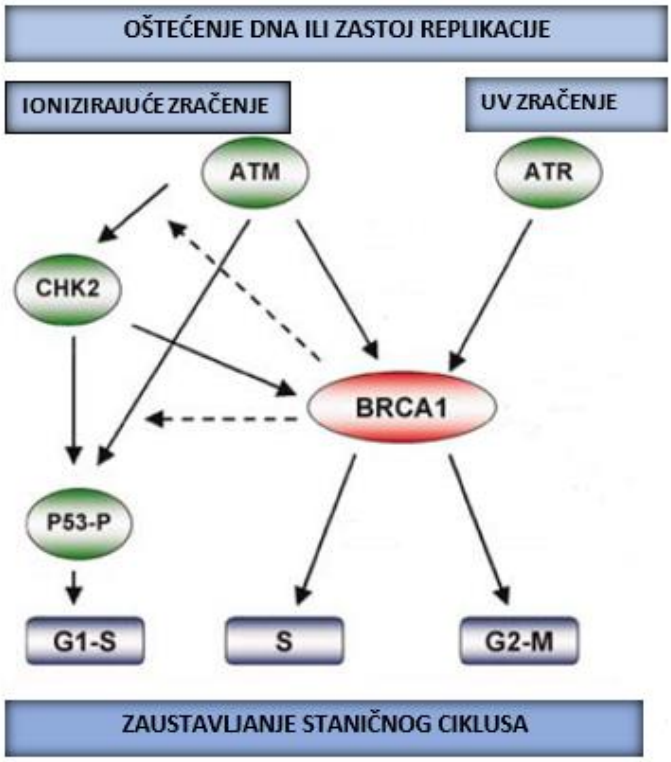
Događaji koji omogućuju prepoznavanje oštećenja molekule DNA uključuju aktivaciju kontrolnih točaka koje odgađaju stanični ciklus prije replikacije (kontrolna točka G1/S), za vrijeme replikacije (kontrolna točka unutar S-faze) ili prije stanične diobe (kontrolna točka G2/M) kako bi se osiguralo da se genetičke pogreške ne prenose na sljedeće generacije i omogućilo vrijeme potrebno za popravak molekule DNA (Roy i sur. 2011).

Istraživanja su pokazala da su ATM i BRCA1 potrebne za efikasnu S-fazu i kontrolnu točku G2/M. Nadalje, BRCA1 regulira kontrolnu točku G2/M aktivacijom kinaze CHK1 kojom se inducira signalna kaskada nizvodno od CHK1 (Slika 6) (Yoshida i Miki, 2004).

14-3-3σ je glavni gen kontrolne točke G2/M te je reguliran pomoću p53 za čiju je aktivaciju odgovoran BRCA1. Također, BRCA1 aktivira kontrolnu točku G2/M i aktivacijom GADD45 proteina, dok aktivacija G1 ciklički-ovisnog kinaznog inhibitora p21 blokira ulazak u S-fazu staničnog ciklusa. Isto tako, BRCA1 aktivira ciklički-ovisni kinazni inhibitor p27KIPI što dovodi do zastoja staničnog ciklusa u fazi G1 (Yoshida i Miki, 2004). Kompleks BRCA1-BRIP1-DNA topoizomeraza 2 vezujući protein 1 (TOPBP1) je važan za kontrolnu točku S-faze

u odgovoru na zastoj ili raspadanje replikacijske viljuške dok je kompleks BRCA1-ABRAXAS-RAP80 uključen u kontrolnu točku G2/M u odgovoru na oštećenje nastalo ionizirajućim zračenjem (Roy i sur. 2011).

BRCA2 posreduje u kontrolnoj točki G2/M interakcijom s faktorom 35 povezanim s BRCA2 (BRAF35). Također, pokazano je da kompleks BRCA2/BRAF35 sudjeluje u modulaciji progresije metafaze (Yoshida i Miki, 2004).



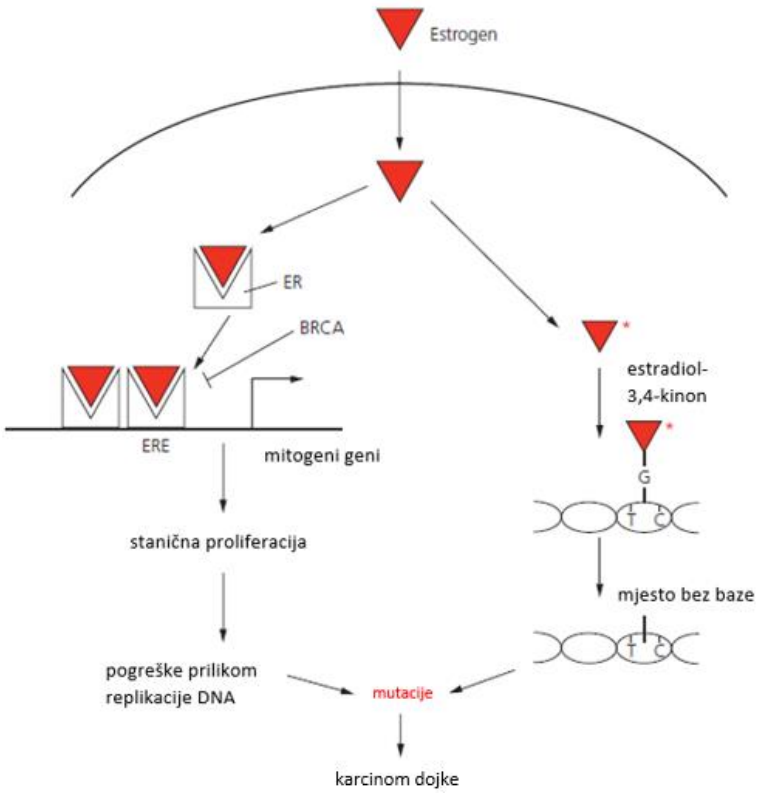
Slika 6. Uloga BRCA1 u regulaciji staničnog ciklusa kao odgovor na oštećenje DNA, zastoj replikacije i zračenja (preuzeto i prilagođeno prema Gudmundsdottir i Ashworth, 2006).

1.3.4. Utjecaj hormona na razvoj karcinoma dojke

BRCA1 i BRCA2 štite genom od pogrešaka za vrijeme replikacije molekule DNA i jasno je da će se u slučaju gubitka njihove funkcije razviti potencijalno onkogene genetičke promjene. Međutim, gubitkom njihove funkcije najveća je predispozicija za razvoj karcinoma dojke ili jajnika što se može objasniti time što su epitelne stanice dojke i jajnika podložne snažnim signalima rasta hormonskom stimulacijom za vrijeme normalnog menstrualnog ciklusa (Roy i sur. 2011). Predložena su dva modela kojima estrogeni djeluju na razvoj karcinoma dojke (Slika 7). Jedan model pretpostavlja da estrogeni promoviraju proliferaciju

stanica dojke što ostavlja manje vremena za popravak molekule DNA pa je i nastanak pogreški prilikom replikacije molekule DNA puno veći. Efekt estrogena je posredovan estrogenskim receptorima koji se vežu kao dimeri na estrogenski element u molekuli DNA (ERE, od eng. *estrogen response element*) i reguliraju gene osjetljive na estrogen. Geni *BRCA1/2* inhibiraju transkripcijsku aktivaciju estrogenskih receptora što znači da suprimiraju proliferativni efekt estrogenskog signaliziranja. Kod pacijenata s nasljednom mutacijom u genima *BRCA1/2* povećan je rizik za razvoj karcinoma dojke zbog gubitka modulacije estrogenskog signaliziranja (Pecorino, 2012).

Drugi model pretpostavlja da su estrogen i njegovi metaboliti genotoksični (Pecorino, 2012). Hormonski posredovan rast za vrijeme svakog menstrualnog ciklusa stvara reaktivne kisikove radikale koji uzrokuju oksidativna oštećenja molekule DNA (Roy i sur. 2011). Estradiol se metabolizira do estradiol-3,4-kinona u stanicama koji se kovalentno veže na adeninsku ili gvaninsku bazu. Takvi adukti destabiliziraju vezu između baza i molekule DNA što rezultira abazičnim mjestom i mutacijama. Osobe s nasljednom mutacijom u genima *BRCA1/2* imaju povećan rizik za razvoj karcinoma dojke jer se povećava genotoksični efekt estrogenskih metabolita budući da nema pravilnog popravka molekule DNA (Pecorino, 2012).



Slika 7. Mehanizam djelovanja estrogena na razvoj karcinoma dojke (preuzeto i prilagođeno prema Pecorino, 2012).

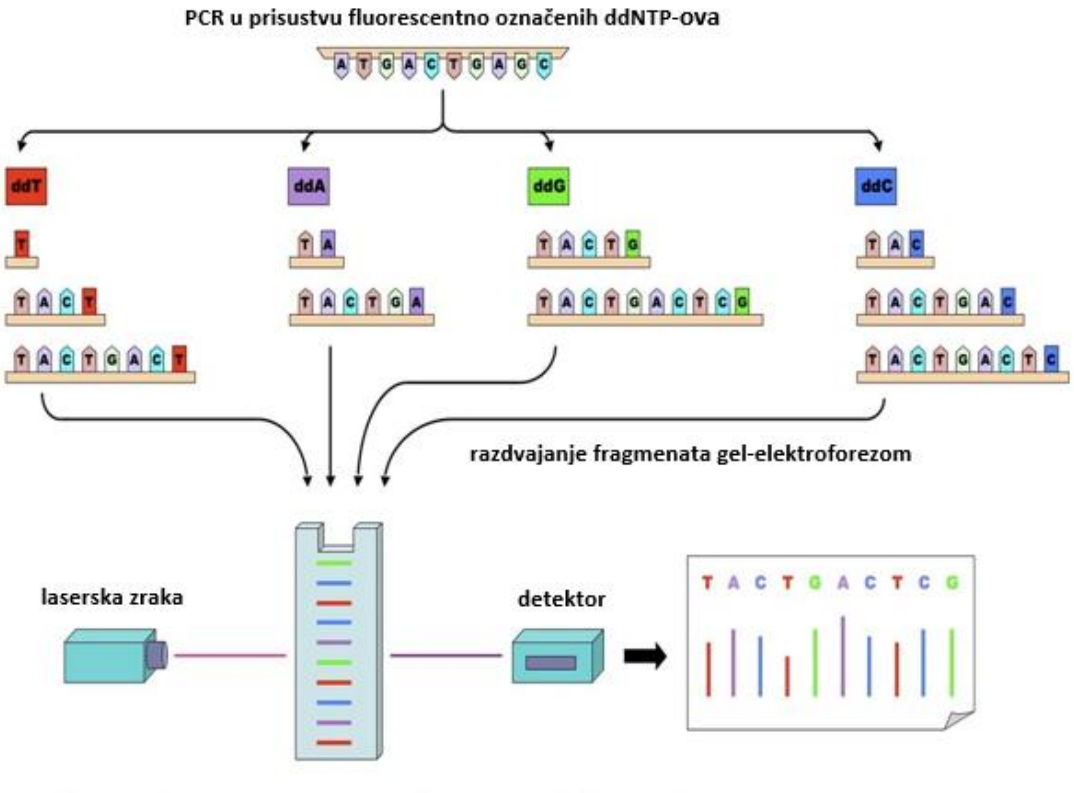
1.4. Sekvenciranje

1.4.1. Razvoj metoda sekvenciranja

Slijed baza u polinukleotidnom lancu sadrži konačnu informaciju o nasljednim i biokemijskim svojstvima živih bića (Heather i Chain, 2015). Prve metode sekvenciranja razvijene su kasnih 70-tih godina prošlog stoljeća. Posljednjih godina osnovne tehnike su unaprjeđene automatizacijom što je omogućilo veću efikasnost sekvenciranja. Tri su osnovna preduvjeta za uspješno sekvenciranje: 1. fragmenti DNA moraju biti pripremljeni i obrađeni tako da su prikladni za sekvenciranje; 2. sekvencer mora svaku bazu očitati na način koji će omogućiti njenu identifikaciju; 3. detekcijske metode moraju osigurati brzu i preciznu identifikaciju baza (Nicholl, 2008).

1977. godine, paralelno su objavljene dvije metode sekvenciranja molekule DNA: Maxam-Gilbertova metoda sekvenciranja (kemijska metoda) i Sangerova metoda sekvenciranja (enzimska, dideoksi ili metoda terminacije sinteze lanaca) (Heather i Chain, 2015). Allan Maxam i Walter Gilbert razvili su metodu koja koristi kemikalije za cijepanje molekule DNA na točno određenom mjestu što generira niz fragmenata koji se razlikuju za jedan nukleotid. S druge strane, Fred Sanger i Alan Coulson su razvili metodu koja koristi enzimsku sintezu lanaca molekule DNA koji završavaju modificiranim nukleotidima. Modificirani nukleotidi su kemijski analozi deoksiribonukleotid trifosfata (dNTP-ovi) - dideoksinukleotid trifosfati (ddNTP-ovi) kojima nedostaje hidroksilna skupina na 3' kraju deoksiriboze čime se zaustavlja daljnji rast lanca molekule DNA jer je onemogućeno povezivanje s 5' fosfatom sljedećeg dNTP-a. Provode se četiri odvojene reakcije tako da svaka sadrži određeni ddNTP i smjesu preostalih dNTP-ova te se dobiju različite duljine lanaca molekule DNA. Analiza fragmenata je slična za obje metode i uključuje gel-elektroforezu i autoradiografiju ukoliko su nukleotidi radioaktivno obilježeni. Enzimska metoda je gotovo u potpunosti zamijenila kemijsku metodu zbog mogućnosti automatizacije i upotrebe fluorescentno označenih ddNTP-ova (Nicholl, 2008). Time je omogućena detekcija sva četiri nukleotida u jednoj reakciji i unaprijeđena je detekcija pomoću kapilarne elektroforeze. Razvojem tehnika kao što je lančana reakcija polimerazom (PCR, od eng. *polymerase chain reaction*) i rekombinantna DNA tehnologija, omogućeno je stvaranje visoke koncentracije DNA potrebne za sekvenciranje. Automatizirana Sangerova metoda se naziva i metodom sekvenciranja prve generacije (eng. *first-generation sequencing*) (Slika 8) (Heather i Chain, 2015) te se smatrala „zlatnim standardom“ sekvenciranja molekule DNA (Kamps i sur. 2017)

Projekt ljudskog genoma (HGP, od eng. *Human Genome Project*) je započeo 1990. godine i trebalo mu je 13 godina i milijarde eura kako bi se sekvencirao čitav ljudski genom (Kamps i sur. 2017). Noviji dideoksi sekvencer, kao što je ABI PRISM (Applied Biosystems) je omogućio istovremeno sekvenciranje stotina uzoraka i time pomogao u dobivanju konačne sekvence ljudskog genoma (Heather and Chain, 2015). Završetkom projekta zaključeno je da je za odgovore na kompleksna biološka pitanja, koja su proizašla nakon sekvenciranja, potrebna naprednija tehnologija i veća količina podataka (Goodwin i sur. 2016).



Slika 8. Automatizirana Sangerova metoda sekvenciranja (preuzeto i prilagođeno prema www.researchgate.net).

1.4.2. Sekvenciranje sljedeće generacije

Glavna ograničenja tradicionalnog sekvenciranja predstavljala je limitirana propusnost (eng. *throughput*) i visoka cijena sekvenciranja cijelog genoma sve do razvoja visoko propusnih (eng. *high-throughput*) platformi za sekvenciranje, odnosno do razvoja metoda sekvenciranja sljedeće generacije (NGS, od eng. *next-generation sequencing*) (Goodwin i sur. 2016). NGS ili masivno paralelno sekvenciranje omogućuje istovremeno sekvenciranje milijuna fragmenata

molekule DNA bez prethodnog znanja o samoj sekvenci. Takva napredna tehnologija je bila prava revolucija naspram tradicionalnih metoda sekvenciranja u kojoj su se relativno kratki fragmenti molekule DNA prethodno umnožili PCR-om i zasebno sekvencirali. S obzirom na visoku cijenu i dugotrajan rad, tradicionalno sekvenciranje se provodilo samo na specifičnim regijama molekule DNA i za specifične uzorke (Kamps i sur. 2017). Također, najčešće je uz Sangerovo sekvenciranje potrebno kombinirati i razne druge metode, a rezultati probira za nasljedne mutacije često budu negativni, čak i u obiteljima sa snažnom sumnjom na prisutnost nasljednih mutacija (Castéra i sur. 2014). NGS omogućuje jeftino stvaranje velike količine podataka po jednoj reakciji sekvenciranja (Metzker, 2010) te se danas čitav genom može sekvencirati u nekoliko dana za nešto više od 1000 eura po genomu (Kamps i sur. 2017).

Početni materijal za sekvenciranje može biti genomska DNA (DNA-seq), glasnička ili nekodirajuća RNA (RNA-seq) ili bilo koji nukleinski/ribonukleinski materijal uz prethodnu obradu. Trenutne NGS tehnologije se dijele u dvije skupine: sekvenciranje kratkih sekvenci (eng. *short-read sequencing*) i sekvenciranje dugih sekvenci (eng. *long-read sequencing*) (Kamps i sur. 2017).

1.4.2.1. Sekvenciranje kratkih sekvenci – druga generacija sekvenciranja

Postoje dva pristupa sekvenciranja kratkih sekvenci – sekvenciranje vezanjem (SBL, od eng. *sequencing by ligation*) i sekvenciranje sintezom (SBS, od eng. *seqeuncing by synthesis*). SBS pristup uključuje veliki broj metoda ovisnih o DNA-polimerazi i mogu se klasificirati kao ciklička reverzibilna terminacija (CRT, od eng. *cyclic reversible termination*) ili kao dodatak jednog nukleotida (SNA, od eng. *single-nucleotide addition*) (Goodwin i sur. 2016).

Prvi komercijalni NGS sekvencer koji se bazirao na SBS pristupu dodatkom jednog nukleotida bio je pirosekvencer 454 (Roche) (Slika 9, B) (Kamps i sur. 2017). Tehnologija pirosekvenciranja se razlikovala od dotadašnjih po tome što nije koristila radioaktivno ili fluorescentno označene dNTP-ove ili oligonukleotide za vizualizaciju. Umjesto toga, koriste se dva enzimska procesa: ATP sulfurilaza prevodi pirofosfat, oslobođen ugradnjom nukleotida, u ATP kojeg enzim luciferaza koristi kao supstrat i na taj način stvara svjetlost proporcionalnu količini oslobođenog pirofosfata. Biblioteke molekula DNA se vežu na zrnca (eng. *beads*) pomoću adapterske sekvence. U prosjeku se jedna molekula DNA veže na jednom zrncu nakon čega se provede emulzijski PCR kako bi se svako zrnce prekrilo klonalnom populacijom DNA. Sekvenciranje se provodi dodatkom enzima i dNTP-ova, a oslobađanje pirofosfata se mjeri senzorima CCD (eng. *charged couple device*) (Heather i Chain, 2016).

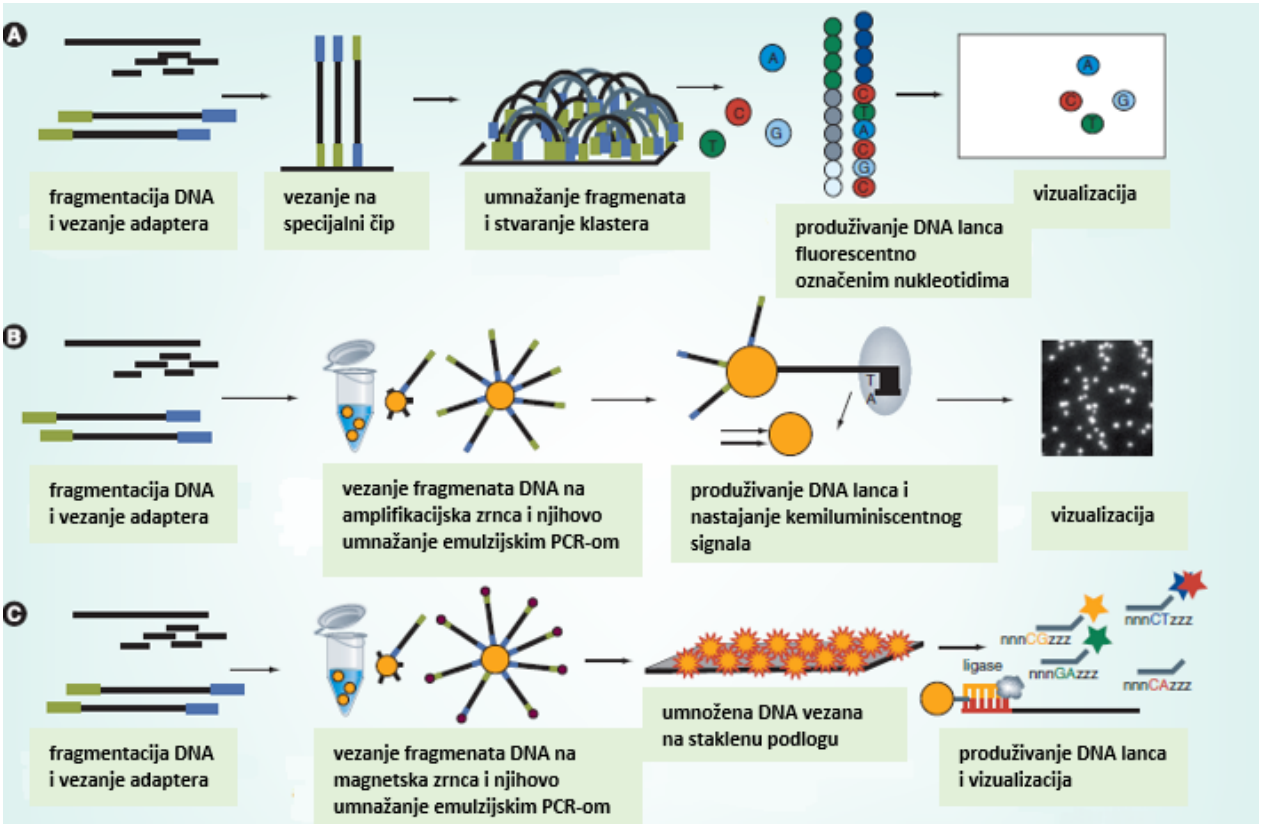
Prednost pirosekvenciranja u odnosu na Sangerovu dideoksi metodu je bila u tome što koristi prirodne nukleotide i može se promatrati u stvarnom vremenu stoga se ne treba provoditi elektroforeza. Fragmenti DNA koje je moguće na ovaj način sekvencirati su duljine od 400 do 500 pb te se istovremeno može sekvencirati više milijuna jažica koje sadrže klonalnu populaciju DNA (Heather i Chain, 2016).

Nakon sekvencera 454, sljedeća tehnika koja se pojavila na tržištu je bila Illumina koja se bazirala na SBS pristupu cikličkom reverzibilnom terminacijom (Slika 9, A) (Goodwin i sur. 2016) i može sekvencirati sekvence do 300 pb (Kamps i sur. 2017). Prvi korak je priprema biblioteke, odnosno, obrada početnih uzoraka. Materijal se mehanički ili enzimatski fragmentira na odsječke veličine koji su kompatibilni sa sekvencerom. Na 5' ili 3' kraj fragmenata se vežu adapteri te se takvi fragmenti umnože PCR-om i pročiste. Biblioteka se nanese na specijalni čip (eng. *flow cell*) za koji se fragmenti biblioteke vežu zahvaljujući komplementarnosti adaptera i oligonukleotida koji su vezani na čip. Potom slijedi stvaranje klastera (eng. *cluster generation*) PCR-om na čvrstoj fazi čime se svaki fragment umnoži u mnogo kopija (eng. *bridge amplification*) koje su označene jedinstvenom nukleotidnom oznakom od 6 do 8 nukleotida – barkodom. S obzirom da se sekvenciranje svih fragmenata odvija paralelno, jedinstveni barkod omogućuje razvrstavanje fragmenata prema početnom uzorku. Illumina koristi modificirane dNTP-ove na čijim se 3' hidroksilnim pozicijama nalaze terminatori koji blokiraju daljnju polimerizaciju. Terminatori sadrže fluorescentnu oznaku koja se detektira ekscitacijom prikladnim laserom sustava CCD nakon ugradnje u DNA lanac. Za vrijeme svakog ciklusa sekvenciranja prisutna su sva četiri modificirana dNTP-a te se nakon ugradnje i detekcije pojedinog nukleotida uklanja terminator, svi neupotrijebljeni nukleotidi budu isprani i započinje idući ciklus sekvenciranja (Heather i Chain, 2016; Illumina, 2016).

Treća opcija koja se pojavila na tržištu je bio sustav SOLiD (eng. *Sequencing by Oligonucleotide Ligation and Detection*) (Applied Biosystems) koji se bazira na SBL pristupu (Slika 9, C). Koriste se probe koje sadrže jednu ili dvije baze i fluorescentno su označene. Početnice su komplementarne adapterskoj sekvenci i omogućuju mjesto od kuda započinje povezivanje proba za početnicu pomoću DNA ligaze. Novi ciklus započinje nakon potpunog uklanjanja kompleksa između početnice i probe ili nakon cijepanja kako bi se uklonio fluorofor i obnovilo mjesto za vezanje proba. Nakon povezivanja, kalup se vizualizira i identificiraju se povezane probe (Goodwin i sur. 2016).

Nakon sekvencera 454, razvijena je još jedna tehnologija koja se bazirala na SBS pristupu dodatkom jednog nukleotida – Ion Torrent (Life Technologies). Kao i kod 454 sekvencera, na zrncima se nalazi klonalna populacija fragmenata DNA proizvedena

emulzijskim PCR-om. Ugradnjom svakog dNTP-a dolazi do oslobađanja H^+ iona. Promjena u pH vrijednosti se detektira senzorom, ali je ograničena točnost interpretacije homopolimernih sekvenci kada dolazi do ugradnje više istih dNTP-ova (Heather i Chain, 2016).



Slika 9. Shema metoda sekvenciranja druge generacije. A – Sustav Illumina; B – Sustav Roche 454; C – Sustav SOLiD (preuzeto i prilagođeno prema Natrajan i Reis-Filho, 2011).

1.4.2.2. Sekvenciranje dugih sekvenci – treća generacija sekvenciranja

Dosada su razvijena dva glavna pristupa sekvenciranja dugih sekvenci – sekvenciranje pojedinačne molekule u stvarnom vremenu (SMRT, od eng. *single-molecule real-time sequencing*) i sintetički pristup koji se oslanja na postojeće tehnike očitanja kratkih sekvenci kako bi se *in silico* konstruirala dugačka sekvenca. Pristup SMRT se razlikuje od pristupa sekvenciranja kratkih sekvenci po tome što se mogu sekvencirati molekule bez prethodne potrebe za amplifikacijom molekule DNA. Sintetički pristup se oslanja na pripremu biblioteke i korištenje barkodova koji omogućuju računalno slaganje većih fragmenata (Goodwin i sur. 2016).

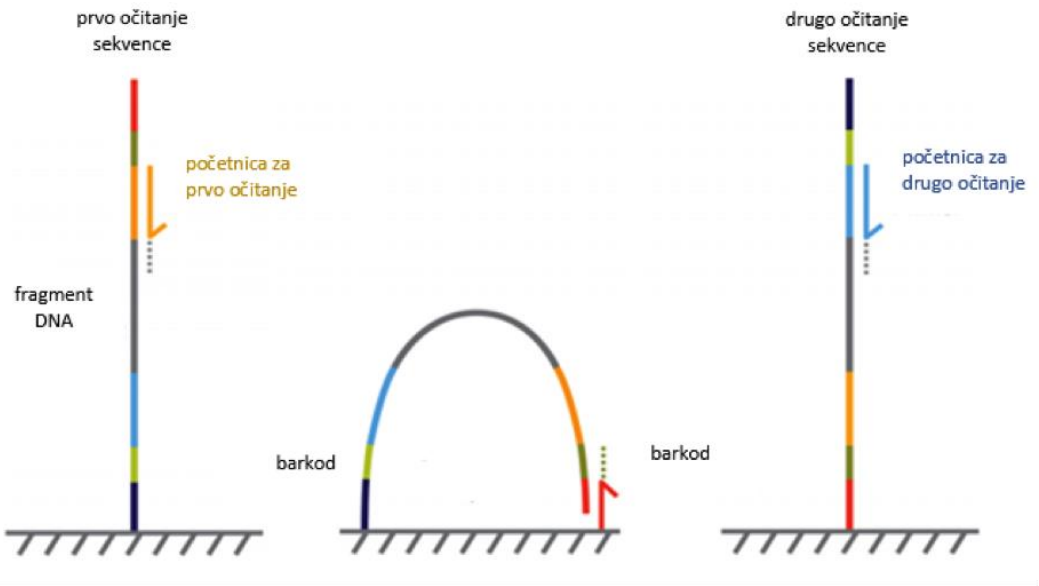
Najkorištenija platforma koja koristi metodu SMRT je bila od tvrtke Pacific Biosciences (PacBio). Instrument koristi posebni čip za sekvenciranje s više tisuća pojedinačnih jažica s

prozirnim dnom koji se nazivaju valovodi (ZMW, od eng. *zero-mode waveguides*). Polimeraza je fiksirana za dno jažice i omogućuje lancu DNA da prolazi kroz ZMW. S obzirom da je mjesto ugradnje konstantno zbog fiksiranog enzima, ugradnja dNTP-a se kontinuirano vizualizira sustavom lasera i kamere koji bilježe boju i trajanje emitirane svjetlosti za vrijeme ugradnje označenog nukleotida na dnu ZMW-a. Polimeraza cijepa fluorofor na dNTP-u prije nego li se ugradi sljedeći dNTP. Platforma SMRT koristi jedinstveni cirkularni kalup koji omogućuje da se svaki kalup sekvencira nekoliko puta (Goodwin i sur. 2016). Uređaji PacBio mogu očitati izrazito duge sekvence do 10 kb duljine, što je posebice korisno za *de novo* sekvenciranje genoma (Heather i Chain, 2016).

2014. godine je razvijen prvi prototip sekvencera *nanopore* – MinION (Oxford Nanopore Technologies). Sekvencer *nanopore* direktno detektira sastav nativne jednolančane molekule DNA (Goodwin i sur. 2016) i omogućuje sekvenciranje vrlo dugih molekula (>10 kb) po nižoj cijeni, ali s puno većom stopom pogrešaka u odnosu na druge sekvencere (Kamps i sur. 2017). Prolazak kroz ionski kanal blokira tok iona čime se smanjuje struja proporcionalno duljini nukleinske kiseline (Heather i Chain, 2016). Promjena u naponu je karakteristična za određenu sekvencu DNA u pori. Uređaj bi mogao sekvencirati od 2 do 4 Tb sekvenci u 2 dana. Također, koristi se takva struktura DNA da se prvo sekvencira nizvodni (eng. *forward*) lanac, a potom uzvodni (eng. *reverse*) lanac te se tako dobije kompletna sekvenca (Goodwin i sur. 2016).

1.4.3. Sekvenciranje sparenih krajeva

Jedan od glavnih napretka u NGS tehnologiji je sekvenciranje sparenih krajeva (PE sekvenciranje, od eng. *paired-end sequencing*) (Slika 10). PE sekvenciranje uključuje sekvenciranje s oba kraja fragmenata DNA te se uzvodno i nizvodno očitani sljedovi poravnavaju kao par (Illumina, 2016). Takvi parovi sekvenca omogućuju točnije poravnavanje s referentnom sekvencom, posebice s područjima koja su bogata ponavljajućim sekvencama, jer se zna udaljenost između svakog sekvenciranog para. Olakšavaju detekciju prekrajanih eksona, indel mutacija i fuzioniranih gena (Heather i Chain, 2016). Također, s obzirom da se točno zna gdje je početak i kraj sekvence, po potrebi je moguće ukloniti PCR-duplikate koji su česti zbog amplifikacije PCR-om za vrijeme pripreme biblioteke i doprinose lažno-pozitivnim rezultatima (Illumina, 2016).



Slika 10. Shema sekvenciranja sparenih krajeva (preuzeto i prilagođeno prema Illumina, 2017).

1.4.4. Obrada podataka

Nakon sekvenciranja, analiziraju se sirovi podaci i kvaliteta očitanih baza. Općenita kontrola kvalitete najčešće se provodi pomoću programa „FastQC“ te su dostupni i mnogi alati kojima se uklanjaju sekvence loše kvalitete, početnice i slično (Kamps i sur. 2017). Kvaliteta očitanih baza smanjuje se prema krajevima sekvenci pa se iz tog razloga krajnje sekvence loše kvalitete uklanjaju kako bi se povećala ukupna kvaliteta podataka (Meldrum i sur. 2011). Određuje se postotak ciljne regije od interesa koja je pravilno obogaćena i sekvencirana. Sekvence se poravnaju (eng. *alignment*) s referentnim genomom i interpretiraju se rezultati cjelokupnog poravnavanja svih očitanih fragmenata (Illumina 2016). Najpoznatiji algoritmi za poravnavanje su „BWA“, „MAQ“, „Bowtie“ i „Novoalign“. Jedna od daljnjih mogućnosti analize je i određivanje nukleotidnih varijanti (eng. *variant calling*) pomoću raznih programskih paketa kao što su „Samtools“, „GATK Unified Genotyper“ i „SOAPsnp“. Nakon očitanja varijanti, rezultati se uspoređuju s bazama podataka (npr. ENSEMBL) kako bi se odredio funkcionalni značaj varijante i eventualno njena povezanost s bolestima (Meldrum i sur. 2011).

1.4.5. Primjena metoda sekvenciranja sljedeće generacije

Danas NGS gotovo u potpunosti zamjenjuje skuplje i dugotrajnije tradicionalno Sangerovo sekvenciranje (Kamps i sur. 2017). NGS platforme se razlikuju ovisno o propusnosti, cijeni, profilu pogrešaka i strukturi očitane sekvence (Goodwin i sur. 2016).

Sekvenciranje kratkih sekvenci ima nisku cijenu po Gb sekvence i visoku točnost (nisku konačnu stopu pogrešaka) te se stoga danas najviše koristi. S druge strane, sekvenciranje dugih sekvenci ima nisku točnost (visoka konačna stopa pogreške) i značajno je skuplja po Gb sekvence, ali omogućuje bolje poravnavanje s bioinformatičkim programima. Stoga je sekvenciranje dugih sekvenci prikladnije za obiteljsko haplotipiziranje, detekciju strukturalnih promjena u DNA i kromosomima, detekciju velikih kromosomskih rearanžmana, translokacija i otkriće novih izoformi transkripata (Kamps i sur. 2017). Platforma Illumina dominira u sekvenciranju kratkih sekvenci, nudi visoku točnost i ima širok raspon primjene: od sekvenciranja genoma i eksoma, primjene u epigenetici, kao što je kromatinska imunoprecipitacija nakon čega slijedi sekvenciranje (ChIP-seq, od eng. *chromatin immunoprecipitation followed by sequencing*), sekvenciranja metilacije DNA (methyl-seq) i primjene u transkriptomici preko sekvenciranja RNA (RNA-seq) (Goodwin i sur. 2016).

Uspjeh prvog ljudskog genoma potaknuo je niz projekata koji su istraživali genom kod različitih oblika i stadija malignih tumora. Dostupnost i pouzdanost sekvenciranja dovela je do integracije genomske znanosti u kliničku primjenu i time je omogućena precizna medicina (Gagan i Van Allen, 2015). 2004. g. Međunarodni projekt genomskog konzorcija (ICGC, od eng. *International Genome Consortium*) je objavio prvi ljudski referentni genom (Metzker, 2010). Projekt Atlas genoma ljudskih malignih tumora (TCGA, od eng. *The Cancer Genome Atlas*) omogućio je stvaranje kataloga koji sadrži na tisuće genoma kod različitih tipova malignih tumora što omogućuje bolje razumijevanje patogeneze bolesti i dovodi do nove ere molekularne patologije i personalizirane medicine (Meldrum i sur. 2011). Sekvenciranjem čitavog genoma (WGS, od eng. *whole genome sequencing*) mnogih organizama napravljena je komparativna analiza velikog razmjera i evolucijske studije (Metzker, 2010). 2010. godine Projekt 1000 genoma (eng. *1000 Genome Project*) omogućio je uvid u varijacije među ljudskom populacijom. Takvo populacijsko sekvenciranje se pokazalo esencijalnim za razumijevanje bolesti ljudi (Goodwin i sur. 2016). 2016. godine, konzorcij sakupljenih eksoma (ExAC, od eng. *The Exome Aggregation Consortium*) objavio je najveći katalog genetičkih varijanti u protein-kodirajućim regijama. Katalog je dobiven analizom 60 706 ljudskih eksoma i takva baza podataka omogućuje veću rezoluciju analize genetičkih varijanti koje su prisutne u niskom udjelu u genomu (Lek i sur. 2017). Kasnije iste godine, objavljena je do sada najveća baza podataka sakupljenih genoma (gnomAD, od eng. *The Genome Aggregation Database*) koja sadrži dvostruko više eksoma od ExAC baze podataka, točnije 123 136 eksoma i 15 496 čitavih genoma. Resekvenciranjem ljudskog genoma omogućeno je bolje razumijevanje kako genetičke razlike utječu na zdravlje ljudi i bolest. Sekvenciranjem transkriptoma (RNA-seq)

identificiraju se i kvantificiraju rijetki transkripti bez prethodnog znanja o određenim genima koji daju informacije o alternativnom izrezivanju i promjenama sekvenci u određenim genima (Metzker, 2010). NGS također omogućuje pogled u regulatorne mehanizme genoma. ChIP-seq služi za otkrivanje interakcija između proteina i molekule DNA, dok se pomoću methyl-seq mogu analizirati modificirane baze, tj. metilirana DNA (Goodwin i sur. 2016).

S medicinskog gledišta je veliki potencijal NGS-a u kliničkoj primjeni kao što je prenatalno testiranje za otkrivanje kromosomskih aneuploidija u fetalnoj DNA, identifikacija rijetkih genetičkih varijanti povezanih s monogenskim Mendelovskim bolestima i uspješna detekcija nasljedne ili somatske mutacije u genima koji dovode do pojave malignih tumora (Meldrum i sur. 2011). U kliničkim primjenama se koristi ciljano sekvenciranje (eng. *targeted sequencing*) zbog visoke pokrivenosti od oko 10 000 X kako bi se potvrdila rijetka mutacija (Goodwin i sur. 2016).

1.4.6. Ciljano sekvenciranje

Ciljanim sekvenciranjem, set gena ili genomskih regija se obogati i sekvencira. Na taj način se omogućuje ušteda vremena i novaca te se analiziraju specifične regije od interesa s puno većom razinom pokrivenosti. Samim time, veća je osjetljivost (eng. *sensitivity*) što omogućuje detekciju genetičkih varijanti koje su prisutne u vrlo niskom udjelu u početnom uzorku (npr. u uzorku genomske DNA iz tumorskog tkiva uklopljenog u parafin, FFPE). Paneli za ciljano sekvenciranje mogu biti komercijalni za analizu određenih gena ili genomskih regija ili se mogu samostalno dizajnirati (eng. *custom design*). Dostupan je veliki broj komercijalnih setova za pripremu biblioteka koji sadrže početnice za analizu specifičnih područja od interesa dok je samostalnim dizajniranjem omogućeno ciljanje specifičnih regija od interesa ovisno o potrebama istraživanja (Illumina, 2016).

Illumina podržava dvije metode ciljanog sekvenciranja – ciljano obogaćivanje (eng. *target enrichment*) i stvaranje amplikona (eng. *amplicon generation*). Ciljanim obogaćivanjem se specifične regije od interesa zarobe hibridizacijom biotiniliranim početnicama nakon čega se izoliraju magnetom. Tom metodom se omogućuje analiza većih genomskih regija od 10 kb do 62 Mb, ovisno o načinu pripreme biblioteke. Metoda stvaranja amplikona obuhvaća amplifikaciju i pročišćavanje regija od interesa korištenjem oligonukleotidnih smjesa za višestruki PCR (eng. *multiplex PCR*). Pogodna je za sekvenciranje manjeg genskog sadržaja, od 2,4 kb do 652,8 kb, i analizu promjena u jednom nukleotidu (Illumina, 2016).

1.5. Gensko testiranje i liječenje karcinoma dojke

Epidemiološki dokazi dosljedno ukazuju da se rizik za oboljenje žene od karcinoma dojke povećava pojavnošću bolesti kod njene majke ili sestre. Što se ranije karcinom dojke dijagnosticira kod njenih rođaka, veći je rizik za njeno oboljenje (Hall i sur. 1990).

Tradicionalno Sangerovo sekvenciranje je znatno skuplje i dugotrajnije od NGS-a. Za mnoge kliničke primjene koriste se genski paneli kako bi se sekvencirao određeni broj gena od interesa. Metoda je isplativa zato što se istovremeno postiže visoka prekrivenost (eng. *coverage*) ciljnih regija od interesa i omogućuje jednostavnija analiza podataka. Dobije se visoka prekrivenost po nukleotidu i visoka specifičnost. Dubina sekvenciranja od 30 X je dovoljna za detekciju nasljednih varijanti koje su prisutne u homozigotnom ili heterozigotnom obliku. U slučaju somatskih mutacija, potrebna je veća pokrivenost (>500) regije od interesa s obzirom da je mutacija prisutna na subklonalnoj razini što rezultira njezinim nižim udjelom pojavljivanja u cjelokupnom uzorku (Kamps i sur. 2017).

Gensko testiranje za mutacije u genima *BRCA1/2* je postao standardni klinički postupak kod žena koje imaju osobnu ili obiteljsku povijest karcinoma dojke ili jajnika. Identifikacija patogene mutacije omogućuje odabir najprikladnije terapijske strategije. 2010. godine objavljen je genski panel za HBOC sindrom koji je nazvan BROCA u čast francuskom doktoru Paulu Broci koji je imaju veliki utjecaj na saznanja o HBOC sindromu. BROCA sadrži 21 gen i omogućuje se visoka osjetljivost detekcije svih tipova mutacija (supstitucije jedne baze, male insercije i delecije, veliki genski rearanžmani). Najčešće se pojavljuju mutacije u genima *BRCA1/2,* ali i ostali geni doprinose nasljednoj predispoziciji za razvoj karcinoma (Kamps i sur. 2017).

Pokazano je da se rizik za razvoj karcinoma među ženama koje su nositelji mutacija u genu *BRCA1* ili *BRCA2* mijenja ovisno o nekim specifičnim genima koji su uključeni u metabolizam spolnih hormona i popravak molekule DNA kao i o hormonskim faktorima (trudnoća, dojenje, kontraceptivi) i životnim stilom. Žene s mutacijama u *BRCA1/2* genima, koje imaju visoki rizik za razvoj karcinoma dojke, mogu se podvrgnuti totalnoj obostranoj profilaktičkoj mastektomiji. Najčešće se podvrgavaju žene koje su nositelji mutacija i imaju obiteljsku pojavnost karcinoma dojke, duktalnog karcinoma *in situ* ili karcinoma jajnika (Teng i sur. 2008). Također, pokazano je i značajno smanjenje razvoja karcinoma dojke kod žena koje su se podvrgnule bilateralnoj profilaktičkoj ooforektomiji, posebice kod nositelja mutacija u *BRCA2*. Iste rezultate pokazao je i tretman tamoksifenom tj. djelomičnim antagonistom estrogena, koji ima najsnažnije djelovanje kod žena mlađih od 50 godina (Petrucelli i sur. 2010).

Za liječenje karcinoma dojke koriste se različiti kemoterapeutici. Jedni od njih su taksani - spojevi koji stabiliziraju mikrotubule i tako blokiraju staničnu proliferaciju i dovode do apoptoze. Najčešći taksani upotrebljavani za liječenje karcinoma dojke su docetaksel i paklitaksel. Također, upotrebljavaju se i spojevi platine koji se direktno vežu na DNA i time dovode do unakrsnog povezivanja unutar lanaca molekule DNA ili dvolančanog loma. Najčešće se upotrebljavaju spojevi cisplatin i karboplatin. Jedna od strategija liječenja karcinoma dojke, koja je još u fazi testiranja, je sintetička letalna terapija koja podrazumijeva korištenje PARP inhibitora. Poli (ADP-riboza) polimeraze (PARP) su enzimi bitni za mehanizme popravka oštećenja molekule DNA mehanizmom izrezivanja baze. Sintetiziraju ADP-ribozni polimer na mjestu oštećenja koji služi za privlačenje kompleksa za popravak (Godet i Gilkes, 2017). PARP inhibitori blokiraju popravak izrezivanjem baze i dovode do akumulacije jednolančanih lomova koji mogu dovesti do dvolančanih lomova koji se u zdravim stanicama popravljaju homolognom rekombinacijom. Tumorske stanice, koje imaju narušenu funkciju proteina BRCA1/2, tretirane PARP inhibitorima imaju narušenu i homolognu rekombinaciju i popravak izrezivanjem baza što je letalno za takve stanice (Pecorino, 2012).

Ovisno o molekularnoj različitosti tumora, različit je pristup liječenju kao i odgovor na liječenje. Tumori nalik bazalnim stanicama imaju najlošiju prognozu, dok luminalni A tip ima najbolju prognozu (Polyak, 2014).

2. CILJ ISTRAŽIVANJA

Cilj ovog istraživanja je samostalna konstrukcija panela za umnažanje fragmenata gena *BRCA1/2* kao pripreme biblioteke za NGS. Konstruirane su početnice za umnažanje kodirajućih regija navedenih gena i optimizirani su uvjeti njihova umnažanja višestrukim PCR-om. Pripremljena je biblioteka na temelju odabranih optimalnih uvjeta umnažanja na uzorku genomske DNA iz krvi i tkiva uklopljenog u parafin i provedena je reakcija sekvenciranja. Dobiveni podaci obrađeni su bioinformatičkim programskim paketima i uspoređena je uspješnost samostalno konstruiranog panela s komercijalnim testom za analizu gena *BRCA1/2*.

3. MATERIJALI I METODE

3.1. Materijali

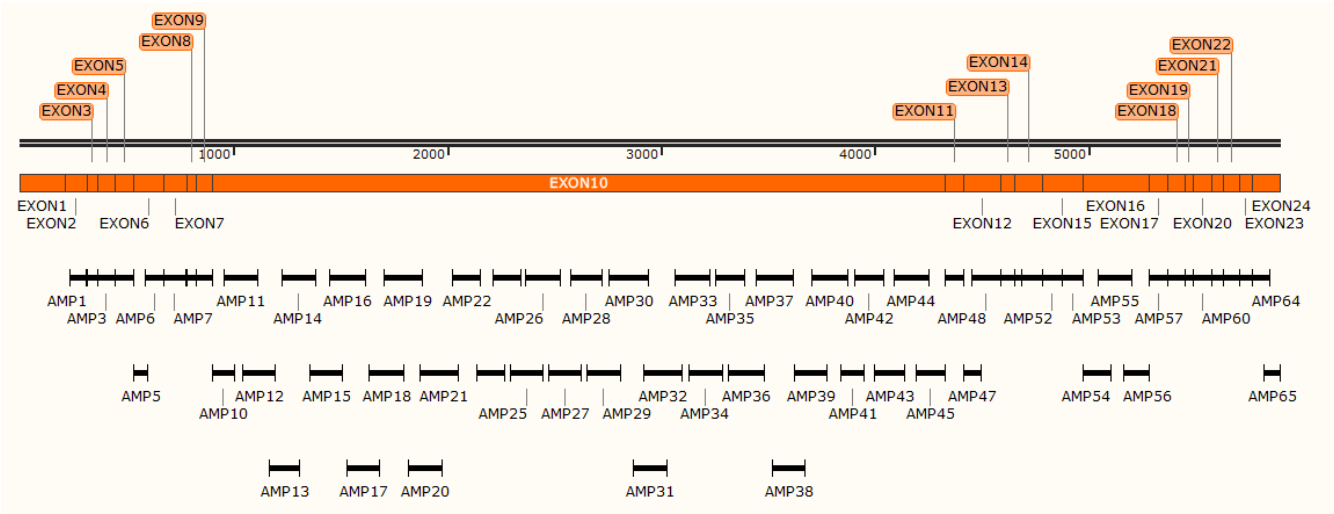
U ovom istraživanju korišten je jedan klinički uzorak krvi za koji postoji sumnja na nasljedni oblik karcinoma dojke i jedan uzorak tumorskog tkiva oboljele osobe uklopljen u parafin (FFPE, od eng. *formalin-fixed paraffin embedded tissue*).

3.2. Metode

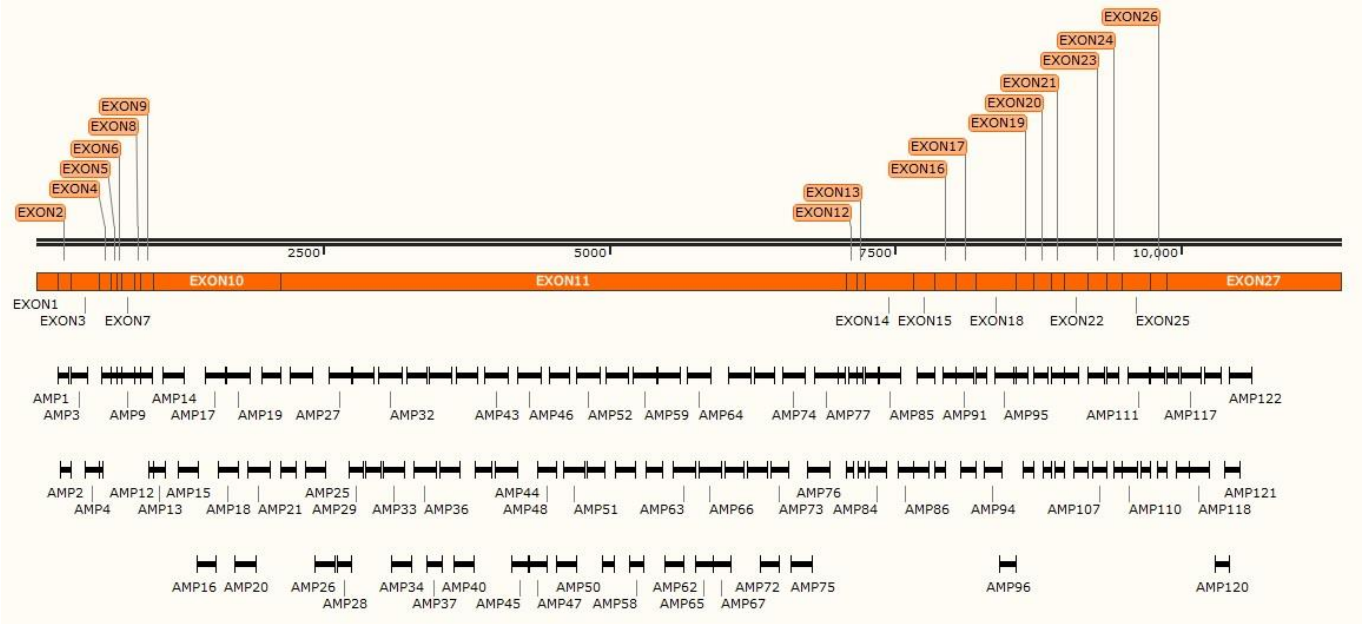
3.2.1. Konstrukcija početnica

Početnice za umnažanje kodirajućih sljedova gena *BRCA1/2* konstruirane su programom za dizajn početnica „Optimus Primer“ (http://op.pgx.ca/). Unosom oznake gena ili koordinata regije od interesa te odgovarajućih parametara za dizajn početnica, kao što su veličina početnica, temperatura mekšanja, veličina konačnog produkta, sadržaj parova GC baza, omogućen je brzi dizajn velikog broja početnica potrebnih za sekvenciranje ciljnih regija. Kako bi se dizajnirale početnice za sve gene ili regije, programski alat koristi 4 ciklusa ponavljanja dizajna početnica gdje se u svakom idućem ciklusu postupno modificiraju parametri dizajna početnica tako da su zadani parametri manje strogi od prethodnih.

Za dizajn početnica za gene *BRCA1* i *BRCA 2*, uneseni su jedinstveni identifikatori gena iz baze podataka NCBI (NG_005905.2 za *BRCA1* i NG_012772 za *BRCA2)* te je zadan parametar željene veličine amplikona od 120 pb. Program je dao listu početnica (Prilog 1 i 2) za umnažanje kodirajućih regija gena koje su provjerene i vizualizirane programom „SnapGene“ (Slika 11 i 12). Ukupno je konstruirano 130 početnica za umnažanje 65 amplikona gena *BRCA1* i 244 početnica za 122 amplikona gena *BRCA2*.



Slika 11. Mapa eksona i amplikona gena *BRCA1* („SnapGene“).



Slika 12. Mapa eksona i amplikona gena *BRCA2* („SnapGene“).

3.2.2. Izolacija genomske DNA

Komercijalno dostupnim kemikalijama, izolirana je genomska DNA. Za izolaciju DNA iz krvi korištene su kemikalije iz DNeasy® Blood & Tissue Kit, QIAGEN, prema protokolu proizvođača (Qiagen, 2006). Ukratko, uzorak krvi liziran je korištenjem proteinaze K, nakon čega je lizat nanesen na DNeasy Min spin kolonu. Centrifugiranjem je DNA selektivno vezana

na membranu i odvojena od ostatka staničnog lizata. Preostala onečišćenja i inhibitori enzima uklonjeni su dvostrukim ispiranjem etanolom te je DNA eluirana u ultra-čistoj vodi

Za izolaciju DNA iz tumorskog tkiva uklopljenog u parafin korišten je komercijalno dostupan komplet GeneRead^{TM} DNA FFPE Kit, QIAGEN, prema protokolu proizvođača (Qiagen, 2014). Postupak je uključivao uklanjanje parafina, liziranje stanica proteinazom K i uklanjanje križnih veza nastalih u molekuli DNA zbog djelovanja formalina. Nakon njihovog uklanjanja, enzimom uracil-N-glikozilazom uklonjeni su deaminirani citozinski ostaci, odnosno umjetno nastali uracili u molekuli DNA. Potom se DNA centrifugiranjem vezala na QIAamp MinElute kolonu, isprana su onečišćenja i DNA je eluirana u ultra-čistoj vodi.

3.2.3. Određivanje koncentracije DNA u otopinama

Koncentracija izolirane DNA određena je koristeći fluorimetrijsku metodu kvantifikacije koncentracije DNA, Qubit^{TM} 3.0 (Thermo Fisher Scientific) te su uzorci pohranjeni na -20 ^{o}C.

Za mjerenje se koriste fluorescentni reagensi koji se specifično vežu na molekulu DNA uslijed čega dolazi do emitiranja zračenja. Klasični UV-spektrofotometri mjere apsorbanciju pri 260 nm, ali nisu selektivni i ne mogu razlikovati potječe li apsorbancija od molekule DNA, RNA ili proteina. Na mjerenje veliki utjecaj imaju i onečišćenja kao što su slobodni nukleotidi i soli. Fluorimetar mjeri emitirano zračenje tek kada se fluorescentna boja veže za molekulu DNA što smanjuje efekt raznih onečišćenja na rezultat mjerenja. Postoje različiti komercijalni kompleti od kojih je svaki specifičan za kvantifikaciju pojedinog analita - DNA, RNA ili proteina. Također, koristi se mali volumen (1-20 µL) uzorka za mjerenje.

Za određivanje koncentracija otopina DNA prethodno je potrebno kalibrirati uređaj standardnim otopinama koje se pripremaju miješanjem reagensa iz komercijalnog kompleta Qubit® dsDNA BR Assay Kit (Thermo Fisher Scientific, 2015). Pripremi se radna otopina razrjeđenjem fluorescentnog reagensa i pufera u omjeru 1:200. U tubicama volumena 0,5 mL, pomiješa se 10 µL otopine standarda 1 sa 190 µL radne otopine te se isto učini i s otopinom standarda 2. Uređaj se kalibrira mjerenjem koncentracije standardnih otopina. Za mjerenje koncentracije uzoraka nepoznate koncentracije, pomiješa se 199 µL radne otopine i 1 µL otopine uzorka te se na uređaju očita masena koncentracija otopina.

3.2.4. Pojedinačne PCR-reakcije

Uspješnost konstrukcije početnica provjerena je lančanom reakcijom polimerazom (PCR) kojom se željeni odsječak DNA umnaža u uvjetima *in vitro* kako bi se dobila dovoljna količina fragmenata DNA za pripremu biblioteke. Metoda zahtijeva termostabilnu DNA polimerazu i DNA početnice specifične za regije DNA od interesa. Reakcija se sastoji od tri faze - faze denaturacije dvolančane DNA u dva jednolančana lanca, faze sparivanja početnica i faze produljivanja lanaca DNA. Kontinuiranim ponavljanjem tih faza, količina umnožene DNA eksponencionalno raste.

Provedena je PCR-reakcija na uređaju za PCR (The Biometra TProfessional Thermocycle, THISTLE SCIENTIFIC LTD) svakog pojedinog para početnica na uzorku genomske DNA iz krvi i tumorskog tkiva uklopljenog u parafin (DNA iz FFPE-a). U pločicama s 96 bunarčića zasebno je napravljeno 65 reakcija provjera početnica za gen *BRCA1* i 122 reakcije provjera početnica za gen *BRCA2*.

Za svaki par početnica napravljena je PCR-reakcija koja je sadržavala genomsku DNA izoliranu iz krvi ili genomsku DNA iz FFPE-a kao kalup, Taq DNA polimerazu, uzvodnu i nizvodnu početnicu, PCR pufer te ultra-čistu vodu. Kao pozitivna kontrola PCR-reakcije, korištene su provjerene početnice za dio gena *SAHH* (S-adenozilhomocistein-hidrolaze) veličine 193 pb dok je kod negativne kontrole izostavljen dodatak genomske DNA. Uvjeti PCR-reakcije kao i konačne koncentracije sastavnih komponenata prikazani su u Tablicama 2 i 3.

Reakcije s nastalim nespecifičnim produktima su ponovljene podizanjem temperature sparivanja početnica na 57 °C, dok je kod PCR-reakcija za uzorak genomske DNA iz FFPE-a temperatura sparivanja početnica postavljena na 57 °C.

Tablica 2. Sastavne komponente i njihove konačne koncentracije korištene za provjeru pojedinačnog para početnica.

	Konačna koncentracija
5 x ALLin™ PCR-pufer	1 x
10 µM uzvodna početnica	0,3 µM
10 µM nizvodna početnica	0,3 µM
genomska DNA iz krvi/FFPE-a	75 ng/100 ng
5 U/µL ALLin™ Taq DNA polimeraza	1 U
mqH_2O	do 25 µL

Tablica 3. Temperaturni program korišten za provjeru pojedinačnog para početnica.

	temperatura / °C	trajanje ciklusa	broj ciklusa
početna denaturacija	95	2 min	1
denaturacija	95	15 sec	30
sparivanje početnica	55/57	15 sec	
produljenje	72	10 sec	
završno produljenje	72	2 min	1

3.2.5. Elektroforeza u agaroznom gelu

Elektroforeza u agaroznom gelu je metoda koja se koristi za razdvajanje, identifikaciju i pročišćavanje nukleinskih kiselina na temelju njihovih veličina, odnosno masa. Temelji se na različitom gibanju molekula nukleinskih kiselina kroz pore agaroznog gela pod utjecajem električnog polja pri čemu se manje molekule kreću brže od većih molekula. Ovisno o koncentraciji agaroze u samom gelu, može se kontrolirati veličina pora, odnosno raspon molekulskih masa nukleinskih kiselina koje se mogu odjeliti. Za detekciju nukleinskih kiselina unutar gela, koriste se interkalirajući reagensi kao što je etidij-bromid. Vezanjem na DNA, intenzitet fluorescencije etidij-bromida znatno se povećava te je vizualizacija fragmenata u gelu moguća pobudom UV-svjetlom transiluminatora (λ=312 nm).

Rezultati PCR-reakcija analizirani su elektroforezom u 2%-tnom agaroznom gelu uz upotrebu boje etidij-bromid. Gelovi su pripremljeni otapanjem 2 g agaroze u 100 mL TAE pufera uz dodatak 5 µL otopine etidij-bromida, zagrijavanjem u mikrovalnoj pećnici do vrenja i izlijevanjem u odgovarajuće kalupe. U jažice je naneseno 5 µL uzorka uz dodatak 1 µL boje za nanošenje uzoraka - Gel Loading Dye, Purple 6x (NEB). Kao marker molekulskih masa, korišten je Quick-Load 2-Log DNA Ladder (NEB). Elektroforeza je provedena u trajanju od 20 minuta uz napon od 100 V korištenjem aparaturu za horizontalnu elektroforezu (Bio-Rad). Umnoženi fragmenti su vizualizirani osvjetljivanjem pomoću transiluminatora (UVItec Ultravue Transilluminator, Progen Scientific).

3.2.6. Optimizacija reakcija višestrukog PCR-a

Veći broj početnica pomiješan je u jednoj PCR-reakciji kako bi se proveo višestruki PCR. Kod višestrukog PCR-a, koristi se više parova početnica s ciljem istovremenog umnažanja više odsječaka DNA. Višestruki PCR je teže optimizirati, ali omogućuje uštedu

vremena i kemikalija. Ključan korak je dizajn početnica sa sličnom temperaturom sparivanja kako bi u jednoj reakciji nastali svi željeni produkti uz minimalan prinos nespecifičnih produkata.

Početnice za oba gena grupirane su u 4 smjese na način da se uzeo svaki četvrti par početnica kako bi najbliži susjedni parovi početnica bili što udaljeniji jedni od drugih na kalupu. Pripremljene pojedine reakcije višestrukog PCR-a sadržavale su različitu količinu genomske DNA izolirane iz krvi ili FFPE-a, 5 U ALLin™ Taq DNA polimeraze, različitu konačnu koncentraciju smjese početnica, 1 x ALLin™ PCR-pufer (5 x) i ultra-čistu vodu. Kako bi se pronašli optimalni uvjeti višestrukog PCR-a, proveden je veći broj višestrukih PCR-reakcija s promjenama u količini genomske DNA, konačnoj koncentraciji smjese početnica, broju ciklusa PCR-reakcija, temperaturi i vremenu sparivanja početnica te vremenu produživanja početnica (Tablica 4 i Tablica 5). Dodatno, napravljene su reakcije gdje su korištene dvije smjese početnica u jednoj reakciji. Kao pozitivna kontrola PCR-reakcije, korištene su provjerene početnice za dio gena *SAHH* (S-adenozilhomocistein-hidrolaze) veličine 193 pb.

Tablica 4. Optimizacija višestrukog PCR-a za genomsku DNA izoliranu iz krvi.

Broj reakcije	Količina gDNA krv	Koncentracija početnica	Broj ciklusa PCR-a	Temperatura i vrijeme sparivanja početnica	Vrijeme produživanja početnica
1	150 ng	ukupno 0,2 µM	20	57 °C, 15 sec	10 sec
2	150 ng	ukupno 0,4 µM	25	57 °C, 15 sec	10 sec
3	220 ng	ukupno 0,4 µM	30	57 °C, 15 sec	10 sec
4	150 ng	svaka 0,15 µM	35	57 °C, 15 sec	10 sec
5	150 ng	svaka 0,15 µM	30	59 °C, 15 sec	10 sec
6	150 ng	svaka 0,15 µM	30	60 °C, 30 sec	10 sec
9	50 ng	svaka 0,15 µM	30	57 °C, 15 sec	10 sec
10	50 ng	svaka 0,15 µM	25	57 °C, 15 sec	10 sec
11	50 ng	svaka 0,15 µM	20	57 °C, 15 sec	10 sec

Tablica 5. Optimizacija višestrukog PCR-a za genomsku DNA izoliranu iz tumorskog tkiva uklopljenog u parafin.

Broj reakcije	Količina gDNA FFPE	Koncentracija početnica	Broj ciklusa PCR-a	Temperatura i vrijeme sparivanja početnica	Vrijeme produživanja početnica
12	200 ng	svaka 0,15 µM	35	57 °C, 15 sec	10 sec
13	200 ng	svaka 0,15 µM	35	57 °C, 15 sec	1 min
14	200 ng	svaka 0,15 µM	35	57 °C, 30 sec	1 min
15	200 ng	svaka 0,15 µM	25	57 °C, 30 sec	1 min
16	200 ng	svaka 0,15 µM	20	57 °C, 30 sec	1 min
17	25 ng	svaka 0,15 µM	35	57 °C, 30 sec	1 min
18	50 ng	svaka 0,15 µM	35	57 °C, 30 sec	1 min
19	100 ng	svaka 0,15 µM	35	57 °C, 30 sec	1 min

FFPE: tumorsko tkivo uklopljeno u parafin, od eng. *formalin-fixed paraffin embedded tissue*

Svaka PCR-reakcija je vizualizirana elektroforezom u 2%-tnom ili 3%-tnom agaroznom gelu. U jažice je naneseno po 15 µL uzorka i 3 µL boje za nanošenje. Elektroforeza je provedena u trajanju od 30 minuta uz napon od 80 V. Kao marker molekulskih masa, korišten je Quick-Load 2-Log DNA Ladder (NEB).

3.2.7. Reakcije višestrukog PCR-a za sekvenciranje sljedeće generacije

Na temelju provedenih reakcija optimizacije višestrukog PCR-a, izabrana su 3 različita reakcijska uvjeta za uzorak krvi i FFPE-a (Tablica 6) te su napravljene po 4 reakcije višestrukog PCR-a pri svakom uvjetu za svaku od 4 smjese početnica. Svaka reakcija je pročišćena pomoću magnetskih zrnaca MagSi-NGS^{PREP} te je izmjerena koncentracija pročišćenih reakcija na Qubit^{TM} 3.0. PCR-reakcije na uzorku genomske DNA iz krvi pri 25 ciklusa PCR-a su ukoncentrirane na uređaju SpeedVac^{TM} (Thermo Fisher Scientific) kako bi im se povećala koncentracija.

Reakcije višestrukog PCR-a, koje su provedene pri istim uvjetima, no s različitim smjesama početnica, pomiješane su u jednakim omjerima DNA te je konačno dobiveno 6 smjesa reakcija višestrukog PCR-a koje su ukupno sadržavale 50 ng DNA (Tablica 6).

Tablica 6. Reakcije višestrukog PCR-a na uzorku genomske DNA iz krvi i FFPE-a za NGS.

Naziv reakcije	Količina gDNA krv/gDNA FFPE	Koncentracija početnica	Broj ciklusa PCR-a	Temperatura i vrijeme sparivanja početnica	Vrijeme produživanja početnica	Koncentracija produkata /ngµL^{-1}			
						S1	S2	S3	S4
Biblioteka 1	50 ng (DNA iz krvi)	svaka 0,15 µM	35	57 °C, 15 sec	10 sec	9,7	5,3	11,2	4,96
Biblioteka 2	50 ng (DNA iz krvi)	svaka 0,15 µM	25	57 °C, 15 sec	10 sec	1,19	1,16	3,02	2,82
Biblioteka 3	50 ng (DNA iz krvi)	svaka 0,15 µM	35	57 °C, 15 sec	10 sec	2,7		4,28	
Biblioteka 4	50 ng (DNA iz FFPE-a)	svaka 0,15 µM	35	57 °C, 30 sec	1 min	27,4	11,9	14,0	22,4
Biblioteka 5	50 ng (DNA iz FFPE-a)	svaka 0,15 µM	45	57 °C, 30 sec	1 min	9,04	4,54	6,74	10,5
Biblioteka 6	50 ng (DNA iz FFPE-a)	svaka 0,15 µM	35	57 °C, 30 sec	1 min	3,1		4,2	

FFPE: tumorsko tkivo uklopljeno u parafin, od eng. *formalin-fixed paraffin embedded tissue*
NGS: sekvenciranje sljedeće generacije, od eng. *next-generation sequencing*

3.2.8. Pročišćavanje produkata višestrukog PCR-a

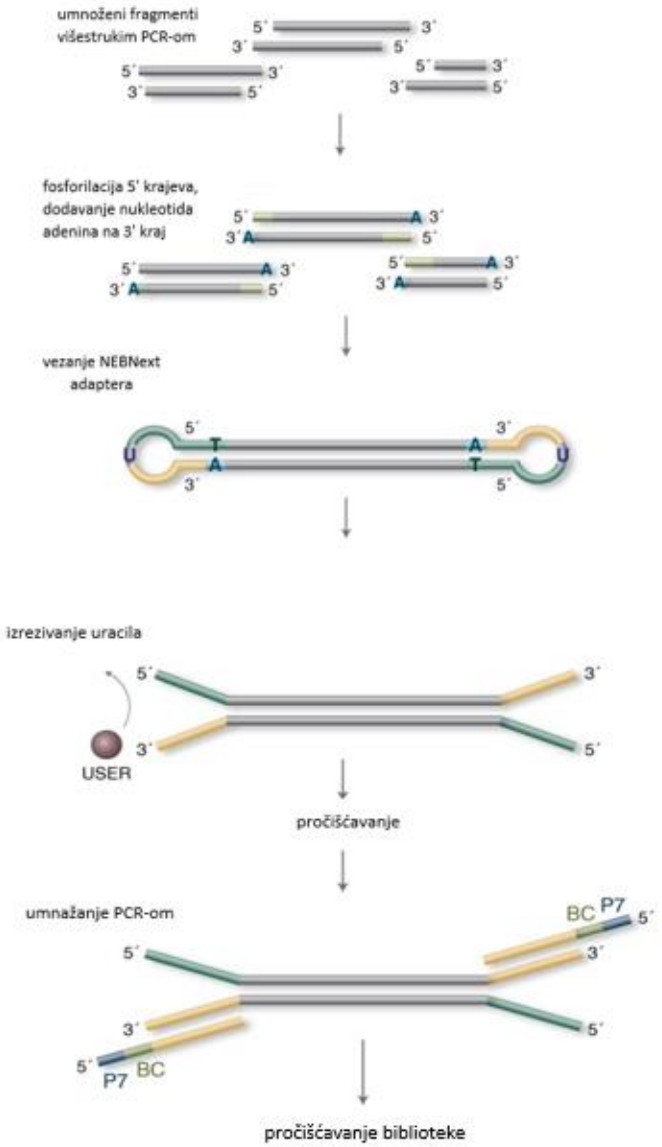
Za pročišćavanje produkata reakcija višestrukog PCR-a korištena su magnetska zrnca MagSi-NGS^{PREP} (AMSBIO) pomoću kojih se uspješno pročišćavaju dvolančani DNA fragmenti od preostalih komponenti PCR-reakcije. Fragmenti DNA vežu se na površinu magnetskih zrnaca ostavljajući neiskorištene nukleotide, početnice, dimere početnica i druga onečišćenja u otopini. Fragmenti se potom eluiraju s magnetskih zrnca u TE puferu ili ultra-čistoj vodi. Ova metoda pročišćavanja DNA omogućuje i pročišćavanje fragmenata u odabranom rasponu veličina ovisno o omjeru dodanog volumena reagensa i volumena uzorka. Povećanjem volumena reagensa, povećava se efikasnost vezanja manjih fragmenata te obrnuto.

3.2.9. Priprema biblioteka za sekvenciranje

Biblioteke su pripremljene pomoću komercijalno dostupnih kemikalija NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England Biolabs, 2016) (Slika 13). Postupak je uključivao fosforilaciju 5' krajeva umnoženih fragmenata iz višestrukog PCR-a i dodavanje nukleotida adenina na 3' kraj (eng. *dA-tailing*). Potom su na krajeve vezani adapteri NEBNext u formi ukosnice, kako bi se smanjilo stvaranje dimera adaptera, i koji na 5' kraju imaju viseći nukleotid timin koji omogućuje to vezanje. Pomoću enzima USER provedeno je

izrezivanje uracila iz ukosnice kako bi se dobili dvolančani slobodni krajevi. Ostatak nevezanih adaptera pročišćen je pomoću magnetskih zrnaca MagSi-NGS^{PREP} i proveden je završni PCR kako bi se dobila zadovoljavajuća količina DNA potrebna za reakciju sekvenciranja. Početnice za PCR su bile specifične za dijelove adaptera i sadržavale su barkodove koji su služili kao referentne točke tijekom analize.

Dobivene biblioteke pročišćene su magnetskim zrncima MagSi-NGS^{PREP} te je napravljen odabir veličine fragmenata s lijeve strane (eng. *left side size selection*) kako bi se uklonili dimeri adaptera nastali u procesu pripreme biblioteke.



Slika 13. Shema pripreme biblioteke pomoću komercijalno dostupnih kemikalija NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®.

3.2.10. Određivanje veličine fragmenata dobivenih produkata

U svrhu određivanja veličine fragmenata dobivenih produkata korišten je napredni sustav gel-elektroforeze pomoću uređaja – Agilent 2100 Bioanalyzer (Agilent Technologies). Uređaj omogućuje preciznije i točnije određivanje veličine fragmenata od klasične elektroforeze u agaroznom gelu kao i njegovu istovremenu kvantifikaciju te procjenu čistoće uzoraka pri čemu se koristi minimalna količina uzoraka (1 μL). Postupak je u potpunosti automatiziran te se odvija na DNA-čipu visoke osjetljivosti (eng. *high sensitivity DNA chip*) koji sadrži mrežu povezanih rezervoara za otopine uzoraka i molekularnih markera veličina i mikrokanala koji se pune smjesom gela i fluorescentne boje. Fluorescentna boja se interkalira u molekulu DNA te se kvantifikacija provodi na temelju intenziteta fluorescencije. Istovremeno se može analizirati 11 uzoraka. Rezultati se mogu prikazati u obliku gela, elektroferograma ili tablice.

Pročišćene biblioteke analizirane su na DNA-čipu visoke osjetljivosti kako bi se procijenila njihova kvaliteta i koncentracija.

3.2.11. Sekvenciranje uzoraka

Pripremljene biblioteke su objedinjene te se prema protokolu proizvođača (Denature and Dilute Libraries Guide, MiniSeq System (Illumina, 2016) dobivena smjesa denaturirala pomoću NaOH i razrijedila do konačnog volumena od 500 μL i koncentracije 1,55 pM.

Sekvenciranje uzoraka provedeno je na platformi MiniSeq™ System, Illumina prema prethodno opisanom postupku u poglavlju 1.4.2.1.

3.2.12. Obrada rezultata

Dobiveni podaci su obrađeni i interpretirani bioinformatičkim programskim paketima. Pomoću programa „BWA Alligner“ (eng. Burrows-Wheeler Aligner) (v.0.7.12) (Li i Durbin, 2009), kratke sekvence su poravnate (eng. *alignment*) prema referentnom genomu. U ovom slučaju korišten je ljudski referentni genom (hg19). Poravnate sekvence spremaju se u jedinstvenu datoteku koja sadrži niz statističkih podataka o kvaliteti poravnanja i očitanih sekvenci.

Pomoću programa „BAMClipper“ (v.1.1.1) (Au i sur. 2017) izuzeti su dijelovi sekvenci koji pripadaju početnicama korištenima u višestrukom PCR-u kako bi se izbjeglo prekrivanje

potencijalnih mutacija sekvencama samih početnica. Na taj način početnice ne ulaze u daljnju analizu te je povećana točnost detekcije genetičkih varijanti.

„SAM Tools“ (SAM, od eng. *Sequence Alignment/Map*) (v.0.1.19) (Li i sur. 2009) je bioinformatički programski paket koji sadrži različite programe za obradu i analizu poravnatih sekvenci kao što je njihovo sortiranje, spajanje, označavanje i slaganje. Ovim programskim paketom moguće je dobiti informacije o pokrivenosti regije od interesa na razini svake pojedine baze, odnosno, koordinate referentnog genoma.

Za detekciju SNV-ova i malih insercija/delecija korišten je program „VarScan“ (v.2.3.9) (Koboldt i sur. 2009). Program omogućuje detekciju statistički značajnih varijanti na temelju zadanog praga (eng. *threshold*) pokrivenosti, udjela sekvenci s mutacijom i kvalitete baza.

Za vizualizaciju i interaktivno istraživanje velikih genomskih skupova podataka korišten je program „IGV“ (eng. *Integrative Genomics Viewer*) (Robinson i sur. 2011). Programom „Qualimap“ (v.2.2.1) (García-Alcalde i sur., 2012) dobiveni su podaci o broju i kvaliteti poravnatih sekvenci. Moguće je analizirati pokrivenost i broj očitanih sekvenci (eng. *read counts*) genskog područja od interesa.

3.2.13. Usporedba s komercijalno dostupnim testovima za analizu mutacija gena *BRCA1* i *BRCA2*

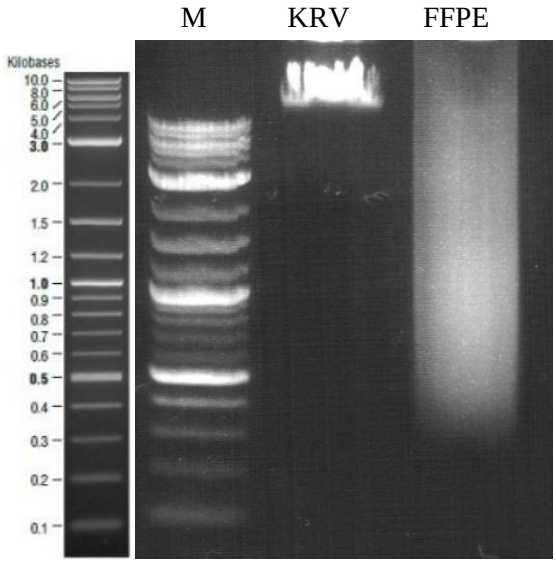
Dobiveni rezultati uspoređeni su s komercijalno dostupnim testom za analizu mutacija gena *BRCA1* i *BRCA2* NEXTflex™ *BRCA1* & *BRCA2* Amplicon Panel for FFPE lllumina® Sequencing (Bioo Scientific, 2016). Test sadrži 260 parova početnica u 4 smjese početnica koje omogućuju amplifikaciju i sekvenciranje svih kodirajućih eksona gena *BRCA1* i *BRCA2*. Amplikoni ciljnih regija su veličine 64-129 pb, dok su regija od interesa zajedno s početnicama veličine 122-196 pb. Ciljne regije se amplificiraju u prvom PCR-u, nakon čega slijedi popravak krajeva i vezanje adaptera, a potom se u drugom PCR-u umnažaju regije od interesa i uvode barkodovi.

Isti uzorci koji su korišteni u ovom eksperimentu prethodno su sekvencirani na platformi MiniSeq™ System, Illumina uz upotrebu komercijalnog testa NEXTflex™ *BRCA1* & *BRCA2* Amplicon Panel for FFPE lllumina® Sequencing. Dobiveni podaci su analizirani bioinformatičkim programskim paketima kako bi se omogućila usporedba dizajniranog panela i komercijalno dostupnog testa.

4. REZULTATI

4.1. Izolacija genomske DNA

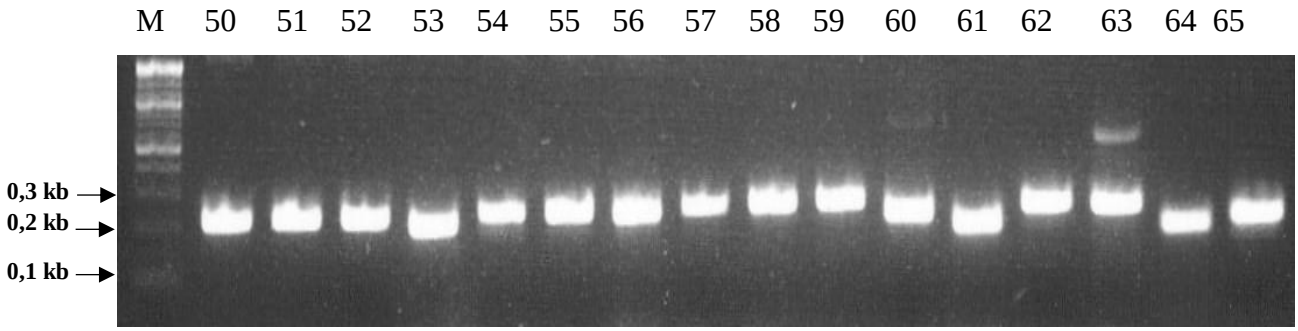
Genomska DNA je izolirana iz 200 µL krvi i uzorka tumorskog tkiva uklopljenog u parafin. Koncentracija je izmjerena fluorimetrijskom metodom Qubit™ 3.0 te je iznosila 16 ng/µL DNA za uzorak krvi, dok je koncentracija izolirane DNA iz tumorskog tkiva uklopljenog u parafin iznosila 287 ng/µL. Napravljena je elektroforeza u 1%-tnom gelu uz upotrebu etidij-bromida za vizualizaciju (Slika 14).



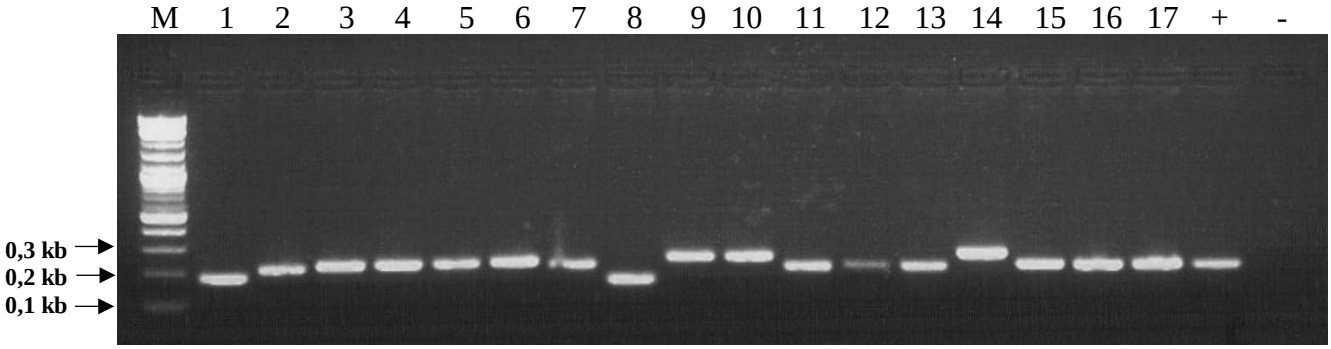
Slika 14. Gel-elektroforeza izolirane genomske DNA iz krvi i tumorskog tkiva uklopljenog u parafin. Na gel je nanesen marker molekulskih masa (M), izolirana genomska DNA iz krvi i tumorskog tkiva uklopljenog u parafin (FFPE).

4.2. Pojedinačne PCR-reakcije

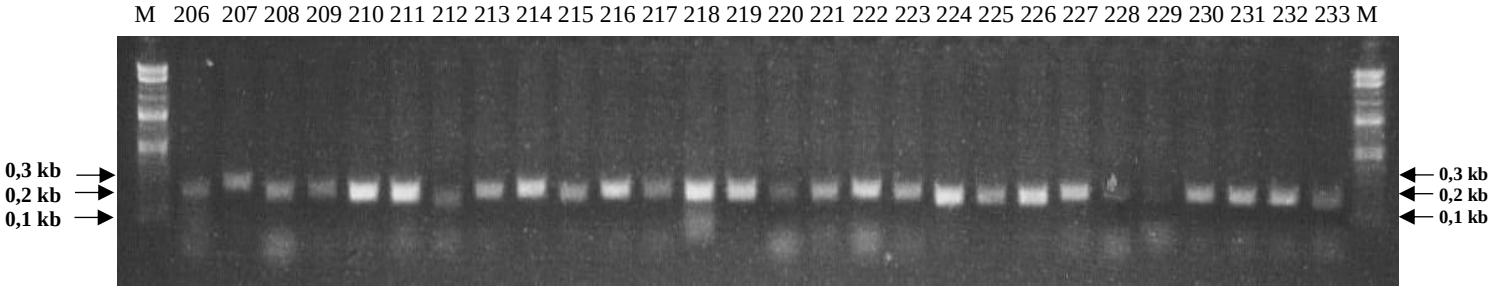
Provjera konstruiranih početnica provedena je zasebnim PCR-reakcijama svakog pojedinog para početnica na uzorku genomske DNA iz krv i na uzorku genomske DNA iz FFPE-a. Ukupno je pripremljeno 374 PCR-reakcija od čega je za svaki uzorak napravljeno 65 reakcija provjera početnica za gen *BRCA1* te 122 reakcije provjera početnica za gen *BRCA2*. Reakcije su analizirane elektoroforezom u 2%-tnom agaroznom gelu uz upotrebu etidij-bromida za vizualizaciju produkata reakcija. Na Slikama od 15 do 18 prikazani su PCR-produkti pojedinih reakcija.



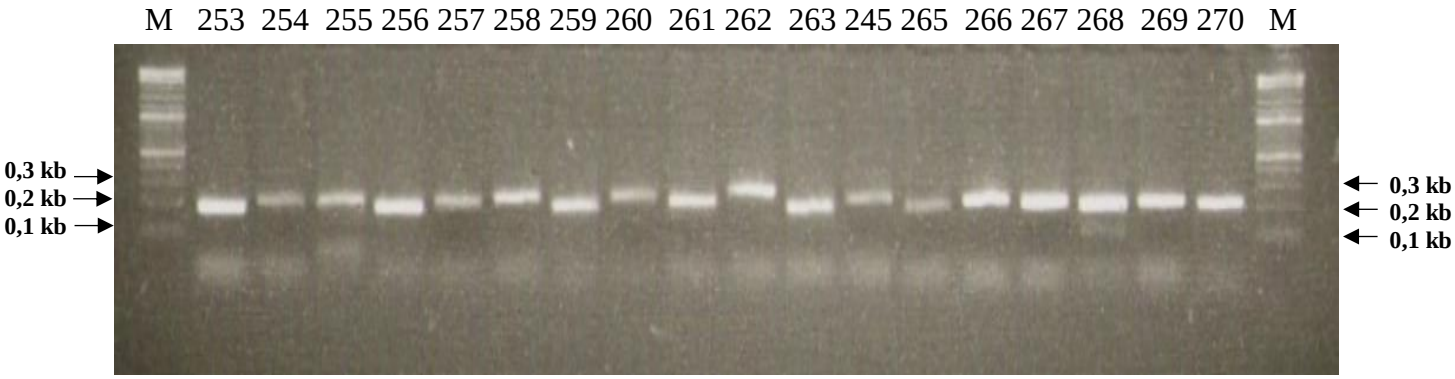
Slika 15. Gel-elektroforeza PCR-produkata gena *BRCA1* na uzorku genomske DNA iz krvi. Na gel je nanesen marker molekulskih masa (M) i produkti PCR-reakcije (50-65).



Slika 16. Gel-elektroforeza PCR-produkata gena *BRCA2* na uzorku genomske DNA iz krvi. Na gel je nanesen marker molekulskih masa (M), produkti PCR-reakcije (1-17), pozitivna kontrola (+) i negativna kontrola (-).



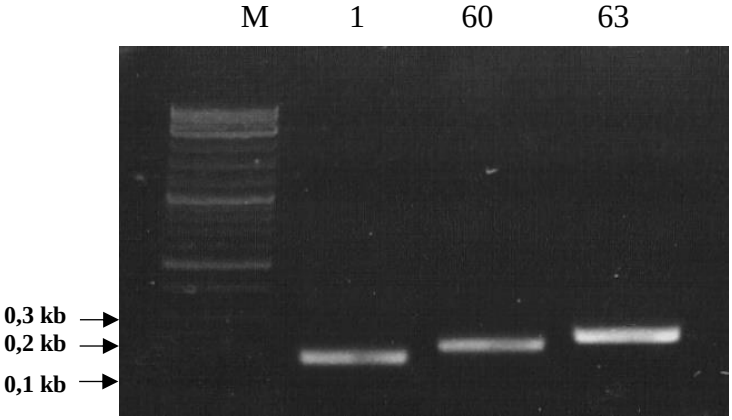
Slika 17. Gel-elektroforeza PCR-produkata gena *BRCA1* na uzorku genomske DNA iz FFPE-a. Na gel je nanesen marker molekulskih masa (M) i produkti PCR-reakcije (206-233).



Slika 18. Gel-elektroforeza PCR-produkata gena *BRCA2* na uzorku genomske DNA iz FFPE-a. Na gel je nanesen marker molekulskih masa (M) i produkti PCR-reakcije (253-270).

Veličina produkta pojedine PCR-reakcije uspoređena je s očekivanom veličinom produkta kako bi se provjerilo nastaje li PCR-reakcijama željeni fragment molekule DNA.

Također, u pojedinim reakcijama bio je vidljiv nastanak nespecifičnog produkta (Slika 15, reakcije 60 i 63). Podizanjem temperature sparivanja početnica s 55 ^{o}C na 57 ^{o}C ponovljene su reakcije 1, 60 i 63 gena *BRCA1* na uzorku genomske DNA iz krvi s ciljem eliminacije nespecifičnih produkata (Slika 19).



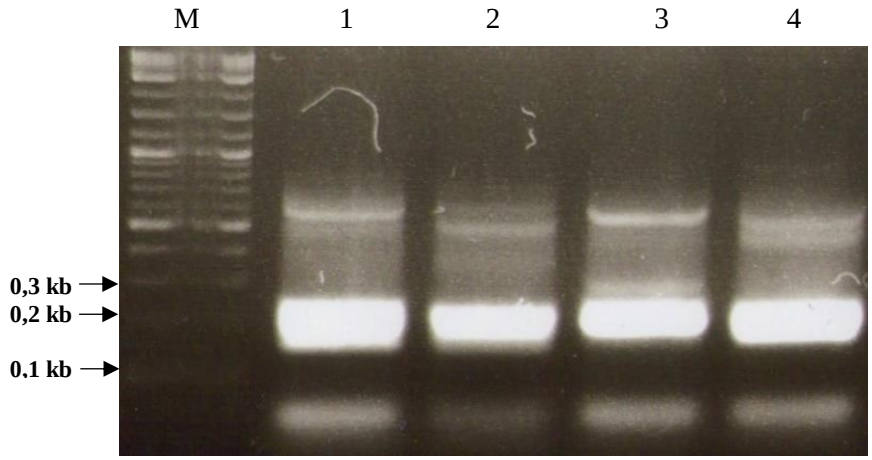
Slika 19. Gel-elektroforeza PCR-produkata ponovljenih reakcija na uzorku genomske DNA iz krvi. Na gel je nanesen marker molekulskih masa (M) i reakcije 1, 60 i 63.

4.3. Reakcije višestrukog PCR-a

Provedeno je više reakcija višestrukog PCR-a na uzorku genomske DNA iz krvi i genomske DNA iz FFPE-a pri različitim uvjetima u svrhu njegove optimizacije (Tablica 2 i Tablica 3). Reakcije su analizirane elektroforezom u 2%-tnom agaroznom gelu uz upotrebu etidij-bromida za vizualizaciju produkata reakcija.

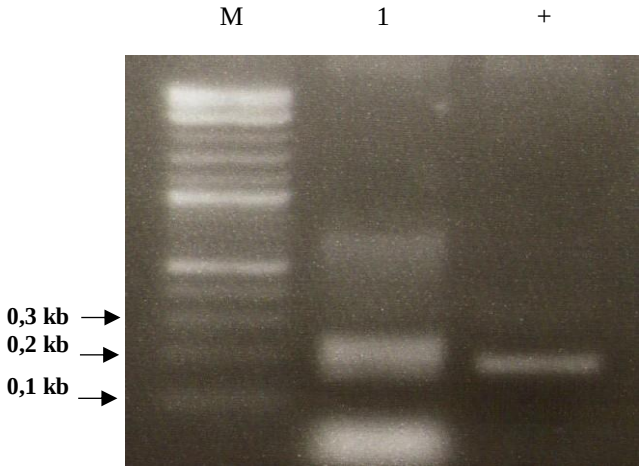
4.3.1. Reakcije višestrukog PCR-a na uzorku genomske DNA iz krvi

Početne reakcije višestrukog PCR-a, pri kojima je korištena ukupna koncentracija početnica od 0,2 μM tj. 0,4 μM (Tablica 4, reakcije 1, 2, 3), nisu bile uspješne, stoga su pripremljene PCR-reakcije koje sadrže 0,15 μM svakog para početnica (Slika 20). Uz željene produkte (vrpce veličine 150-250 pb), u reakciji nastaju i nespecifični produkti (vrpce veće molekulske mase) te je na gelu vidljiva prisutnost dimera početnica (vrpca manja od 100 pb).



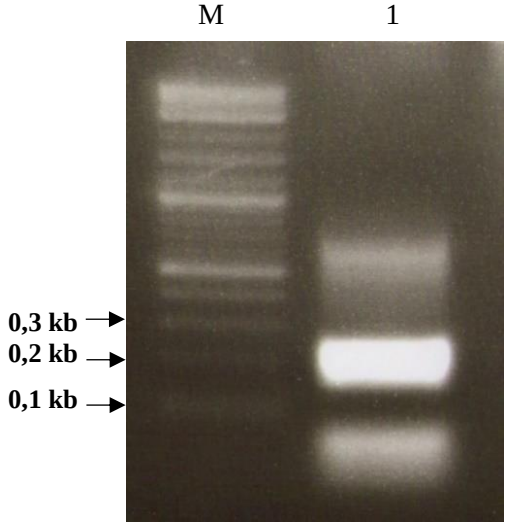
Slika 20. Gel-elektroforeza produkata višestrukog PCR-a uz korištenje 150 ng genomske DNA iz krvi kao kalup, temperature i vremena sparivanja početnica od 57 °C, 15 sec i 35 ciklusa PCR-a. Na gel je nanesen marker molekulskih masa (M) i reakcije višestrukog PCR-a 4 smjese početnica (1, 2, 3, 4).

Kako bi se smanjila količina nespecifičnih produkata, povišena je temperatura sparivanja početnica s 57 °C na 59 °C, a broj ciklusa reakcije PCR-a smanjen na 30 (Slika 21).



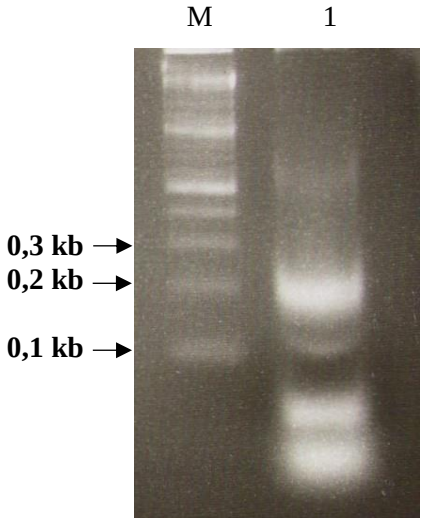
Slika 21. Gel-elektroforeza produkta višestrukog PCR-a uz korištenje 150 ng genomske DNA iz krvi kao kalup, temperature i vremena sparivanja početnica od 59 °C, 15 sec i 30 ciklusa PCR-a. Na gel je nanesen marker molekulskih masa (M), reakcija višestrukog PCR-a 1. smjese početnica (1) i umnoženi dio gena *SAHH* kao pozitivna kontrola (+).

Kako bi se pokušalo potpuno ukloniti nespecifične produkte, pripremljena je reakcija u kojoj je temperatura sparivanja početnica iznosila 60 °C te se produljilo vrijeme sparivanja početnica s 15 sec na 30 sec (Slika 22). Na gelu su još uvijek prisutni nespecifični produkti, no u manjem omjeru naprema željenom produktu.



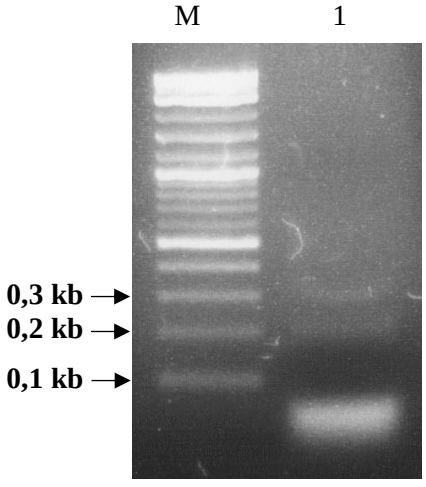
Slika 22. Gel-elektroforeza produkata višestrukog PCR-a uz korištenje 150 ng genomske DNA iz krvi kao kalup, temperature i vremena sparivanja početnica od 60 °C, 30 sec i 30 ciklusa PCR-a. Na gel je nanesen marker molekulskih masa (M) i reakcija višestrukog PCR-a 1. smjese početnica (1).

Daljnja optimizacija provedena je u svrhu smanjenja količine početnog kalupa za reakciju višestrukog PCR-a te je pripremljena reakcija s 50 ng genomske DNA, a broj ciklusa PCR-a je iznosio 30 (Slika 23). Na gelu je vidljiva zadovoljavajuća količina željenog produkta.

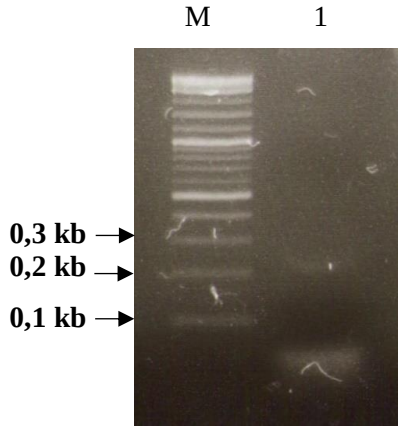


Slika 23. Gel-elektroforeza produkta višestrukog PCR-a uz korištenje 50 ng genomske DNA iz krvi kao kalup, temperature i vremena sparivanja početnica od 57 °C, 15 sec i 30 ciklusa PCR-a. Na gel je nanesen marker molekulskih masa (M) i reakcija višestrukog PCR-a 1. smjese početnica (1).

Nadalje, optimizacija broja ciklusa višestrukog PCR-a s provedena je pripremom dvije reakcije pri identičnim uvjetima kao i prethodne uz smanjenje broja ciklusa PCR-a na 25 (Slika 24) i 20 (Slika 25).



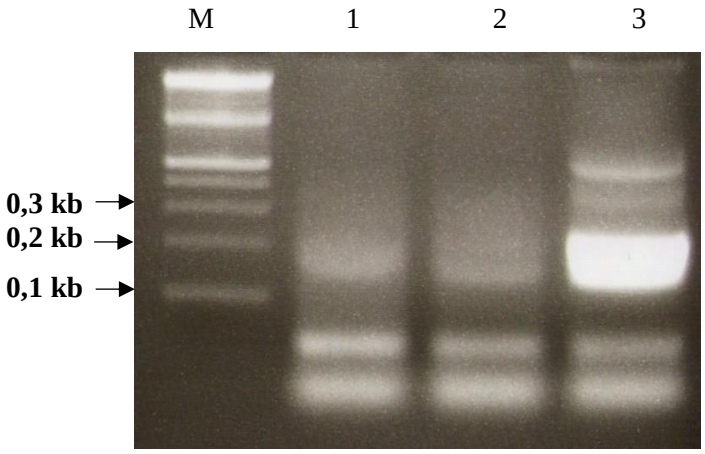
Slika 24. Gel-elektroforeza produkta višestrukog PCR-a uz korištenje 50 ng genomske DNA iz krvi kao kalup, temperature i vremena sparivanja početnica od 57 °C, 15 sec i 25 ciklusa PCR-a. Na gel je nanesen marker molekulskih masa (M) i reakcija višestrukog PCR-a 1. smjese početnica (1).



Slika 25. Gel-elektroforeza produkta višestrukog PCR-a uz korištenje 50 ng genomske DNA iz krvi kao kalup, temperature i vremena sparivanja početnica od 57 °C, 15 sec i 20 ciklusa PCR-a. Na gel je nanesen marker molekulskih masa (M) i reakcija višestrukog PCR-a 1. smjese početnica (1).

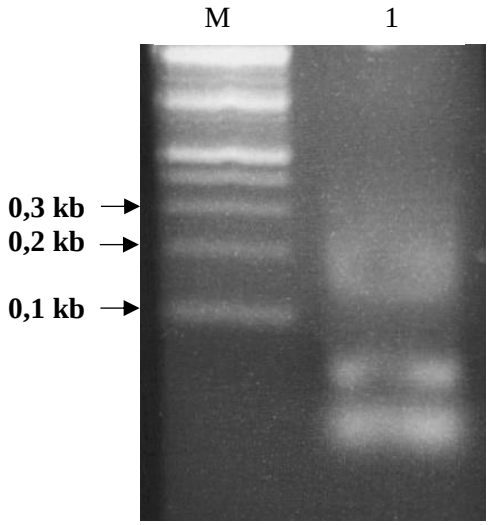
4.3.2. Reakcije višestrukog PCR-a na uzorku genomske DNA iz FFPE-a

Za početnu reakciju višestrukog PCR-a na genomskoj DNA iz FFPE-a korišteno je 200 ng kalupa, 35 ciklusa PCR-a te je kao pozitivna kontrola korištena genomska DNA iz krvi (Slika 26). Vidi se puno veća količina željenog produkta na uzorku genomske DNA iz krvi u odnosu na genomsku DNA iz FFPE-a, ali je kod genomske DNA iz FFPE-a nastala manja količina nespecifičnih produkata.



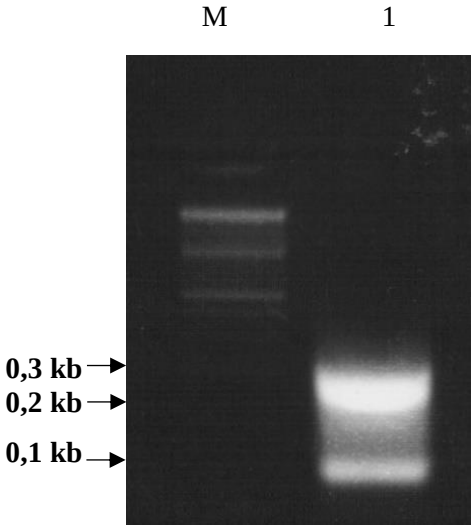
Slika 26. Gel-elektroforeza produkata višestrukog PCR-a genomske DNA iz krvi i DNA iz FFPE-a uz korištenje 200 ng DNA kao kalup, temperature i vremena sparivanja početnica od 57 °C, 15 sec, vrijeme produživanja početnica 10 sec i 35 ciklusa PCR-a Na gel je nanesen marker molekulskih masa (M) i reakcija višestrukog PCR-a 1. i 2. smjese početnica na uzorku DNA iz FFPE-a (1 i 2) i 1. smjese početnica na uzorku DNA iz krvi (3).

U svrhu povećanja prinosa željenog produkta napravljena je reakcija u kojoj je vrijeme produživanja početnica promijenjeno s početnih 10 sec na 1 min (Slika 27).



Slika 27. Gel-elektroforeza produkta reakcije višestrukog PCR-a uz korištenje 200 ng genomske DNA iz FFPE-a kao kalup, temperature i vremena sparivanja početnica od 57 °C, 15 sec, vrijeme produživanja početnica 1 min i 35 ciklusa PCR-a. Na gel je nanesen marker molekulskih masa (M) i reakcija višestrukog PCR-a 2. smjese početnica (1).

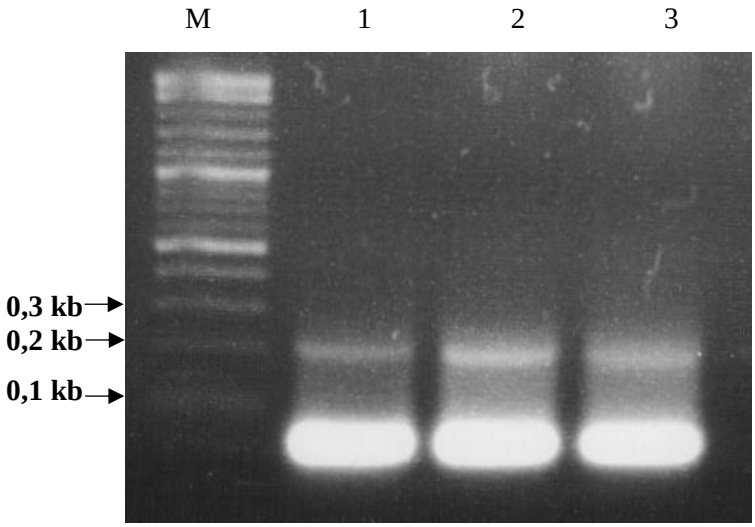
U sljedećoj reakciji (Slika 28), uz dulje vrijeme produživanja početnica, produženo je i vrijeme sparivanja početnica s 15 sec na 30 sec. Na gelu je vidljiva znatno veća količina željenog produkta od prethodnih reakcija uz kraće vrijeme sparivanja početnica.



Slika 28. Gel-elektroforeza produkta reakcije višestrukog PCR-a uz korištenje 200 ng genomske DNA iz FFPE-a kao kalup, temperature i vremena sparivanja početnica od 57 °C, 30 sec, vrijeme produživanja početnica 1 min i 35 ciklusa PCR-a. Na gel je nanesen marker molekulskih masa (M) i reakcija višestrukog PCR-a 1. smjese početnica (1).

Za optimizaciju broja ciklusa PCR-a, napravljene su reakcije s 20 odnosno 25 ciklusa PCR-a, ali na gelu nije bila vidljiva zadovoljavajuća količina željenog produkta.

Optimizacija početne količine kalupa za reakciju višestrukog PCR-a, provedena je variranjem količina dodane genomske DNA iz FFPE-a – 25 ng, 50 ng i 100 ng (Slika 29).

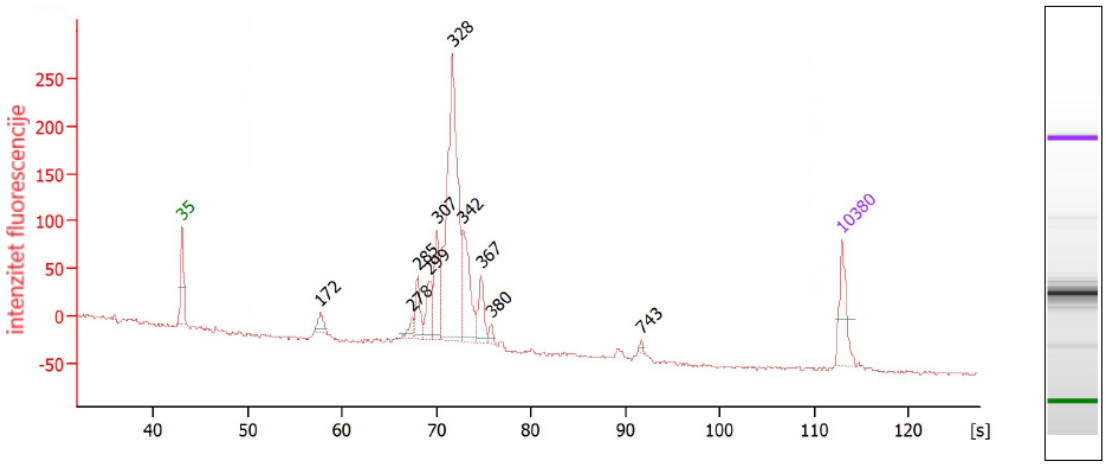


Slika 29. Gel-elektroforeza produkta reakcije višestrukog PCR-a uz dodatak različite količine DNA iz FFPE-a kao kalup, temperature i vremena sparivanja početnica od 57 °C, 30 sec, vrijeme produživanja početnica 1 min i 35 ciklusa PCR-a. Na gel je nanesen marker molekulskih masa (M) i reakcije višestrukog PCR-a smjese 3 početnica s 25 ng (1), 50 ng (2) i 100 ng (3) DNA iz FFPE-a.

4.4. Rezultati DNA-čipa visoke osjetljivosti pripremljenih biblioteka

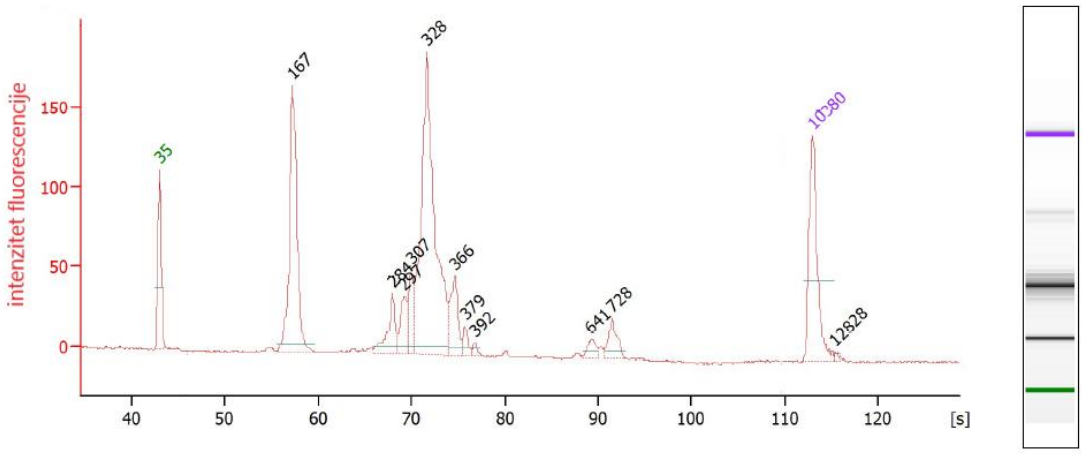
Nakon što je prema opisanom postupku za pripremu biblioteke za sekvenciranje pripremljeno 6 biblioteka, kojima su prethodili različiti uvjeti višestrukog PCR-a (Tablica 6), uzorci su analizirani na DNA-čipu visoke osjetljivosti (Agilent 2100 Bioanalyzer). Sustav očitava dva unutarnja standarda – marker niže veličine od 35 pb i više veličine od 10 380 pb te se na temelju njih i markera molekularnih veličina kvantificira veličina fragmenata u uzorku.

Biblioteka 1 (Slika 30), pokazala je najveću distribuciju pikova veličine od 278 pb do 380 pb. To je u skladu s očekivanim budući da je veličina dodanih adaptera 120 pb, a amplikoni dobiveni višestrukim PCR-om su veličine od 150 pb do 270 pb, odnosno prosječno 210 pb. Pik veličine 172 pb može se pripisati produktu koji je nastao vezanjem adaptera na zaostale dimere početnica iz reakcija višestrukog PCR-a, a pik od 743 pb potječe od nespecifičnih produkata.



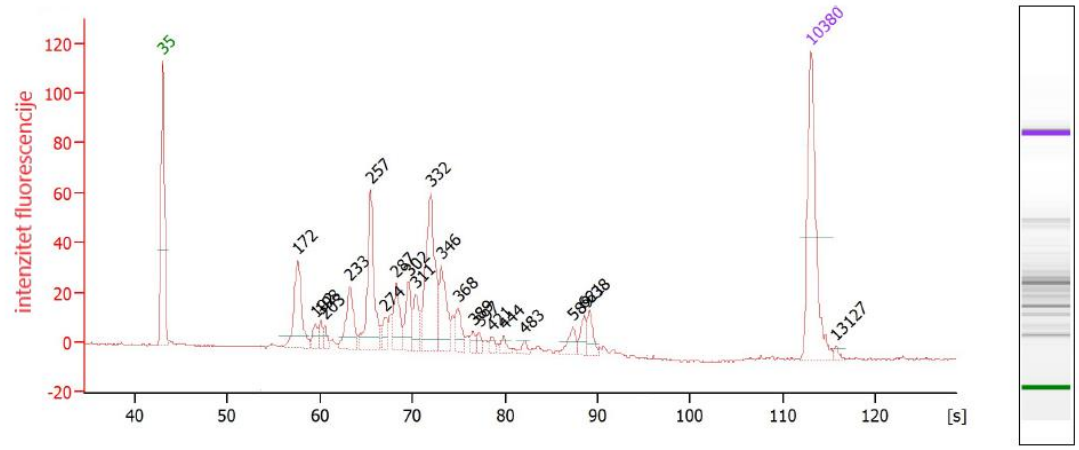
Slika 30. Elektroferogram fragmenata dobivenih analizom biblioteke 1 na DNA-čipu visoke osjetljivosti. Biblioteka je pripremljena iz 4 smjese višestrukih PCR-produkata provedenih na uzorku DNA iz krvi uz 35 ciklusa umnažanja. Niži marker (zeleno), viši marker (ljubičasto).

Biblioteka 2 (Slika 31), pokazuje najveću distribuciju pikova od 284 pb do 392 pb. Uočava se pik od 167 pb koji odgovara vezanim adapterima na zaostale dimere početnica iz reakcija višestrukog PCR-a te pikovi od 641 i 728 pb koji odgovaraju nespecifičnih produktima.



Slika 31. Elektroferogram fragmenata dobivenih analizom biblioteke 2 na DNA-čipu visoke osjetljivosti. Biblioteka je pripremljena iz 4 smjese višestrukih PCR-produkata provedenih na uzorku DNA iz krvi uz 25 ciklusa umnažanja. Niži marker (zeleno), viši marker (ljubičasto).

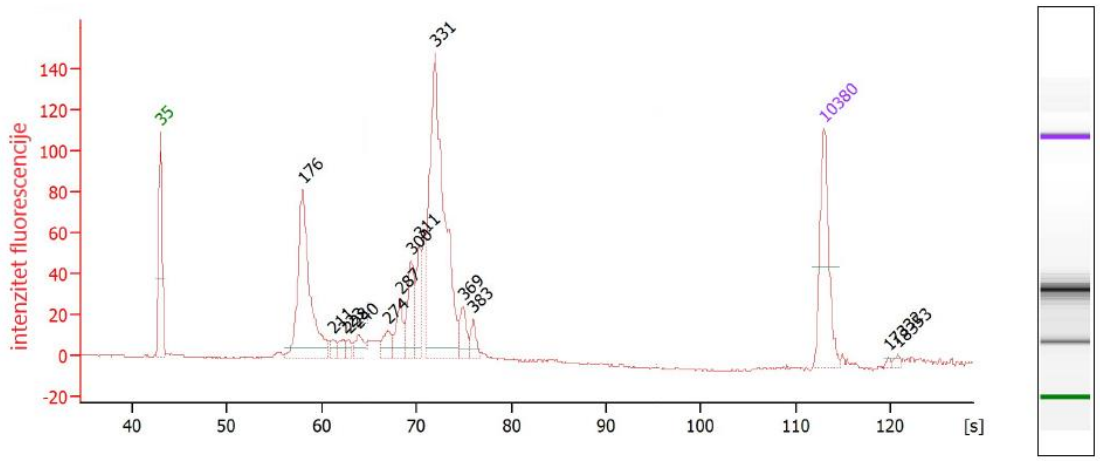
Biblioteka 3 (Slika 32) koja je za razliku od prethodnih pripremljena iz dvije smjese početnica, pokazuje široku distribuciju pikova od 172 pb do 638 pb. Usporedbom s prethodna dva grafa, vidi se da je nastalo znatno manje fragmenata očekivane veličine.



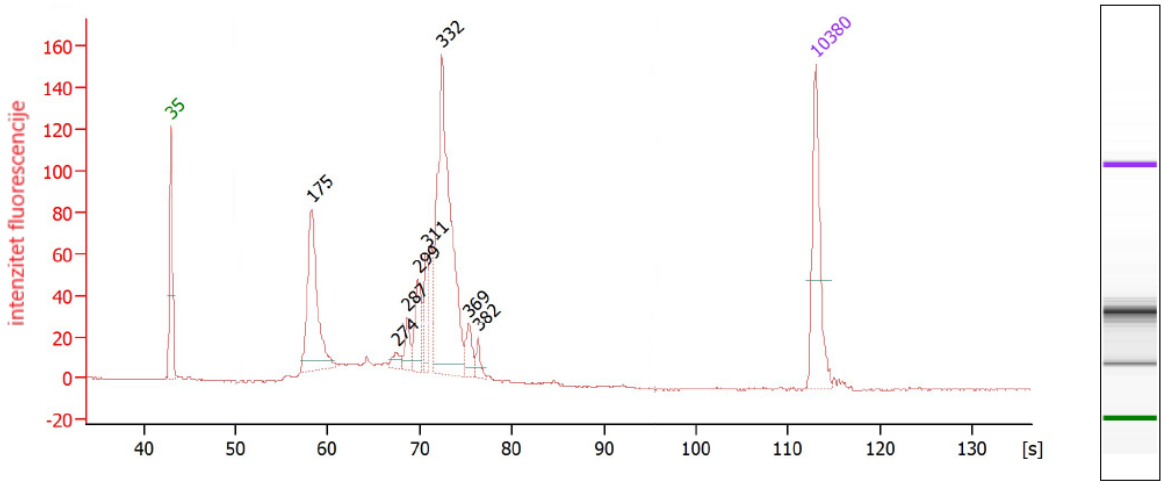
Slika 32. Elektroferogram fragmenata dobivenih analizom biblioteke 3 na DNA-čipu visoke osjetljivosti. Biblioteka je pripremljena iz 2 smjese višestrukih PCR-produkata provedenih na uzorku DNA iz krvi uz 35 ciklusa umnažanja. Niži marker (zeleno), viši marker (ljubičasto).

Biblioteke 4, 5 i 6 pripremljene su korištenjem genomske DNA iz FFPE-a kao kalupa za višestruki PCR. Biblioteke 4 i 5 (Slika 33 i Slika 34) pokazuju najveću distribuciju pikova od 274 pb do 383 pb, odnosno 274 do 382 pb, koji odgovaraju očekivanoj veličini biblioteke. Kao i u prethodnim bibliotekama, uočavaju se i pikovi od 176 pb tj. 175 pb koji su rezultat vezanja adaptera na dimere početnica.

Nadalje, u bibliotekama koje su pripremljene korištenjem genomske DNA iz FFPE-a ne nastaju nespecifični produkti kao što je to u slučaju biblioteka pripremljenih na uzorku genomske DNA iz krvi.

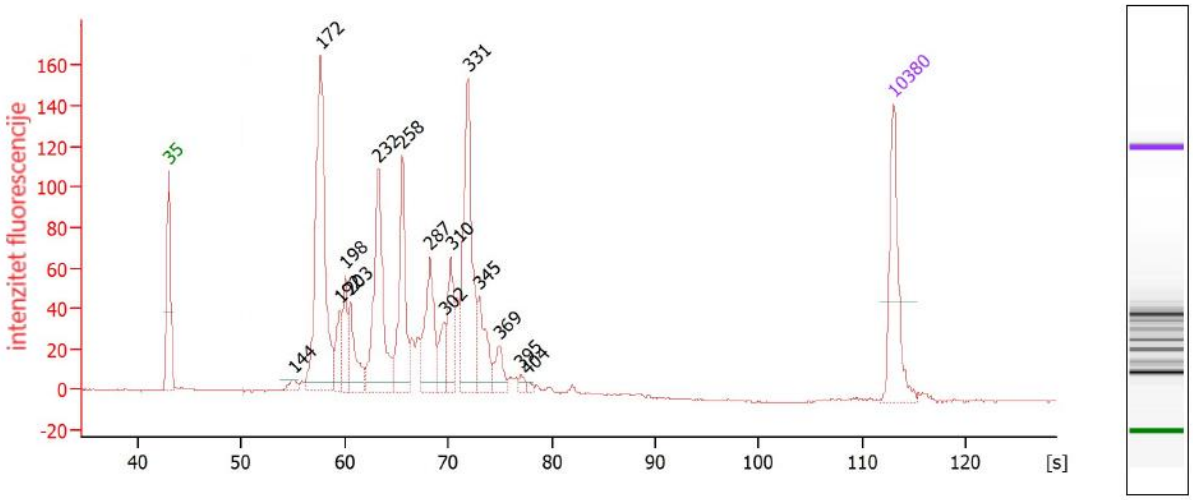


Slika 33. Elektroferogram fragmenata dobivenih analizom biblioteke 4 na DNA-čipu visoke osjetljivosti. Biblioteka je pripremljena iz 4 smjese višestrukih PCR-produkata provedenih na uzorku DNA iz FFPE-a uz 35 ciklusa umnažanja. Niži marker (zeleno), viši marker (ljubičasto).



Slika 34. Elektroferogram fragmenata dobivenih analizom biblioteke 5 na DNA-čipu visoke osjetljivosti. Biblioteka je pripremljena iz 4 smjese višestrukih PCR-produkata provedenih na uzorku DNA iz FFPE-a uz 45 ciklusa umnažanja. Niži marker (zeleno), viši marker (ljubičasto).

Biblioteka 6 (Slika 35), kojoj su prethodile reakcije višestrukog PCR-a s dvije smjese početnica na genomskoj DNA iz FFPE-a kao kalupu, pokazuje široku distribuciju pikova od 144 pb do 404 pb.



Slika 35. Elektroferogram fragmenata dobivenih analizom biblioteke 6 na DNA-čipu visoke osjetljivosti. Biblioteka je pripremljena iz 2 smjese višestrukih PCR-produkata provedenih na uzorku DNA iz FFPE-a uz 35 ciklusa umnažanja. Niži marker (zeleno), viši marker (ljubičasto).

4.5. Rezultati sekvenciranja

Prekonoćnim radom sekvencera, provedeno je sekvenciranje s oba kraja fragmenata DNA što je rezultiralo sa sveukupno 314 ciklusa očitavanja baza. Ukupno je očitano 16 290 718 sekvenci. Svaka pripremljena biblioteka je imala svoj specifični barkod na temelju kojih ih je uređaj razvrstao za daljnju analizu. Razvrstano je ukupno 10 439 957 sekvenci koje su veće od 100 pb.

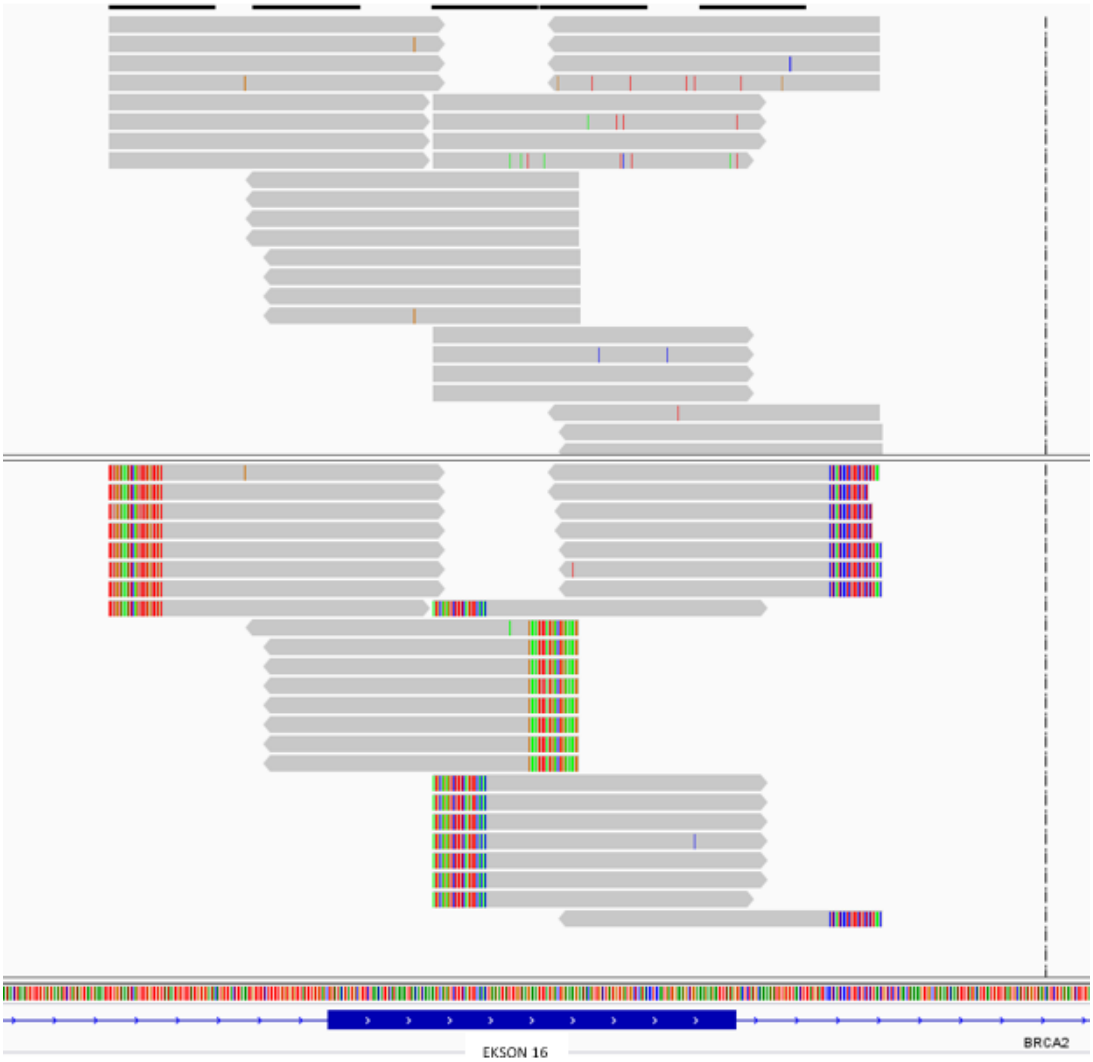
Ranije opisanim programskim paketom „BWA Alligner“, sekvence su posložene prema ljudskom referentnom genomu. Pomoću programa „Qualimap“ iščitane su najbitnije statističke informacije poravnanja sekvenci (Tablica 7).

Tablica 7. Statističke informacije poravnanja sekvenci dobivene programom „Qualimap“.

		DNA iz krvi			DNA iz FFPE-a		
		35 ciklusa	25 ciklusa	35 ciklusa/ 2 smjese	35 ciklusa	45 ciklusa	35 ciklusa/ 2 smjese
broj očitanih sekvenci		2 518 016	1 311 104	2 019 223	1 560 940	1 756 747	1 273 927
postotak mapiranih sekvenci		96,24%	93,06%	91,15%	94,81%	90%	89,61%
prosječna duljina očitane sekvence		137 pb	144 pb	143 pb	145 pb	143 pb	143 pb
postotak mapiranih sekvenci unutar regije od interesa		84,03%	81,70%	85,38%	80,47%	62,30%	75,78%
prosječna pokrivenost unutar regije od interesa		9114	4624	6504	5380	4528	3491
prosječna pokrivenost po kromosomu	**kromosom 13**	8398	4779	6411	5834	4117	3561
	kromosom 17	20989	9656	16075	10958	10997	8878

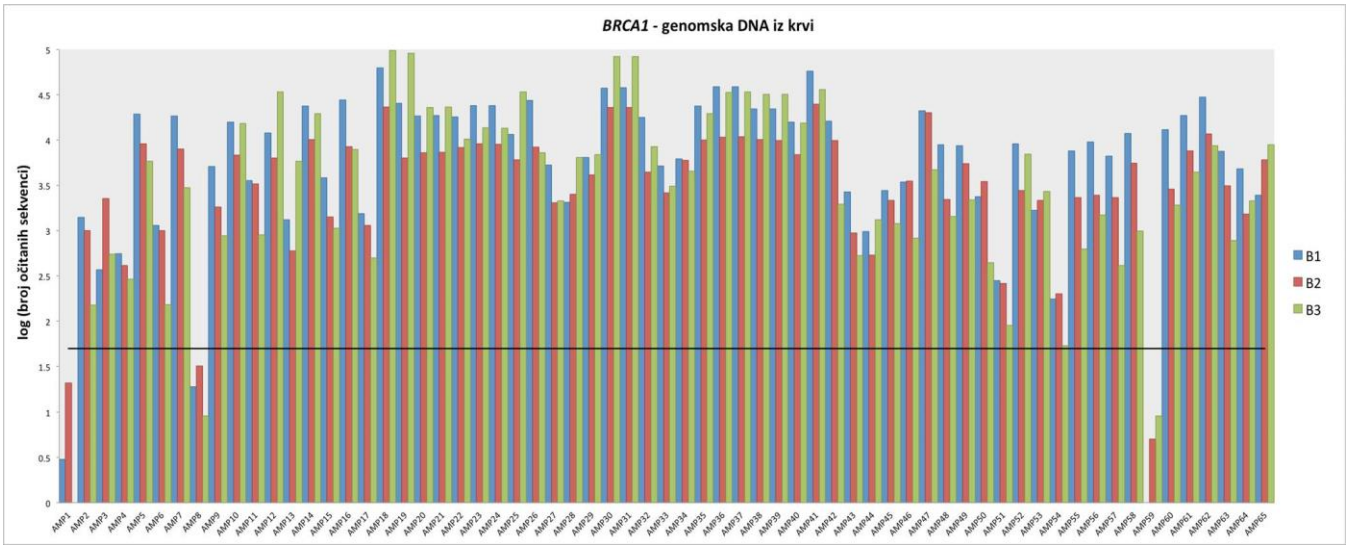
Iz tablice se vidi visoki postotak mapiranih sekvenci na referentni ljudski genom (90% i više). Također, postotak mapiranih sekvenci unutar regija od interesa je visok kod uzoraka na genomskoj DNA iz krvi (>80%), dok je nešto niži kod uzoraka na DNA iz FFPE-a (>60%). Primjećuje se da je prosječna pokrivenost unutar regije od interesa u korelaciji s brojem očitanih sekvenci. Pokrivenost na razini genoma odgovara lokacijama gena *BRCA1* i *BRCA2* na kromosomu 17, odnosno 13.

Programom „BAMClipper“ izuzet je dio svake očitane sekvence koji pripada početnicama korištenim u višestrukom PCR-u. Na primjeru eksona 16 gena *BRCA2* (Slika 36), vide se dijelovi sekvenca koji odgovaraju neizuzetim, odnosno izuzetim početnicama.

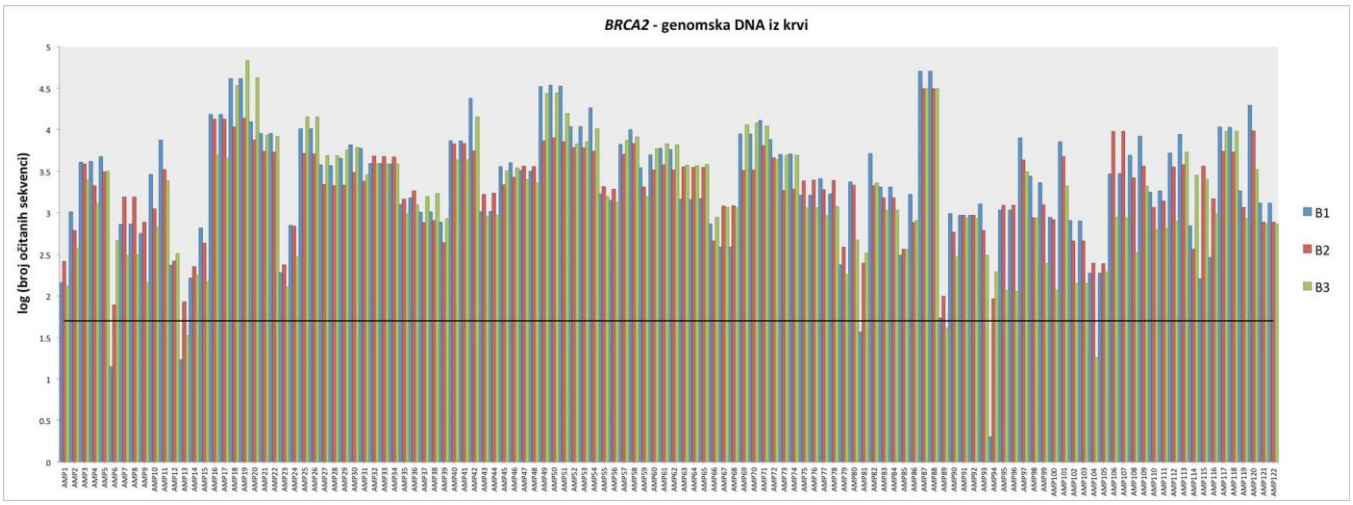


Slika 36. Prikaz poredanih sekvenci na ekson 16 gena *BRCA2* s neizuzetim početnicama (gore) i izuzetim početnicama (dolje). Za vizualizaciju korišten je program „IGV“. Šarene prugice predstavljaju izuzete dijelove.

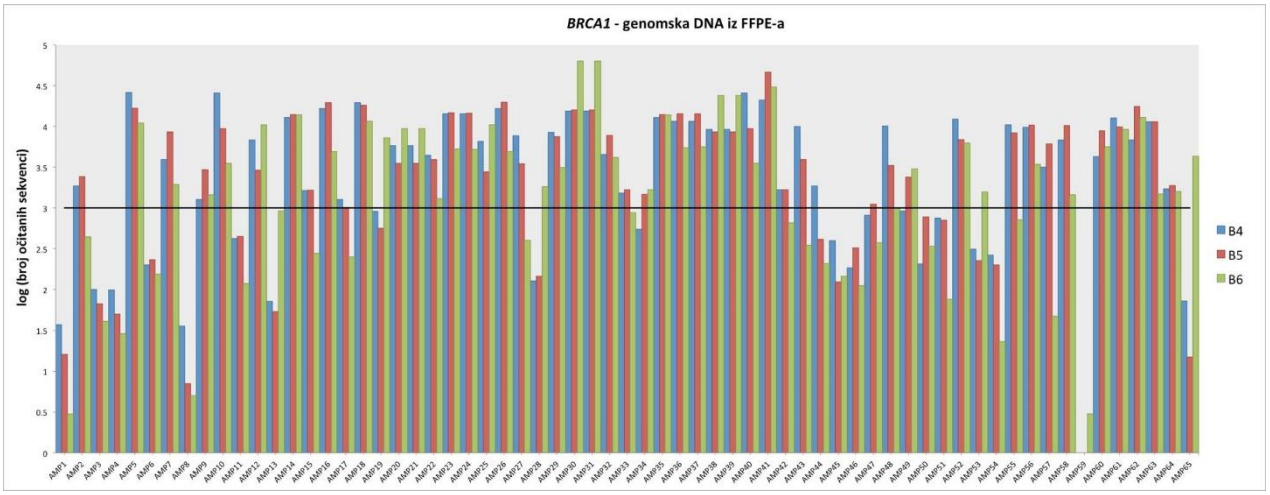
U svrhu analize umnažanja i očitanja svakog pojedinog amplikona, tj. fragmenta DNA koji pripada regiji od interesa, konstruiran je grafički prikaz broja očitanja po svakom pojedinom amplikonu. Za uzorak genomske DNA iz krvi zadan je prag od minimalno 50 očitanja, a za uzorak genomske DNA iz FFPE-a minimalno 1000 očitanja.



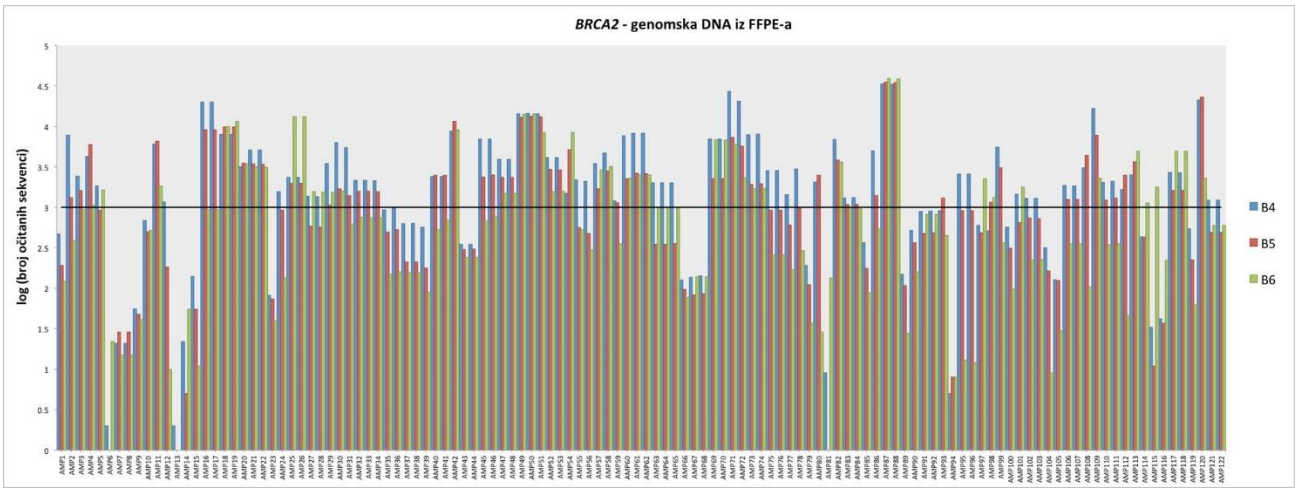
Slika 37. Grafički prikaz broja očitanih sekvenci po pojedinom amplikonu gena *BRCA1* na uzorku genomske DNA iz krvi. B1 – DNA iz krvi, 35 ciklusa, 4 smjese početnica; B2 – DNA iz krvi, 25 ciklusa, 4 smjese početnica, B3 – DNA iz krvi, 35 ciklusa, 2 smjese početnica; crna linija – prag od 50 očitanja.



Slika 38. Grafički prikaz broja očitanih sekvenci po amplikonu gena *BRCA2* na uzorku genomske DNA iz krvi. B1 – DNA iz krvi, 35 ciklusa, 4 smjese početnica; B2 – DNA iz krvi, 25 ciklusa, 4 smjese početnica, B3 – DNA iz krvi, 35 ciklusa, 2 smjese početnica; crna linija – prag od 50 očitanja.



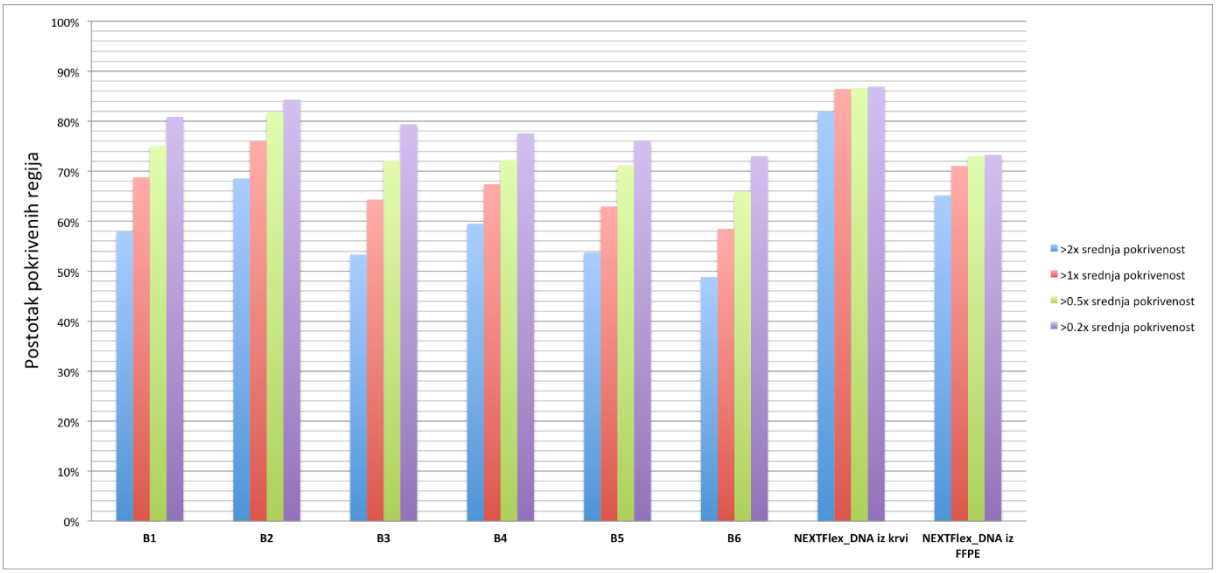
Slika 39. Grafički prikaz broja očitanih sekvenci po amplikonu gena *BRCA1* na uzorku genomske DNA iz FFPE-a. B4 – DNA iz FFPE-a, 35 ciklusa, 4 smjese početnica, B5 - DNA iz FFPE-a, 45 ciklusa, 4 smjese početnica, B6 – DNA iz FFPE-a, 35 ciklusa, 2 smjese početnica; crna linija – prag od 1000 očitanja.



Slika 40. Grafički prikaz broja očitanih sekvenci po amplikonu gena *BRCA2* na uzorku genomske DNA iz FFPE-a. B4 – DNA iz FFPE-a, 35 ciklusa, 4 smjese početnica, B5- DNA iz FFPE-a, 45 ciklusa, 4 smjese početnica, B6 – DNA iz FFPE-a, 35 ciklusa, 2 smjese početnica; crna linija – prag od 1000 očitanja.

Kod uzoraka na genomskoj DNA iz krvi (Slika 37 i 38), vidi se da su pojedini amplikoni gena *BRCA1* i *BRCA2* ispod granične vrijednosti dok je kod uzorka na genomskoj DNA iz FFPE-a (Slika 39 i 40) znatno veći broj amplikona ispod granične vrijednosti.

Na temelju pokrivenosti na razini svake pojedine baze, konstruiran je grafički prikaz uniformnosti dizajniranog panela i komercijalnog testa (NEXTflex) (Slika 41).



Slika 41. Grafički prikaz uniformnosti dizajniranog panela i komercijalnog testa (NEXTflex). B1 – DNA iz krvi, 35 ciklusa, 4 smjese početnica; B2 – DNA iz krvi, 25 ciklusa, 4 smjese početnica, B3 – DNA iz krvi, 35 ciklusa, 2 smjese početnica; B4 – DNA iz FFPE-a, 35 ciklusa, 4 smjese početnica, B5- DNA iz FFPE-a, 45 ciklusa, 4 smjese početnica, B6 – DNA iz FFPE-a, 35 ciklusa, 2 smjese početnica.

Kod komercijalnog testa na uzorku genomske DNA iz krvi oko 80% regije od interesa pokriveno je više od srednje vrijednosti pokrivenosti, dok je kod dizajniranog panela ta vrijednost manja, oko 65-75%, ovisno o biblioteci. Za biblioteke pripremljene genomskom DNA iz FFPE-a, ta vrijednost je nešto manja te iznosi 58-67%, tj. 70% za komercijalni test. Primjećuje se manja uniformnost komercijalnog testa na uzorku genomske DNA iz FFPE-a u odnosu na genomsku DNA iz krvi, kao i generalna manja uniformnost dizajniranog panela u odnosu na komercijalni test.

Pomoću programa „VarScan“ definirane su statistički značajne varijante uz korištenje zadanih postavki koje uključuju graničnu vrijednost od 20% prisutnosti određene varijante u uzorku. Uspoređeni su rezultati analize varijanti biblioteka pripremljenih komercijalnim kemikalijama (NEXTflex) i dizajniranim panelom (Tablica 8 i 9).

Tablica 8. Analiza varijanti dobivenih komercijalnim kemikalijama (NEXTflex) i dizajniranim panelom na uzorku genomske DNA izolirane iz krvi. U tablici su navedeni identifikacijski brojevi varijanti iz baze podataka dbSNP.

DNA iz krvi				
	Komercijalni test	**B1**	**B2**	**B3**
Broj pronađenih varijanti	8	9	8	9
rs1799943	+	+	+	+
rs1801406	+	+	+	+
rs206075	+	+	+	+
rs206076	+	+	+	+
rs1799955	+	+	+	+
rs169547	+	+	+	+
rs95354262	+	+	+	+
rs799923	+	+	+	+

Iz tablice se vidi da je komercijalnim testom identificirano sveukupno 8 varijanti te da su sve varijante pronađene i u pripremljenim uzorcima. U uzorku B1 i B3 pronađena je još jedna dodatna varijanta.

Tablica 9. Analiza varijanti dobivenih komercijalnim kemikalijama (NEXTflex) i dizajniranim panelom na uzorku genomske DNA izolirane iz tumorskog tkiva uklopljenog u parafin. U tablici su navedeni identifikacijski brojevi varijanti iz baze podataka dbSNP.

DNA iz FFPE-a				
	Komercijalni test	**B4**	**B5**	**B6**
Broj pronađenih varijanti	14	15	13	12
rs9534262	+	-	-	-
rs799917	+	+	+	+
rs273900734	+	-	-	-
rs206076	+	+	+	+
rs206075	+	+	+	+
rs1799966	+	+	+	+
rs1799949	+	+	+	+
rs169547	+	+	+	+
rs16942	+	+	+	+
rs16941	+	+	+	+
rs16940	+	+	+	+
rs144848	+	+	+	+
rs1060915	+	+	+	+
neobilježena varijanta	+	-	-	-

Komercijalnim testom pronađeno je sveukupno 14 varijanti. U uzorcima je identificirano 11 varijanti detektiranih komercijalnim testom uz dodatak 4, 2 ili 1 nove varijante, ovisno o uzorku.

5. Rasprava

Cilj opisanih eksperimenata bio je dizajn genskog panela za pripremu biblioteka i sekvenciranje gena *BRCA1* i *BRCA2*. Ovi geni odabrani su jer na tržištu postoji potreba za pristupačnijim sekvenciranjem gena *BRCA1/2,* a s obzirom na dostupnost mnogobrojnih komercijalnih testova na tržištu, olakšana je usporedba i procjena uspješnosti konstrukcije samostalnog panela. Glavni pristup pri dizajnu panela uključivao je umnažanje regija od interesa višestrukim PCR-om jer, u odnosu na ostale pristupe, omogućuje bržu i jeftiniju pripremu biblioteka uz veću pokrivenost regija od interesa, što je od izuzetne važnosti pri detekciji SNV-a, posebice kod uzoraka tumorskog tkiva uklopljenog u parafin (Enyedi i sur. 2016). U skladu s time, korištena su dva različita tipa početnog uzorka - DNA iz krvi te DNA izolirana iz tkiva uklopljenog u parafin, kako bi se ispitale glavne razlike u pripremi biblioteka te analizi podataka ovisno o tipu uzorka.

Za početak, pojedinačnom provjerom konstruiranih parova početnica dobiveni su produkti u svim reakcijama te su njihove veličine uspoređene s očekivanom veličinom produkata. U pojedinim reakcijama nastali nespecifični produkti uspješno su uklonjeni podizanjem temperature sparivanja čime se povećala specifičnost vezanja početnica na odgovarajuća vezna mjesta u molekuli DNA.

Optimizacija višestrukog PCR-a započeta je variranjem ukupne koncentracije početnica te je ustanovljeno da koncentracija od 0,15 μM svakog para početnica rezultira uspješnim nastankom željenih produkata reakcije, što je u skladu s prethodno opisanim rezultatima (Markoulatos i sur. 2002). Naime, umnažanje cjelokupnih kodirajućih regija gena za pojedini uzorak provedeno je u četiri zasebne reakcije s odgovarajućom smjesom početnica kako bi se smanjio udio nespecifičnih produkata zbog prevelikog broja početnica u jednoj reakciji. Za ispitivanje efikasnosti ovakvog pristupa umnažanja ciljne regije, provedene su i reakcije s dvije smjese početnica.

Reakcijama na uzorku DNA iz krvi nastaju i nespecifični produkti kao rezultat velikog broja početnica u jednoj reakciji te njihovog nespecifičnog vezanja duž cijele molekule genomske DNA. Dodatno, u provedenim reakcijama zaostaju i dimeri početnica koji zahtijevaju dodatne korake uklanjanja s obzirom da mogu predstavljati probleme pri daljnjoj konstrukciji panela te analizi podataka nakon sekvenciranja. Jedno od mogućih rješenja je, između ostalog, korištenje PCR-a s vrućim startom (eng. *hot start PCR*) koji smanjuje prinos nespecifičnih produkata i nastanak dimera početnica (Markoulatos i sur. 2002). U reakcijama višestrukog PCR-a na uzorku DNA iz FFPE-a, pri istim uvjetima kao i na uzorku DNA iz krvi,

nastaje znatno manja količina produkta (Slika 26) zbog degradacije molekule DNA i nastanka križnih veza između DNA i proteina koje ometaju pristup DNA polimerazi. Ova svojstva rezultat su dulje pohrane uzorka i niske pH vrijednosti formalina koji se koristi za fiksaciju tkiva (Dietrich i sur. 2013). Zbog razlika u prirodi između dva korištena tipa uzorka, za umnažanje regije od interesa koristeći DNA iz FFPE-a, produženo je vrijeme sparivanja i produživanja početnica. Naime, kratki fragmenti molekule DNA mogu se vezati na DNA polimerazu i na taj način ju inhibirati pa povećanje vremena produživanja početnica pogoduje njenoj aktivnosti i omogućuje kompletnu polimerizaciju fragmenata DNA (Dietrich i sur. 2013). Iako zahtijeva zaseban pristup prilikom umnažanja, DNA iz FFPE-a zbog svoje degradacije ne pogoduje nastanku nespecifičnih produkata (Slika 28).

Daljnja optimizacija višestrukog PCR-a, u svrhu smanjenja količine početnog DNA materijala, provedena je zbog često ograničene početne količine materijala, osobito kod uzoraka tumorskih tkiva. Konačna optimizacija minimalnog broja ciklusa višestrukog PCR-a provedena je zbog sklonosti DNA polimeraze uvođenju pogrešaka čime se povećava ukupni broj grešaka što je više ciklusa PCR-reakcija. Najveći problem predstavljaju GC i AT bogata područja u molekuli DNA (Pan i sur. 2014). Sukladno tome, daljnja priprema biblioteka za sekvenciranje provedena je koristeći reakcije s različitim brojem ciklusa umnažanja kako bi se ispitao njihov utjecaj na rezultate sekvenciranja.

Na temelju provedene optimizacije, odabrana su po tri različita protokola umnažanje uzorka genomske DNA iz krvi i umnažanje DNA iz FFPE-a te je pripremljeno sveukupno 6 različitih biblioteka. Kvaliteta biblioteka je procijenjena pomoću DNA-čipa visoke osjetljivosti. Očekivano, biblioteke pripremljene iz četiri reakcije višestrukog PCR-a na uzorku DNA iz krvi uz fragmente željene veličine sadrže i nespecifične produkte, dok to nije slučaj kod biblioteka pripremljenih na uzorku DNA iz FFPE-a (Slika 30, 31, 33, 34). Na temelju njihovih veličina, nespecifični produkti od 167 pb do 176 pb se mogu pripisati vezanim adapterima na dimere početnica zaostale iz reakcija višestrukog PCR-a, dok su produkti od 641 pb do 743 pb rezultat nespecifičnog vezanja početnica u reakciji višestrukog PCR-a. U slučaju biblioteka koje su pripremljene iz dvije, a ne četiri reakcije višestrukog PCR-a (Slika 32 i 35), produkti pokazuju široku distribuciju veličina zbog većeg nastanka nespecifičnih produkata jer je dodano dvostruko više početnica po reakciji višestrukog PCR-a.

Nakon sekvenciranja pripremljenih biblioteka, dobivena očitanja razvrstana su prema jedinstvenom barkodu koji je karakterističan za pojedinu biblioteku. Uklonjena su očitanja manja od 100 pb s obzirom da su željeni amplikoni zajedno s početnicama veličine od 151 pb do 270 pb. Ukupni broj očitanih sekvenci i razvrstanih sekvenci razlikuje se za 5 850 761

očitanja. Ta razlika upućuje na vezanje i sekvenciranje niskih nespecifičnih produkata biblioteka koji se pripisuju vezanim adapterima na dimere početnica.

Poravnavanjem sekvenci na ljudski referentni genom, na uzorku DNA iz krvi broj očitanih sekvenci značajno se mijenja ovisno o broju ciklusa višestrukog PCR-a (Tablica 7). Razlog tome najvjerojatnije se nalazi u razlici između prinosa niskih nespecifičnih produkata koja je vidljiva između dvije biblioteke (Slika 30 i 31). Preostale razlike između očitanja najvjerojatnije su rezultat nedovoljno preciznog miješanja biblioteka u jednakom omjeru.

Općenito, postotak mapiranih sekvenci na ljudski referentni genom je visok za sve uzorke i iznosi >90%. Postotak mapiranih sekvenci unutar regije od interesa također je zadovoljavajući (62-85%) te je očekivano viši kod biblioteka DNA iz krvi u odnosu na biblioteke DNA iz FFPE-a. Najniži postotak mapiranih sekvenci unutar regije od interesa (62,30%) vidljiv je kod biblioteke pripremljene pri najvećem broju ciklusa, tj. 45 ciklusa (Tablica 7). Usporedbom s ostalim bibliotekama, moguća je pretpostavka da se povećanjem broja ciklusa povećava i vjerojatnost uvođenja krivih baza te se time smanjuje mogućnost mapiranja određenog fragmenta na referentni genom što u konačnici rezultira nižim postotkom mapiranih sekvenci na regiju od interesa.

Nadalje, iz poravnatih sekvenci su izuzeti dijelovi sekvenci koji pripadaju početnicama korištenima u reakcijama višestrukog PCR-a iz razloga što bi sekvence početnica mogle pokriti potencijalne SNV-ove u genomu koji se nalaze na mjestu vezanja početnica i dovesti do lažno-negativnih mutacija (Shin i sur. 2016). Iz tog razloga, prilikom konstrukcije početnica za umnažanje kodirajućih dijelova gena *BRCA1/2*, vodilo se računa o tome da se amplikoni međusobno preklapaju kako bi čitava kodirajuća regija oba gena bila pokrivena amplikonima kada se iz analize izuzmu sekvence početnica.

Kako bi se procijenila uspješnost umnažanja i očitanja svakog pojedinog amplikona, za biblioteke DNA iz krvi zadan je prag od minimalno 50 očitanja po amplikonu, dok je za uzorke DNA iz FFPE-a zadan prag od minimalno 1000 očitanja (Enyedi i sur. 2016). Prag je niži u slučaju DNA iz krvi jer je cilj detekcija nasljednih varijanti koje su prisutne u svim stanicama organizma, dok je kod DNA iz FFPE-a znatno viši jer je uzorak tumora u pravilu heterogen pa osim tumorskih stanica sadrži i zdrave stanice, a sama mutacija je prisutna na subklonalnoj razini što rezultira njenim nižim udjelom pojavljivanja u cjelokupnom uzorku te time zahtjeva veću pokrivenost za njezinu detekciju. Na temelju broja očitanja po svakom pojedinom amplikonu (Slika 37-40), u slučaju DNA iz krvi, ukupno 9 amplikona se nalazi ispod zadanog praga dok u slučaju DNA iz FFPE-a, čak 89 amplikona se nalazi ispod zadanog praga. Iz tih podataka može se zaključiti da je dizajnirani panel primjenjiv za uzorke DNA iz krvi, dok je za

analizu DNA iz FFPE-a potrebna dodatna optimizacija. Treba napomenuti da su zadani pragovi očitanja vrlo strogi te se u praksi koristite i niži pragovi čime bi se daleko manji broj amplikona nalazio ispod zadane razine očitanja (Dong i sur. 2015., Kang i sur. 2016). Isto tako, izborom specijalnog čipa za sekvenciranje većeg kapaciteta ili smanjenjem ukupnog broja uzorka po reakciji sekvenciranja, postigla bi se veća prosječna pokrivenost.

Usporedbom uniformnosti dizajniranog panela s komercijalnim testom (Slika 41), može se zaključiti kako je postignuta slabija uniformnost, odnosno veća je razlika pokrivenosti među pojedinim bazama dizajniranog panela. Razlog tomu je vidljivo veća brojnost određenih amplikona naspram ostalih, dok pojedini amplikoni čak u potpunosti nedostaju. Općenito, kao i u dizajniranom panelu, i kod komercijalnog testa se uočava manja uniformnost na uzorku genomske DNA iz FFPE-a u odnosu na genomsku DNA iz krvi, što odražava razliku između prethodno spomenutih svojstava različitih uzoraka.

Usporedbom statistički značajnih varijanti identificiranih komercijalnim testom i dizajniranim panelom kod uzoraka krvi, najbolje rezultate pokazuje biblioteka pripremljena korištenjem najmanjeg broja ciklusa gdje su detektirane iste varijante kao i kod komercijalnog testa. Preostale biblioteke pripremljene korištenjem DNA iz krvi, pokazuju dodatne varijante (Tablica 8). S obzirom da je riječ o bibliotekama pri kojima se koristio veći broju ciklusa PCR-a, moguće je da su dodatne varijante posljedica ugradnje pogrešne baze DNA polimerazom prilikom reakcija višestrukog PCR-a. Potvrda ove teze mogla bi se provesti klasičnim Sangerovim sekvenciranjem, no ta metoda provjere nije dio ovog istraživanja.

Kod biblioteka pripremljenih korištenjem DNA iz FFPE-a, niti u jednom uzorku nisu pronađene sve iste varijante kao kod komercijalnog testa. U sve tri biblioteke identificirano je 11 varijanti kao i u komercijalnom testu, uz dodatak različitih novih varijanti. Izostanak triju potvrđenih varijanti može se pripisati premalom broju amplikona koji prekrivaju regije eksona 17 gena *BRCA2* i eksona 11 gena *BRCA1* gdje se nalaze detektirane varijante. Uz to, treća izostala varijanta nalazi se u intronskoj regiji koja pri dizajnu panela nije obuhvaćena. Uzrok dodatnim varijantama i premalom broju pojedinih amplikona mogao bi biti i nastanak abazičnih mjesta u molekuli DNA zbog hidrolize šećernih veza pri niskom pH zbog čega DNA polimeraza sprječava daljnju amplifikaciju molekule DNA ili ugrađuje pogrešnu bazu. Također, u uzorku DNA iz FFPE-a često dolazi i do deaminacije citozinskih baza čime nastaju uracilske lezije koje se ne popravljaju u potpunosti prilikom izolacije DNA pa i one mogu uzrokovati nastanak varijanti prilikom PCR-reakcija. Korištenje polimeraze visoke vjernosti (eng. *high fidelity polymerase*) mogla bi se smanjiti količina artefakta koji nastaju zbog lezija u molekuli DNA (Do i Dobrovic, 2015).

Dakle, dizajnirani panel se u odnosu na komercijalni test pokazao vrlo uspješnim u detekciji varijanti kod uzoraka DNA iz krvi dok je kod uzoraka DNA iz FFPE-a uspješnost nešto manja stoga je potrebna dodatna optimizacija. Provjerom i ponovnim dizajnom početnica koje nisu dale zadovoljavajuću brojnost amplikona, mogla bi se poboljšati detekcija varijanti te postići jednaka točnost kao i komercijalnim testom. Također, istraživanjima je pokazano da je uspjeh PCR-reakcija kod DNA iz FFPE-a povezan s veličinom amplikona te da bi za povećanje količine kalupa i uspješnije umnažanje amplikoni trebali biti <120 pb (Do i Dobrovic, 2015). Slično se može primijetiti i kod komercijalnog testa gdje je dužina amplikona s početnicama od 122 pb do 196 pb pa bi sljedeći korak optimizacije dizajniranog panela za ovakav tip uzorka svakako trebao ići u smjeru smanjenja veličine amplikona.

Znanje stečeno ovim istraživanjem koristit će se u razvoju proizvoljnih genskih panela za analizu bilo kojih gena od interesa. Takvi genski paneli u budućnosti bi mogli naći široku primjenu u istraživanjima i medicinskoj dijagnostici zbog veće isplativosti i točnosti u odnosu na tradicionalne metode koje se danas rutinski koriste.

6. Zaključci

- Uspješno su dizajnirane početnice za umnažanje kodirajućih sljedova gena *BRCA1* i *BRCA2*.
- Na uzorku genomske DNA iz krvi i tumorskog tkiva uklopljenog u parafin, optimizirani su uvjeti umnažanja fragmenata gena *BRCA1* i *BRCA2* višestrukim PCR-om obzirom na količinu genomske DNA, konačnu koncentraciju smjese početnica, broj ciklusa višestrukih PCR-reakcija, temperaturu i vrijeme sparivanja početnica te vrijeme produživanja početnica.
- Na temelju odabranih optimalnih uvjeta višestrukog PCR-a, uspješno su pripremljene po 3 biblioteke na uzorku genomske DNA iz krvi i tkiva uklopljenog u parafin.
- Usporedbom sekvenciranih uzoraka s komercijalno dostupnim testom za analizu mutacija gena *BRCA1* i *BRCA2*, zaključuje se da je genski panel uspješno dizajniran, ali je potrebna dodatna optimizacija, posebice kod uzoraka na genomskoj DNA iz tkiva uklopljenog u parafin.

7. Literatura

Agilent Technologies (2013): Agilent High Sensitivity DNA Kit Guide.

Ahmadloo S., Nakaoka H., Hayano T., Hosomichi K., You H., Utsuno E., Sangai T. i sur. (2017): Rapid and cost-effective high-throughput sequencing for identification of germline mutations of *BRCA1* and *BRCA2*. Journal of Human Genetics, 62, 561–567

Alberts, B. (2013): Molecular biology of the cell. Garland Science

Amsbio (2016): MagSi-NGS^{PREP} Plus.

Au C. H., Ho D. N., Kwong A., Chan T. L. i Ma E. S. K. (2017): BAMClipper: removing primers from alignments to minimize false-negative mutations in amplicon next-generation sequencing. Scientific Reports, 7, 1567

Bioo Scientific (2016): NEXTflex™ *BRCA1* & *BRCA2* Amplicon Panel for FFPE lllumina® Sequencing.

Castéra L., Krieger S., Rousselin A., Legros A., Baumann J. J., Bruet O., Brault B. i sur. (2014): Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes. European Journal of Human Genetics, 22, 1305–1313

Cavanagh H. i Rogers, K. M. A. (2015): The role of *BRCA1* and *BRCA2* mutations in prostate, pancreatic and stomach cancers. Hereditary Cancer in Clinical Practice, 13, 16

Cooper G. M. i Hausman R. E. (2004): The cell. Oxford University Press

D'Argenio V., Esposito M. V., Telese A., Precone V., Starnone F., Nunziato M., Cantiello P. i sur. (2015) : The molecular analysis of *BRCA1* and *BRCA2*: Next-generation sequencing supersedes conventional approaches. Clinica Chimica Acta, 446, 221–225

Dietrich D., Uhl B., Sailer V., Holmes E. E., Jung M., Meller S. i Kristiansen G. (2013): Improved PCR Performance Using Template DNA from Formalin-Fixed and Paraffin-Embedded Tissues by Overcoming PCR Inhibition. PLOS ONE, 8, 1-10

Do H. i Dobrovic A. (2015): Sequence artifacts in DNA from formalin-fixed tissues: Causes and strategies for minimization. Clinical Chemistry, 61, 64-71

Donepudi M. S., Kondapalli K., Amos S. J. i Venkanteshan P. (2014): Breast cancer statistics and markers. Journal of Cancer Research and Therapeutics, 10, 506–11

Dong Z., Dong H., Zhong X., Peng Z., Zhu X., Sun Y., Chen Y. i sur. (2015): Development of a Comprehensive NGS Workflow for the Analysis of Tumor *BRCA1* and *BRCA2* Mutations and Large Rearrangements. Journal of Genetics and Genome Research, 2, 1-5

Enyedi M. Z., Jaksa G., Pintér L., Sükösd F., Gyuris Z., Hajdu A., Határvölgyi E. i sur. (2016): Simultaneous detection of *BRCA* mutations and large genomic rearrangements in germline DNA and FFPE tumor samples. Oncotarget, 7, 61845-61859

Fackenthal J. D. i Olopade O. I. (2007): Breast cancer risk associated with *BRCA1* and *BRCA2* in diverse populations. Nature Reviews Cancer, 7, 937–948

Gagan J. i Van Allen E. M. (2015): Next-generation sequencing to guide cancer therapy. Genome Medicine, 7, 80

García-Alcalde F., Okonechinkov K., Carbonell J., Cruz L. M., Götz S., Tarazona S., Dopazo J., Meyer T. F. i Conesa A. (2012): Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics, 28, 2678–2679

Gelband H., Jha P., Sankaranarayanan R. i Horton S. (2015): Cancer: Disease Control Priorities. World Bank Group, 1–363

Godet I. i Gilkes D. M. (2017): *BRCA1* and *BRCA2* mutations and treatment strategies for breast cancer. Integrative Cancer Science and Therapies, 106, 201–207

Goodwin S., McPherson J. D. i McCombie W. R. (2016): Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics, 17, 333–351

Gudmundsdottir K. i Ashworth A. (2006): The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability. Oncogene, 25, 5864–5874

Hall J. M., Lee M. K., Newman B., Morrow J. E., Anderson L. A., Huey B. i King M. C. (1990): Linkage of early-onset familial breast cancer to chromosome 17q21. Science, 250, 1684–1689

Hanahan D., Weinberg R. A. (2011): Hallmarks of cancer: the next generation. Cell, 144, 646-674

Heather J. M. i Chain, B. (2016): The sequence of sequencers: The history of sequencing DNA. Genomics, 107, 1–8

Illumina (2015): Overview of Indexed Sequencing on the NextSeq, MiSeq, and HiSeq Platforms.

Illumina (2016): An Introduction to Next-Generation Sequencing Technology.

Illumina (2016): Benefits of NGS Targeted Resequencing.

Illumina (2016): MiniSeq System Denature and Dilute Libraries Guide.

Kamps R., Brandão R. D., van den Bosch B. J., Paulussen A. D. C., Xanthoulea S., Blok M. J. i Romano A. (2017): Next-Generation Sequencing in Oncology: Genetic Diagnosis, Risk Prediction and Cancer Classification. International Journal of Molecular Sciences, 18, 308

Kang H. P., Maguire J. R., Chu C., Haque I. S., Lai H., Mar-Heyming R., Ready K. i sur. (2016): Design and validation of a next generation sequencing assay for hereditary *BRCA1* and *BRCA2* mutation testing. PeerJ, 4, e2162

Koboldt D. C., Chen K., Wylie T., Larson D. E., McLellan M. D., Mardis E. R., Weinstock G. M., Wilson R. K. i Ding L. (2009): VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics, 25, 2283–2285

Lek M. i sur. (2017): Analysis of protein-coding genetic variation in 60,706 humans. Nature, 106, 201–207

Li H. i Durbin R. (2009): Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25, 1754–1760

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G. i Durbin R. (2009): The Sequence Alignment/Map format and SAMtools. Bioinformatics, 25, 2078–2079

Life technologies (2015): Qubit® dsDNA BR Assay Kits.

Makki J. (2015): Diversity of breast carcinoma: Histological subtypes and clinical relevance. Clinical Medicine Insights: Pathology, 8, 23–31

Markoulatos P., Siafakas N. i Moncany M. (2002): Multiplex polymerase chain reaction: A practical approach. Journal of Clinical Laboratory Analysis, 16, 47-51

Mehrgou A. i Akouchekian M. (2016): The importance of *BRCA1* and *BRCA2* genes mutations in breast cancer development. Medical Journal of the Islamic Republic of Iran, 1–12

Meldrum C., Doyle M. A. i Tothill R. W. (2011): Next-Generation Sequencing for Cancer Diagnostics: a Practical Perspective. The Clinical Biochemist Reviews, 32, 177–195

Metzker M. L. (2010): Sequencing technologies — the next generation. Nature Reviews Genetics, 11, 31–46

Miki Y., Swensen J., Shattuck-Eidens D., Futreal P. A., Harshman K., Tavtigian S., Liu Q. i sur. (1994): A Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene *BRCA1*. Science, 266, 66–71

Natrajan R. i Reis-Filho J. S. (2011): Next-generation sequencing applied to molecular diagnostics. Expert Review of Molecular Diagnostics, 11, 425–444

New England Biolabs (2016): NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®.

Nicholl D. S. T. (2008): An Introduction to Genetic Engineering. Cambridge University Press

Pan W., Byrne-Steele M., Wang C., Lu S., Clemmons S., Zahorchak R. J. i Han J. (2014): DNA polymerase preference determines PCR priming efficiency. BMC Biotechnology, 14, 1-16

Pecorino L. (2012): Molecular Biology of Cancer. Oxford University Press

Petrucelli N., Daly M. B. i Feldman G. L. (2010): Hereditary breast and ovarian cancer due to mutations in *BRCA1* and *BRCA2*. Genetics in Medicine, 12, 245–259

Polyak K. (2007): Breast cancer: origins and evolution. The Journal of Clinical Investigation, 117, 3115-3163

Powell S. N. i Kachnic L. A. (2003): Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication fidelity and the cellular response to ionizing radiation. Oncogene, 22, 5784–5791

Qiagen (2006): DNeasy® Blood & Tissue Handbook.

Qiagen (2014): GeneRead™ DNA FFPE Handbook.

Robinson J. T., Thorvaldsdóttir H., Winckler W., Guttman M., Lander E. S., Getz G. i Mesirov J. P. (2011): Integrative Genomics Viewer. Nature Biotechnology, 29, 24–6

Roy R., Chun J. i Powell S. N. (2011): BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nature Reviews Cancer, 12, 68–78

Shin S., Hwang I. S., Lee S. T. i Choi J. R. (2016): Evaluation of an amplicon-based next-generation sequencing panel for detection of *BRCA1* and *BRCA2* genetic variants. Breast Cancer Research and Treatment, 158, 433-440

Teng L., Zheng Y. i Wang H. (2008): *BRCA1/2* associated hereditary breast cancer. Journal of Zhejiang University SCIENCE B, 9, 85–89

Thermo Fisher Scientific (2015): Qubit® dsDNA BR Assay Kit.

Wallace A. J. (2016): New challenges for *BRCA* testing: a view from the diagnostic laboratory. European Journal of Human Genetics, 24, 10–18

Yoshida K. i Miki Y. (2004): Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. Cancer Science, 95, 866–871

http://op.pgx.ca/

http://www.onkologija.hr

www.researchgate.net

http://www.wcrf.org

http://www.who.int

8. Prilozi

Popis priloga:

- Prilog 1. Tablica dizajniranih početnica gena *BRCA1* i njihovih parametara
- Prilog 2. Tablica dizajniranih početnica gena *BRCA2* i njihovih parametara

Prilog 1. Tablica dizajniranih početnica gena *BRCA1* i njihovih parametara

Naziv amplikona	Duljina uzvodne početnice (pb)	Duljina nizvodne početnice (pb)	GC%-tak uzvodne početnice	GC%-tak nizvodne početnice	Tm uzvodne početnice (°C)	Tm nizvodne početnice (°C)	Veličina amplikona bez početnica (pb)	Veličina amplikona s početnicama (pb)
BRCA1_exon2	25	28	36	21	53	52	161	214
BRCA1_exon3	24	23	38	43	55	56	143	190
BRCA1_exon4	24	20	42	55	56	57	198	242
BRCA1_exon5	23	25	43	36	54	54	198	246
BRCA1_exon6_1	25	24	44	38	56	54	145	194
BRCA1_exon6_2	21	23	52	39	56	52	177	221
BRCA1_exon7	27	23	37	43	55	55	203	253
BRCA1_exon8	25	26	36	31	54	50	196	247
BRCA1_exon9	25	21	44	43	57	54	195	241
BRCA1_exon10_1	24	27	44	41	56	58	175	225
BRCA1_exon10_2	25	22	36	45	54	57	182	229
BRCA1_exon10_3	25	25	36	44	56	58	180	230
BRCA1_exon10_4	24	25	50	36	59	52	155	204
BRCA1_exon10_5	21	24	48	42	57	55	175	220
BRCA1_exon10_6	26	24	42	46	55	57	170	220
BRCA1_exon10_7	23	22	43	41	53	54	175	220
BRCA1_exon10_8	25	24	40	42	55	56	158	207
BRCA1_exon10_9	25	25	40	40	55	55	157	207
BRCA1_exon10_10	25	24	40	42	55	56	155	204
BRCA1_exon10_11	25	24	44	42	55	56	139	188
BRCA1_exon10_12	24	25	42	40	56	55	156	205
BRCA1_exon10_13	25	25	44	40	56	56	153	203
BRCA1_exon10_14	22	22	45	45	56	54	166	210
BRCA1_exon10_15	24	25	42	40	56	54	154	203
BRCA1_exon10_16	24	22	42	41	57	54	161	207
BRCA1_exon10_17	25	24	44	42	58	56	160	209
BRCA1_exon10_18	25	25	40	40	55	56	134	184
BRCA1_exon10_19	25	26	40	38	56	56	131	182
BRCA1_exon10_20	25	22	40	41	55	55	133	180
BRCA1_exon10_21	24	25	46	40	56	56	154	203
BRCA1_exon10_22	26	25	42	40	56	56	159	210
BRCA1_exon10_23	25	25	40	40	56	56	151	201
BRCA1_exon10_24	25	26	40	46	55	59	144	195
BRCA1_exon10_25	20	25	45	40	54	57	159	204
BRCA1_exon10_26	25	22	40	41	55	54	156	203
BRCA1_exon10_27	25	25	40	40	56	57	159	209
BRCA1_exon10_28	24	24	46	46	57	58	147	195

BRCA1_exon10_29	25	26	40	35	55	55	140	191
BRCA1_exon10_30	23	23	43	48	55	58	164	210
BRCA1_exon10_31	25	25	44	40	57	56	153	203
BRCA1_exon10_32	24	23	50	43	59	56	112	159
BRCA1_exon10_33	24	25	42	48	55	58	139	188
BRCA1_exon10_34	25	25	40	48	56	59	142	192
BRCA1_exon10_35	24	24	42	46	54	56	160	208
BRCA1_exon10_36	25	25	40	40	55	56	143	193
BRCA1_exon11	25	21	44	48	58	54	199	245
BRCA1_exon12_1	20	25	60	40	58	56	165	210
BRCA1_exon12_2	24	25	46	40	58	56	203	252
BRCA1_exon13	23	25	48	40	56	56	148	196
BRCA1_exon14_1	23	20	48	45	56	55	180	223
BRCA1_exon14_2	25	27	40	37	55	55	173	225
BRCA1_exon15_1	24	20	42	45	55	55	124	168
BRCA1_exon15_2	25	25	40	44	57	57	149	199
BRCA1_exon16_1	22	26	41	31	54	54	158	206
BRCA1_exon16_2	22	24	46	46	56	56	156	202
BRCA1_exon16_3	27	25	41	48	56	58	121	173
BRCA1_exon17	25	23	40	48	55	58	127	175
BRCA1_exon18	23	25	43	44	55	57	162	210
BRCA1_exon19	20	20	40	45	51	54	169	209
BRCA1_exon20	24	22	42	45	55	57	144	190
BRCA1_exon21	22	24	59	38	60	55	143	189
BRCA1_exon22	24	19	46	58	56	56	132	175
BRCA1_exon23	23	25	44	40	53	56	134	182
BRCA1_exon24_1	23	22	52	41	58	52	162	207
BRCA1_exon24_2	25	21	48	57	57	58	144	190

Prilog 2. Tablica dizajniranih početnica gena *BRCA2* i njihovih parametara

Naziv amplikona	Duljina uzvodne početnice (pb)	Duljina nizvodne početnice (pb)	GC%-tak uzvodne početnice	GC%-tak nizvodne početnice	Tm uzvodne početnice (ºC)	Tm nizvodne početnice (ºC)	Veličina amplikona bez početnica (pb)	Veličina amplikona s početnicama (pb)
BRCA2_exon2_1	29	20	38	45	56	53	128	177
BRCA2_exon2_2	25	25	40	40	55	56	163	213
BRCA2_exon3_1	25	25	44	40	56	55	169	219
BRCA2_exon3_2	25	23	36	48	55	55	139	187
BRCA2_exon4_1	20	27	45	41	53	58	165	212
BRCA2_exon4_2	25	25	44	44	56	56	182	232
BRCA2_exon5	20	25	40	40	51	56	151	196
BRCA2_exon6	22	24	45	42	54	55	194	240
BRCA2_exon7	22	22	41	45	53	55	172	216
BRCA2_exon8	21	25	38	44	52	56	223	269
BRCA2_exon9_1	25	25	44	40	55	50	144	194
BRCA2_exon9_2	25	22	36	45	54	54	185	232
BRCA2_exon10_1	26	24	31	42	53	53	154	204
BRCA2_exon10_2	24	25	42	40	56	56	178	227
BRCA2_exon10_3	20	25	40	40	54	55	175	220
BRCA2_exon10_4	22	25	45	40	56	55	167	214
BRCA2_exon10_5	24	25	42	40	54	55	177	226
BRCA2_exon10_6	21	25	52	40	56	55	173	219
BRCA2_exon10_7	25	20	40	55	56	56	202	247
BRCA2_exon10_8	25	21	44	52	58	58	189	235
BRCA2_exon10_9	20	22	50	41	55	53	199	241
BRCA2_exon10_10	20	22	40	41	53	55	218	260
BRCA2_exon11_1	26	25	31	36	52	54	143	194
BRCA2_exon11_2	25	25	40	40	55	54	188	238
BRCA2_exon11_3	25	25	40	40	55	56	176	226
BRCA2_exon11_4	25	25	40	40	56	57	168	218
BRCA2_exon11_5	26	25	42	44	57	56	198	249
BRCA2_exon11_6	25	25	40	44	56	56	116	166
BRCA2_exon11_7	23	25	35	40	54	55	121	169
BRCA2_exon11_8	25	24	40	40	55	56	192	241
BRCA2_exon11_9	25	25	44	40	57	55	136	186
BRCA2_exon11_10	25	22	44	41	57	53	202	249
BRCA2_exon11_11	25	24	40	42	56	55	180	229
BRCA2_exon11_12	25	25	40	40	55	55	169	219
BRCA2_exon11_13	23	25	48	40	56	56	176	224
BRCA2_exon11_14	25	24	40	42	55	55	192	241
BRCA2_exon11_15	20	23	40	43	52	55	128	171
BRCA2_exon11_16	25	25	40	40	56	55	196	246

BRCA2_exon11_17	25	24	32	42	54	56	180	229
BRCA2_exon11_18	25	25	40	40	56	57	174	224
BRCA2_exon11_19	25	25	40	40	56	55	184	234
BRCA2_exon11_20	24	21	42	52	56	56	139	184
BRCA2_exon11_21	20	25	40	40	52	56	205	250
BRCA2_exon11_22	24	25	42	40	55	56	199	248
BRCA2_exon11_23	25	25	40	40	57	56	139	189
BRCA2_exon11_24	26	25	23	40	49	55	207	258
BRCA2_exon11_25	25	22	40	41	55	55	150	197
BRCA2_exon11_26	25	22	40	41	55	54	165	212
BRCA2_exon11_27	20	22	40	41	53	53	175	217
BRCA2_exon11_28	25	25	40	44	56	57	175	225
BRCA2_exon11_29	22	25	41	44	53	57	188	235
BRCA2_exon11_30	21	25	43	48	53	58	198	244
BRCA2_exon11_31	26	22	42	50	56	56	160	208
BRCA2_exon11_32	25	24	48	50	58	59	102	151
BRCA2_exon11_33	25	25	40	40	56	56	197	247
BRCA2_exon11_34	25	25	40	40	56	56	168	218
BRCA2_exon11_35	22	25	41	40	55	54	119	166
BRCA2_exon11_36	25	20	40	40	55	51	203	248
BRCA2_exon11_37	25	25	40	44	54	56	141	191
BRCA2_exon11_38	20	25	40	40	51	56	194	239
BRCA2_exon11_39	25	25	40	44	55	57	164	214
BRCA2_exon11_40	22	24	41	42	53	56	194	240
BRCA2_exon11_41	24	25	50	40	59	56	201	250
BRCA2_exon11_42	25	25	40	40	56	55	161	211
BRCA2_exon11_43	24	25	42	40	55	56	199	248
BRCA2_exon11_44	25	25	40	40	56	56	151	201
BRCA2_exon11_45	25	25	40	44	56	56	173	223
BRCA2_exon11_46	22	25	41	40	52	56	192	239
BRCA2_exon11_47	25	25	44	44	56	56	177	227
BRCA2_exon11_48	23	25	39	44	54	56	177	225
BRCA2_exon11_49	25	25	40	44	56	56	167	217
BRCA2_exon11_50	25	24	44	42	56	56	162	211
BRCA2_exon11_51	25	25	44	40	56	56	198	248
BRCA2_exon11_52	25	24	40	42	56	56	180	229
BRCA2_exon11_53	25	25	40	40	55	55	197	247
BRCA2_exon11_54	23	20	43	40	56	52	201	244
BRCA2_exon11_55	24	25	42	40	57	56	201	250
BRCA2_exon11_56	25	25	36	40	53	56	194	244
BRCA2_exon12_1	28	22	36	41	55	54	199	249
BRCA2_exon12_2	24	21	42	52	55	56	187	232
BRCA2_exon13_1	21	22	43	45	52	55	163	206

BRCA2_exon13_2	21	23	43	43	52	55	145	189
BRCA2_exon14_1	26	25	35	40	54	56	196	247
BRCA2_exon14_2	25	22	44	50	57	58	152	199
BRCA2_exon14_3	25	25	40	40	56	55	178	228
BRCA2_exon14_4	24	25	42	40	57	55	164	213
BRCA2_exon15_1	24	21	38	52	55	57	165	210
BRCA2_exon15_2	25	25	40	40	56	56	172	222
BRCA2_exon16_1	25	24	28	42	53	55	167	216
BRCA2_exon16_2	25	25	40	44	57	55	156	206
BRCA2_exon17_1	25	25	36	36	53	55	184	234
BRCA2_exon17_2	23	25	43	48	55	58	183	231
BRCA2_exon18_1	29	22	31	41	55	54	160	211
BRCA2_exon18_2	24	25	42	40	55	56	161	210
BRCA2_exon18_3	27	23	41	39	57	52	169	219
BRCA2_exon18_4	25	22	40	41	56	53	164	211
BRCA2_exon18_5	22	24	50	50	58	58	195	241
BRCA2_exon19_1	23	22	39	41	53	55	174	219
BRCA2_exon19_2	25	21	36	57	54	58	172	218
BRCA2_exon20_1	23	25	48	40	57	55	170	218
BRCA2_exon20_2	21	25	43	40	54	55	214	260
BRCA2_exon21_1	24	25	38	40	54	55	191	240
BRCA2_exon21_2	25	24	40	42	55	55	158	207
BRCA2_exon22_1	25	24	32	42	52	55	155	204
BRCA2_exon22_2	24	25	42	36	54	55	221	270
BRCA2_exon23_1	23	25	35	44	54	57	167	215
BRCA2_exon23_2	25	25	40	40	55	55	161	211
BRCA2_exon24_1	25	22	40	41	55	55	139	186
BRCA2_exon24_2	25	25	40	40	56	56	183	233
BRCA2_exon25_1	26	24	31	42	53	56	174	224
BRCA2_exon25_2	25	25	40	40	56	56	190	240
BRCA2_exon25_3	24	25	42	40	55	55	128	177
BRCA2_exon26_1	23	24	35	46	53	57	147	194
BRCA2_exon26_2	22	25	41	44	54	57	196	243
BRCA2_exon27_1	24	25	38	40	57	55	190	239
BRCA2_exon27_2	23	25	43	40	55	56	114	162
BRCA2_exon27_3	25	25	40	44	55	57	174	224
BRCA2_exon27_4	25	24	44	42	57	57	171	220
BRCA2_exon27_5	25	25	44	44	56	56	147	197
BRCA2_exon27_6	25	21	40	52	56	56	130	176
BRCA2_exon27_7	25	22	40	41	54	54	134	181
BRCA2_exon27_8	24	22	42	41	55	54	201	247

9. Životopis

Osobni podaci

- Ime i prezime: Martina Sajfert
- Datum rođenja: 1.4.1993.

Obrazovanje

- 2015. – 2018. **Diplomski sveučilišni studij molekularne biologije**
  Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet, Zagreb
- 2012. – 2015. **Preddiplomski sveučilišni studij molekularne biologije**
  Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet, Zagreb
- 2008. – 2012**.** **II. gimnazija, Zagreb**

Aktivnosti

- 2015./2016. demonstrator na kolegiju Animalna fiziologija
- 2015. predstavnik Biološkog odsjeka na Smotri Sveučilišta u Zagrebu
- 2015. polaznik 1. Hrvatskog simpozija biologa u zdravstvu
- 2013., 2014. i 2016. aktivni sudionik i organizator manifestacije Noć biologije

Stipendije

- 2016./2017. dobitnica Stipendije Grada Svete Nedelje prema kriteriju izvrsnosti