

Humane genetske spremembe in njihovo določanje: trenutno stanje in obeti za prihodnost

Human Genetic Variants and Their Analysis: Current State and Future Perspectives

Lara Slavec, Ksenija Geršak, Nataša Karas Kuželički,
Katarina Trebušak Podkrajšek

Izvleček

Humane genetske spremembe se pojavljajo v celotnem genomu, tako v kodirajočih kot tudi v nekodirajočih predelih. Opredeljujejo genetsko raznolikost med posamezniki in so v nekaterih primerih lahko vzrok različnih genetskih bolezni. Številne se pojavijo že v otroštvu. V prvem delu preglednega prispevka opisujemo različne vrste genetskih sprememb, od manjših sekvenčnih sprememb do velikih kromosomskih nepravilnosti. Sledi predstavitev pomembnejših pristopov (citogenetskih in molekularnih genetskih), ki se uporabljajo za določanje humanih genetskih sprememb v klinični praksi. Večji poudarek namenjamo metodam sekvenciranja, saj so omogočile vpogled v nukleotidno zaporedje celotnega genoma in pomenijo revolucijo na področju molekularnega genetskega diagnosticiranja. Izpostavili smo težave, ki se porajajo pri določanju genetskih sprememb z metodami sekvenciranja naslednje generacije (angl. *next generation sequencing*, NGS), in predstavili prednosti najnovejših metod sekvenciranja – tehnologij dolgih odčitkov. Dandanes precejšnega deleža humanih genetskih sprememb, povezanih z razvojem genetskih bolezni, še vedno niso odkrili. Zato predstavljamo nekaj možnih pristopov, s katerimi bi lahko izboljšali odkrivanje genetskih sprememb na individualni ravni.

Ključne besede: genetske spremembe, genetsko testiranje, sekvenciranje naslednje generacije, tehnologije dolgih odčitkov.

Abstract

Human genetic variants occur throughout the whole genome, both in the coding and non-coding regions. In addition to defining genetic diversity among individuals, they may, in some instances, cause various genetic diseases, many of which have their onset in childhood. In the first part of the review article, we describe different types of genetic variants, ranging from minor sequence changes to major chromosomal aberrations. This is followed by a presentation of some of the most important approaches (cytogenetic and molecular genetic) that are used to detect genetic variants in clinical practice. We put greater emphasis on sequencing methods, as they allowed us to gain insight into the nucleotide sequence of the entire genome, and therefore revolutionised the field of molecular genetic diagnostics. We also highlight the problems that arise in determining genetic variants through next-generation sequencing (NGS) methods and outline the benefits of the latest sequencing techniques – long-read sequencing. Today, a substantial proportion of human genetic variants associated with the development of genetic diseases, are still unknown. Therefore, we conclude the article by presenting some possible approaches that could improve the detection of genetic variants at the individual level.

Key words: genetic variants, genetic testing, next-generation sequencing, long-read sequencing.

Uvod

Genetske spremembe (spremembe v zapisu DNK) se v človeškem genomu pojavljajo v različnih razsežnostih, od majhnih sprememb, npr. sprememb enega nukleotida, do velikih sprememb, npr. kromosomskih nepravilnosti (1). Humane genetske spremembe opredeljujejo genetsko raznolikost med posamezniki tako znotraj populacije kot tudi med populacijami (2). Vplivajo namreč na fenotipske razlike med ljudmi ter povzročajo različne odzive na zdravila in vplive okolja. Nekatere spremembe so lahko tudi vzrok nastanka različnih bolezni, med katerimi številne nastopijo že v otroštvu. Včasih lahko z določanjem genetskih sprememb ocenimo nagnjenost k bolezni ali napovemo njen potek (3).

Genetske spremembe so lahko bodisi prirojene bodisi pridobljene. Prirojene se preko zarodnih celic staršev prenašajo na potomce in se nahajajo v vseh celicah posameznika. Nasprotno pa pridobljene spremembe nastanejo le v določenih celicah posameznika, in sicer kadar koli v življenju. Če nastanejo v somatskih celicah, se ne prenašajo na potomstvo. Čeprav je večina humanih genetskih sprememb benignih (niso vpletene v razvoj bolezni), so v določenem deležu tudi patogene in sodelujejo pri nastanku genetskih bolezni (4).

Poznamo štiri glavne tipe humanih genetskih bolezni: monogenske bolezni (posledica napak v zapisu DNK enega gena), kromosomske bolezni (posledica sprememb v številu ali strukturi kromosomov), kompleksne oz. večfaktorske bolezni (posledica kombinacije delovanja več genetskih in okoljskih dejavnikov) in mitohondrijske bolezni (posledica napak v mitohondrijski DNK) (5).

Raziskovalci so se v preteklosti posluževali različnih strategij odkrivanja genetskih označevalcev, ki vplivajo na nastanek genetskih bolezni, npr. analiz vezanega dedovanja, genetskih asociacijskih raziskav/študij primerov in kontrol (npr. analiz kandidatnih genov v populaciji, analiz kandidatnih genov

znotraj družin (npr. test TDT) in asociacijskih raziskav na celotnem genomu (GWAS)) (4,6). Čeprav so ti raziskovalni pristopi dokaj preprosti, je odkrivanje patogenih različic DNK in njihovih posledic lahko zelo zapleteno in zahteva desetletja raziskovanja. Vsekakor pa je uspešnost prepoznavanja vzročnih genetskih različic določenih bolezni v veliki meri odvisna od napredka genomskih tehnologij (7).

Vrste humanih genetskih sprememb

Genetske spremembe se pojavljajo na celotnem genomu. Humani genom pogosto razdelimo na kodirajoče regije DNK (manj kot 2 % genoma), ki se prepišejo v mRNK in prevedejo v proteine, ter nekodirajoče regije DNK (približno 98 % genoma) (8). Včasih so raziskovalci nekodirajočo DNK imenovali tudi »odpadna« DNK (angl. *junk DNA*), dandanes pa je znano, da ima velik del nekodirajoče DNK pomembno regulatorno vlogo (npr. kodira različne RNK, sodeluje pri aktivaciji genov, urejanju strukture kromatina itd.) (9). Različne vrste genetskih sprememb se v diploidnem humanem genomu, ki obsega približno 6 Gbp, pojavljajo z različnimi frekvencami. Njihov doprinos k nastanku genetskih bolezni je različen (7). V nadaljevanju na kratko opisujemo vrste humanih genetskih sprememb. Nekatere od shematsko prikazujemo na Sliki 1.

Različice posameznih nukleotidov (angl. *single nucleotide variants, SNV*) oz. polimorfizmi posameznih nukleotidov (angl. *single nucleotide polymorphisms, SNP*)

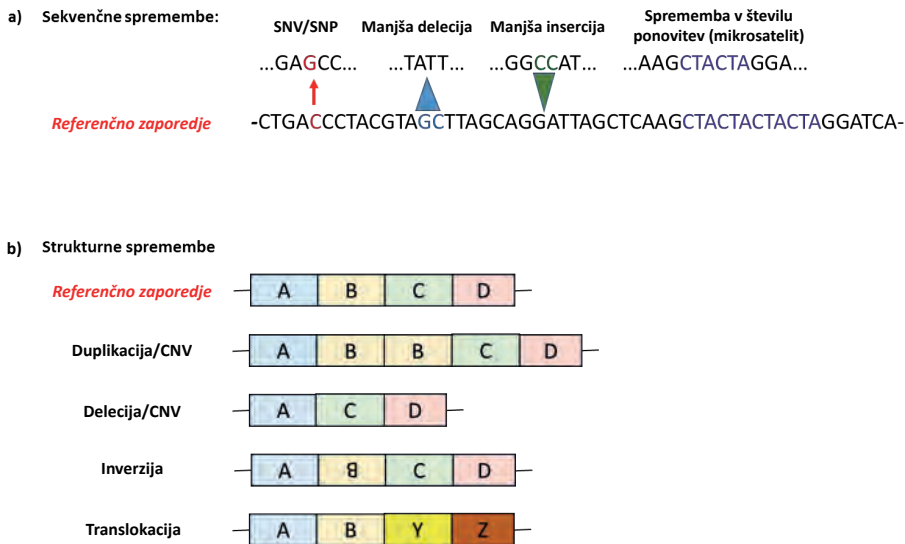
Različice posameznih nukleotidov (SNV) ali polimorfizmi posameznih nukleotidov (SNP) so substitucije oz. zamenjave enega nukleotida (10).

Poznamo dva različna tipa zamenjav, in sicer tranzicije (zamenjava purina s purinom (A↔G) ali pirimidina s pirimidinom (T↔C)) in transverzije (zamenjava dvoobročnega purina z enoobročnim pirimidinom ali obratno) (11). SNP so najpogostejše in najbolj definirane spremembe v humanem genomu. V povprečju se pojavi en SNP na 1000 baznih parov, kar pomeni, da jih je v celotnem humanem genomu približno 4–5 milijonov (12). Po številčnosti presegajo vse ostale genetske spremembe v razmerju skoraj 7 proti 1 (7).

Zamenjava enega nukleotida v kodirajočih regijah DNK lahko vodi do sinonimnih oz. tihih sprememb (spremenjen kodon kodira isto aminokislino) ali do nesinonimnih sprememb, kot so drugačnosmiselne spremembe (spremenjen kodon kodira drugačno aminokislino – bodisi podobno ali zelo različno) in nesmiselne spremembe (sprememba nukleotida vodi v nastanek prezgodnjega stop kodona ali pa stop kodon odstrani). Predvsem nesinonimne, v nekaterih primerih pa tudi sinonimne spremembe lahko pomembno spremenijo končni proteinski produkt in vplivajo na njegovo funkcijo (10,11,13). Če se SNV pojavi v nekodirajočem delu DNK, npr. v promotorski regiji gena ali v intronskem zaporedju, lahko v prvem primeru spremenijo vezavno mesto za transkripcijski dejavnik, v drugem pa povzročijo nepravilno izrezovanje intronov. Tudi na ta način lahko spremenijo funkcijo proteina (10).

Manjše insercije in delecije (angl. *indels*)

Pri insercijah gre za vrinek nukleotidov, pri delecijah pa za njihovo izgubo. Manjše insercije in delecije so dolge 1–49 bp (14). Tako kot SNV so tudi te genetske spremembe v genomu zelo pogoste. Pojavijo se namreč od 700.000- do 800.000-krat (7). Če je število nukleotidov, ki se vrine oz. izreže iz kodirajoče sekvence DNK, večkratnik števila tri, pride do dodajanja oz. odstranjanja



SLIKA 1. NEKATERE GENETSKE SPREMEMBE, KI JIH NAJDEMO V HUMANEM GENOMU. A) SEKVENČNE SPREMEMBE, KI VKLJUČUJE ENEGA ALI NEKAJ NUKLEOTIDOV IN B) STRUKTURNE SPREMEMBE, KI VKLJUČUJEJO VEČJE SEGMENTE DNK (Y IN Z STA SEGMENTA DNK IZ DRUGEGA KROMOSOMA); V OBEH PRIMERIH SO SPREMEMBE PRIKAZANE V PRIMERJAVI Z REFERENČNIM ZAPOREDJEM.

FIGURE 1. SOME OF THE GENETIC VARIANTS FOUND IN THE HUMAN GENOME. A) SEQUENCE VARIANTS INVOLVING ONE OR MORE NUCLEOTIDES AND B) STRUCTURAL VARIANTS INVOLVING LARGE DNA SEGMENTS (Y AND Z REPRESENT DNA SEGMENTS FROM ANOTHER CHROMOSOME); IN BOTH CASES, THE VARIANTS ARE PRESENTED IN RELATION TO THE REFERENCE SEQUENCE.

celotnih kodonov in s tem posameznih aminokislin. Končni proteinski produkt ima spremenjeno funkcijo, a je v nekaterih primerih še vedno lahko funkcionalen. Če pa to število ni deljivo s tri, pride do spremembe bralnega okvirja in posledično do popolne spremembe strukture proteina. Posledica je nastanek nefunkcionalnega proteina ali ta sploh ne nastane (10,11,15).

Strukturne spremembe (angl. structural variants)

O strukturnih spremembah govorimo, ko so spremembe med dvema genomoma večje od 50 bp (14). V humanem genomu jih je 23.000–28.000 (7). Med strukturne spremembe uvrščamo večje delecije in insercije, inverzije (zamenjave smeri odseka DNK na nekem položaju), duplikacije (del kromosoma se podvoji), translokacije (premestitve kromosomskih odsekov znotraj kromo-

soma ali med dvema kromosomoma) in različice v številu kopij (angl. *copy-number variations*, CNV) (10,16). CNV so definirani kot večji odseki DNK, ki se v različnih humanih genomih pojavijo v različnem številu kopij (od nič do več različic). Največji CNV vključujejo ponovitve več genov oz. alelov (16).

Spremembe v številu ponovitev (angl. repeat variants)

Humani genom vsebuje tudi veliko število ponovitvenih sekvenc. Nahajajo se predvsem v nekodirajočih regijah DNK (telomerah, centromerah itd.) in predstavljajo približno 50 % genoma. Število ponovitev se med humanimi genomi zelo razlikuje (10,17). Ponavljajoča se zaporedja lahko razdelimo v dve veliki skupini: mobilni ali transpozicijski elementi (angl. *transposable elements*) in tandemске ponovitve (angl. *tan-*

dem repeats). Transpozicijski elementi so razpršena ponovljena zaporedja, za katera je značilno, da se prestavljajo po genomu. Poznamo dve skupini transpozicijskih elementov:

- retroelementi, ki se prestavljajo po genomu z vmesno pretvorbo v RNK-intermediat. Med transpozicijo se ne izrežejo iz DNK, zato se število ponovitev povečuje (mehanizem »kopiraj in prilepi«). Mednje uvrščamo:
 - kratke razpršene jedrne elemente (angl. *short interspersed nuclear element*, SINE);
 - dolge razpršene jedrne elemente (angl. *long interspersed nuclear element*, LINE);
 - dolge terminalne ponavljajoče se retrotranspozone oz. LTR-retrotranspozone (angl. *long terminal repeat retrotransposons*);
- DNK-transpozoni, ki se iz donorske DNK izrežejo. Pri vstavitvi na drugo mesto v genomu zato ne pride do povečanja števila ponovitev (mehanizem »izreži in prilepi«) (18).

V nasprotju z naključno razpršenimi ponovitvami je za tandemске ponovitve značilno, da se določen odsek DNK zaporedoma ponavlja. Enote ostanejo medsebojno povezane in se ne prestavljajo po genomu. Med tandemске ponovitve štejemo:

- satelitno DNK (ponavljajoča se enota je dolga 100–500 bp);
- minisatelite (ponovitve segmenta, dolgega 10–100 bp) in
- mikrosatelite (ponovitve segmenta, dolgega 1–9 bp).

Med minisatelite uvrščamo npr. spremenljivo število tandemskih ponovitev (angl. *variable number tandem repeat*, VNTR), med mikrosatelite pa npr. kratko tandemsko ponovitev (angl. *short tandem repeat*, STR) (10,17,19). Čeprav se tandemске ponovitve navadno nahajajo v kromosomskih regijah brez genov oz. tako imenovanih »genskih puščavah«, so nekatere prisotne tudi v funkcionalnih delih genoma (kodirajočih ali regulatornih). Predvsem mikrosateliti so zaradi visoke stopnje interindividualne variabilnosti

pomembni biološki označevalci. Poleg tega, da so vpleteni v nastanek nekaterih genetskih bolezni, se v forenziki uporabljajo za identifikacijo oseb, koristni pa so tudi za študije kratkoročne evolucije človeštva (17,19).

Spremembe v številu kromosomov (angl. *numerical abnormalities*)

Med humane genetske spremembe uvrščamo tudi spremembe v številu kromosomov. V primeru, da celica vsebuje več kot dva popolna seta humanega haploidnega genoma (69 kromosomov ali več), govorimo o poliploidiji. Primer je triploidija (trije haploidni seti kromosomov), ki naj bi se pojavila pri 2–3 % nosečnosti in navadno vodi v zgodnji spontani splav (20,21). Če število kromosomov ni mnogokratnik haploidnega števila kromosomov, govorimo o aneuploidiji. Večina aneuploidij ni združljivih z življenjem. Izjeme so nekatere trisomije (npr. trisomija 21 ali Downov sindrom) in monosomije (npr. monosomija X ali Turnerjev sindrom) (21,22).

Kadar genetske spremembe povzročijo nastanek nefunkcionalnega proteina ali povzročijo, da proteinski produkt sploh ne nastane, gre za škodljive (patogene) genetske spremembe, ki vodijo v določeno klinično sliko oz. nastanek genetske bolezni. Te bolezenske genetske spremembe imenujemo tudi mutacije (23).

Čeprav je med humanimi genetskimi spremembami daleč največ SNV, je analiza podatkov iz podatkovne baze Human Genome Mutation Database (HGMD) pokazala, da pri približno tretjini vseh patogenih genetskih sprememb ne gre za enonukleotidne substitucije (24). Dandanes je znano, da strukturne spremembe, ki prizadenejo gene, bistveno prispevajo k nastanku bolezni. Te namreč doprinejajo več razlik v baznih parih med dvema humanima genomoma kot katera koli druga vrsta humanih genetskih sprememb (25).

Strukturne spremembe navadno zavzemajo večje spremembe v sekvenci DNK, ki so zato tudi bolj škodljive (26). Tudi če se pojavljajo v nekodirajočih regijah DNK (kar je za njih značilno), lahko tam dodajo, spremenijo ali odstranijo regulatorno sekvenco in s tem spremenijo izražanje genov (27). V primerjavi s SNV naj bi imele strukturne spremembe vsaj 28-krat večjo verjetnost, da bodo vplivale na gensko izražanje (26). Če so strukturne spremembe prisotne v eksonih protein-kodirajočih genov, navadno pride do veliko bolj škodljivih posledic, saj izguba celotnega eksona onemogoči sintezo proteina (27). Pri identifikaciji genetskih dejavnikov tveganja za nastanek bolezni se moramo tako poleg SNV osredotočiti tudi na druge humane genetske spremembe in jih ne smemo podcenjevati.

Določanje humanih genetskih sprememb

Humane genetske spremembe določamo zaradi pridobivanja genetskih informacij posameznika in morebitnega kliničnega ukrepanja. Obstaja več zdravstvenih razlogov (indikacij), zakaj se odločamo za izvajanje genetskih preiskav. Primeri indikacij so:

- potrditev genetske spremembe pri simptomatskem bolniku,
- potrditev genetske spremembe pri brezsimptomnem bolniku,
- ugotavljanje prenašalstva za genetsko spremembo,
- testiranje pred rojstvom (predimplantacijsko in prenatalno),
- presejanje v populaciji in
- forenzična identifikacija.

Velika večina indikacij je tudi glavni razlog za izvajanje genetskih preiskav pri otrocih in njihovih družinskih članih. Ker obstajajo različne indikacije za genetsko testiranje, je z njimi povezana tudi pravilna izbira ustreznega testa. Poznamo diagnostične in prediktivne teste, teste prenašalstva ter farmakogenomske in genetske identifikacijske teste (28).

Pristope določanja humanih genetskih sprememb v grobem delimo na dve skupini: citogenetske in molekularne genetske (29). V Tabeli 1 so zbrane pomembne metode, ki se v klinični praksi uporabljajo za določanje humanih genetskih sprememb, in njihove osnovne značilnosti.

Citogenetske in molekularnocitogenetske preiskave

Te preiskave so usmerjene v prepoznavanje znanih in še neopisanih kromosomskih sprememb, torej velikih strukturnih sprememb in sprememb v številu kromosomov, ter različic v številu kopij (CNV) (29,30). Naštevamo in opisujemo jih v Tabeli 1.

Molekularne genetske preiskave

Za razliko od citogenetskih metod z molekularnimi genetskimi metodami zaznavamo manjše sekvenčne spremembe (29). Izjema je metoda hkratnega pomnoževanja od ligacije odvisnih sond (MLPA), ki omogoča tudi identifikacijo večjih znanih delecij, insercij in duplikacij (31). Poleg tega je dandanes razvoj najnovejših metod sekvenciranja (metod 3. generacije) usmerjen v zmožnost zaznavanja večjih genetskih sprememb, tudi strukturnih (32). Molekularne metode, ki se uporabljajo v praksi, so našteje in opisane v Tabeli 1. Nekatere med njimi se uporabljajo manj kot druge. Primer je PCR z analizo polimorfizmov dolžin restriksijskih fragmentov (PCR-RFLP), ena prvih genotipizacijskih metod, ki je danes zaradi razvoja enostavnejših in zanesljivejših metod (npr. PCR v realnem času (qPCR)) skoraj obsolentna.

Metode se bistveno razlikujejo v tem, ali omogočajo prepoznavanje samo znanih sprememb ali tudi tistih, ki še niso bile opisane (29). Vsem je skupno,

Metoda	Osnovni princip metode	Tip analiziranih genetskih sprememb
citogenetske in molekularno-citogenetske preiskave		
klasična kariotipizacija	analiza kariograma – mikroskopske slike metafaznih kromosomov, ki so obarvani in razporejeni po velikosti	znane in še neopisane kromosomske spremembe (velike strukturne spremembe ali spremembe v številu kromosomov)
fluorescentna hibridizacija in situ (FISH)	označevanje in prepoznavanje posameznih delov kromosomov z uporabo fluorescentnih sond	znane mikrolelecije in spremembe v številu kromosomov
primerjalna genomna hibridizacija na mikromrežah (aCGH)	molekularna kariotipizacija – tekmovalne fluorescentno označene preiskovančeve DNK in referenčne DNK za vezavo na mikromrežo; analiza sprememb v številu kopij na podlagi sprememb v intenzivnosti signala	znane in še neopisane različice v številu kopij (CNV)
molekularne genetske preiskave		
verižna reakcija s polimerazo (PCR) poleg opisanih metod obstajajo še številne druge: npr. hkratni PCR (angl. <i>multiplex</i> PCR), vgnezdjeni PCR (angl. <i>nested</i> PCR) itd.	PCR z analizo polimorfizmov dolžin restrikcijskih fragmentov (PCR-RFLP): PCR pomnožitev odseka z znano genetsko spremembo, specifična cepitev z restrikcijskim encimom in detekcija s pomočjo elektroforeze	znane sekvenčne spremembe v restrikcijskem mestu
	alelnospecifični PCR: pomnoževanje odseka DNK z dvema setoma oligonukleotidnih začetnikov, komplementarnima mutiranemu ali normalnemu alelu, in detekcija odsekov s pomočjo elektroforeze	znane sekvenčne spremembe (SNV in manjše delecije/insercije)
	PCR v realnem času (qPCR): pomnoževanje specifičnega odseka DNK z oligonukleotidnimi začetniki in z dvema fluorescentno označenima sondama, komplementarnima mutiranemu ali normalnemu alelu; vsaka sonda ima vezano svoje barvilo, kar omogoča razlikovanje med različnimi genotipi.	znane sekvenčne spremembe (predvsem SNV)
DNK mikromreže	na čip z vezanimi enoverižnimi tarčnimi odseki DNK hibridiziramo fluorescentno označeno testno DNK; sledi spiranje nevezanih fragmentov, presvetlitev z laserjem in detekcija fluorescence.	prepoznavanje velikega števila sekvenčnih sprememb hkrati (znanih in še neopisanih)
preizkus ligiranja oligonukleotidov (OLA)	po PCR pomnožitvi odsekov DNK uporabimo dve različno dolgi oligonukleotidni sondi, ena je komplementarna normalnemu zaporedju, druga pa mutiranemu zaporedju; uporabimo še skupno, označeno sondo, ki se na tarčno zaporedje veže poleg prve oz. druge sonde; v primeru popolnega prileganja dveh sond pride do njune ligacije; sledita denaturacija DNK in detekcija z elektroforezo (ali encimskoimunsko)	znane sekvenčne spremembe (predvsem SNV)
metoda hkratnega pomnoževanja od ligacije odvisnih sond (MLPA)	temelji na PCR reakciji z uporabo enega para oligonukleotidnih začetnikov in posebnih DNK-sond (vsaka je sestavljena iz dveh delov, ki vsebujeta sekvenco, komplementarno neposredni bližini tarčnega odseka, ter sekvenco, kamor se vežejo oligonukleotidni začetniki); reakcija je uspešna le, ko pride do hibridizacije obeh delov sond in njune ligacije; sledi ločitev pomnožkov in detekcija s kapilarno elektroforezo	znane večje in manjše delecije in duplikacije
analiza taliine krivulje visoke ločljivosti (HRM)	po PCR pomnožitvi odsekov DNK dvojno vijačnico DNA označimo z interkalirajočim barvilom in počasi segrevamo do tališča sekvence, kar vodi do denaturacije, ki se kaže z zmanjšanja signala fluorescence	znane in še neopisane sekvenčne spremembe
izvedbe sekvenciranja	sekvenciranje po Sangerju oz. dideoksi metoda (1. generacija sekvenciranja): metoda je podrobno opisana v podglavju molekularne genetske preiskave	znane in še neopisane sekvenčne spremembe (predvsem SNV, manjše delecije in insercije)
	sekvenciranje naslednje generacije (NGS) (2. generacija sekvenciranja): metoda podrobno opisana v podglavju Molekularne genetske preiskave	znane in še neopisane sekvenčne spremembe (SNV in manjše delecije/insercije) obsežnih genomskih področij, omejeno zaznavanje večjih strukturnih sprememb
	tehnologije dolgih odčitkov (3. generacija sekvenciranja): osnovni princip metode, opisan v poglavju Omejitve pri določanju humanih genetskih sprememb s sekvenciranjem naslednje generacije.	znane in še neopisane sekvenčne ter strukturne spremembe

TABELA 1. GENETSKE PREISKAVE, KI SE UPORABLJAJO V KLINIČNI PRAKSI, NJIHOV OSNOVNI PRINCIP IN TIP ANALIZIRANIH GENETSKIH SPREMOMB (29–31,33,36–40).

TABLE 1. GENETIC METHODS USED IN CLINICAL PRACTICE, THEIR BASIC PRINCIPLE AND THE TYPE OF ANALYZED GENETIC VARIANTS (29–31,33,36–40).

da se pričnejo z izolacijo nukleinskih kislin (DNK, RNK) iz izbranega biološkega vzorca (polne krvi, brisa bukalne sluznice, kostnega mozga itd.). Navadno metode vključujejo pomnoževanje odsekov DNK (oz. RNK) z verižno reakcijo s polimerazo (PCR). Na ta način dobimo iz minimalne količine začetne DNK milijon ali več kopij odsekov, ki so pomembni za testiranje. Uspešnost reakcije PCR lahko preverimo z elektroforeznimi tehnikami, ki jih uporabljamo za ločevanje fragmentov DNK glede na dolžino. Temu sledi detekcija sekvenčnih sprememb (33).

Včasih so bile zelo pomembne presejalne metode, kot so analiza talilne krivulje visoke ločljivosti (HRM), PCR z analizo konformacijskih polimorfizmov enoverižnih DNK (PCR-SSCP), visokotlačna tekočinska kromatografija z denaturacijo (dHPLC), denaturirajoča gradientna gelska elektroforeza (DGGE) idr. Te metode omogočajo analizo omejenih genomske regije pri večjem številu preiskovancev hkrati. Z izjemo HRM se z razvojem modernih metod sekvenciranja v klinični praksi vse manj uporabljajo (34). Se pa še vedno uporabljajo v raziskovalne namene.

Revolucijo na področju molekularnega genetskega diagnosticiranja so namreč prinesle metode sekvenciranja oz. metode določanja nukleotidnega zaporedja. Vsem metodam sekvenciranja je skupno, da se končno zaporedje odseka primerja z zaporedjem referenčnega genoma. Na ta način prepoznamo mesto razlike v zaporedju (35).

Prva generacija sekvenciranja, to je sekvenciranje po Sangerju, je še danes zlati standard za določanje nukleotidnega zaporedja (34). Postopek metode se začne s PCR pomnožitvijo odseka genomske DNK, ki ga želimo sekvencirati. Nato izvedemo sintezo fragmentov, pri čemer uporabimo pomnožen odsek DNK, polimerazo, en oligonukleotidni začetnik, deoksinukleotide (dNTP) in štiri različno fluorescentno označene dideoksinukleotide (ddNTP). DdNTP v nasprotju z dNTP na 3' koncu deoksiriboze nimajo hidroksilne skupi-

ne, zato se z njihovo vezavo izgrajevanje verige DNK konča. DNTP in ddNTP tekmujejo za vgrajevanje v novonastalo verigo DNK. Vsakič ko se vgradi ddNTP, se sinteza verige zaustavi, zato nastanejo različno dolgi fragmenti. Ti se nato s pomočjo kapilarne elektroforeze ločijo po velikosti. Fragmenti potujejo po kapilari in na koncu dosežejo detektor. Tam barvilo na 3' koncu ddNTP osvetli laser, pri čemer pride do oddajanja fluorescence. Sledi tolmačenje rezultatov na podlagi valovne dolžine in intenzivnosti fluorescence. Pristop omogoča detekcijo znanih in še neopisanih sprememb (38).

V zadnjih letih je glavna osnovna metoda izbire za določanje nukleotidnega zaporedja sekvenciranje naslednje generacije. Druga generacija sekvenciranja v primerjavi s sekvenciranjem po Sangerju omogoča masovno vzporedno sekvenciranje večjih genomske regije (34). Različni proizvajalci uporabljajo različne izvedbe NGS. Vsem je skupno, da DNK najprej fragmentiramo in pripravimo naključne, prekrivajoče se odseke podobnih dolžin. Sledi vezava adapterjev na oba konca fragmentov, kar omogoči kasnejšo hibridizacijo odsekov na trdno površino. Na ta način pripravimo knjižnico. Preverimo kakovost in kvantiteto knjižnice. Nato fragmente v knjižnici klonalno pomnožimo (emulzijski PCR ali premostitveni PCR). Na koncu poteče sekvenciranje vsakega klona (sekvenciranje s sintezo ali sekvenciranje z ligacijo). Sekvence analiziramo z bioinformatičnimi orodji in preverimo prisotnost sekvenčnih sprememb. Tako lahko določimo nukleotidno zaporedje vseh kodirajočih regij DNK – eksoma (angl. *whole exome sequencing*, WES) ali celotne sekvence DNA – genoma (angl. *whole genome sequencing*, WGS) (39). Za NGS je v primerjavi z dideoksi metodo značilna večja hitrost pridobivanja podatkov in nižja cena analize na nukleotid. Pomanjkljivosti tehnike pa sta krajša dolžina sekvenc, ki jih hkrati sekvenciramo (za zagotavljanje čim večje kakovosti prebranih odčitkov), in večja stopnja napak (manjša točnost) (34,40). Vseeno omogoča zelo zanesljivo detekcijo SNV ter malih insercij in delecij (40).

Omejitve pri določanju humanih genetskih sprememb s sekvenciranjem naslednje generacije

Med ključnimi parametri platform NGS, ki vplivajo na zanesljivost pridobljenih rezultatov, so dolžina in natančnost odčitkov ter globina sekvenciranja (41). Med sekvenciranjem lahko nastajajo naključne napake, ki jih je v nekaterih primerih težko prepoznati. Ta problem v praksi rešujemo s povečanjem globine sekvenciranja, torej povečanjem končnega števila sekvenciranih odčitkov. Z večjim številom odčitkov je večja tudi zanesljivost prepoznavanja prave baze na posameznem odčitku, saj napaka sekvenciranja v tem primeru postane statistično nepomembna. Žal pa je večje število odčitkov povezano tudi z višjimi stroški sekvenciranja. Izraz pokritost se nanaša na globino sekvenciranja izbranih tarčnih področij (42). Trenutno žal v klinični praksi ni jasnega priporočila glede minimalne pokritosti. Ta je namreč odvisna od meje detekcije, ki jo želimo doseči, in je na primer pri določanju pridobljenih sprememb izredno visoka, tolerance glede lažno pozitivnih oziroma negativnih rezultatov in seveda napake izbranega instrumentalnega pristopa.

Določanje strukturnih sprememb

NGS je precej uspešna metoda za določanje SNV in manjših sekvenčnih sprememb. Največji problem so strukturne spremembe, ki jih s kratkimi dolžinami odčitkov, značilnimi za metode NGS, ne uspemo identificirati (7). Določanje večjih insercij, delecij ali duplikacij z NGS lahko izboljšamo tako s povečanjem globine sekvenciranja kot tudi z uporabo obojestranskih odčitkov, ki nam dajo informacijo o približni razdalji med dvema parnima odčitkom. Na ta način lahko sklepamo, da oba odčitka prihajata iz istega izhodiščnega fragmenta in tako identificiramo določene strukturne spremembe

(41). Vseeno pa kljub vsem izboljšavam tehnologije NGS veliko strukturnih sprememb zgrešimo. Pomembno je tudi dejstvo, da je določanje genetskih sprememb z NGS posredno, saj je odvisno od primerjave preiskovančevega zaporedja DNK z zaporedjem referenčnega genoma (trenutno GRCh38). Tako o resničnem zaporedju strukturne spremembe sklepamo, ne da bi jo točno določili (43,44).

Pogosto kratkih dolžin odčitkov ne moremo točno prilegati določenim delom humanega genoma. Primer so srednje velike strukturne spremembe (manjše od 2000 bp), inverzije, ponovitvene sekvence, GC- ali AT-bogata področja ter različni deli genoma, kjer se nahajajo zaporedja s homolognimi elementi (npr. CNV). V teh primerih se odčitki prilegajo na več lokacij v genomu, kar povzroči napake pri identifikaciji določenih genetskih sprememb. Krajši kot so odčitki, večji del genoma je podvržen netočnemu (dvoumnemu) prileganju (14,43,45).

Pomanjkljivosti referenčnih genomov

Po sekvenciranju prvega humanega genoma (osnova za sedanjí referenčni genom – GRCh38) in kasneje naslednjih genomov so raziskovalci z začudenjem ugotovili, da med njimi obstajajo velike razlike. Genoma dveh posameznikov se med seboj ne razlikujeta le v milijonih SNV, temveč tudi v več deset tisočih večjih strukturnih spremembah (večjih od 50 bp) (7). Ta informacija je zelo pomembna, saj genetiki uporabljajo referenčni humani genom za primerjavo s preučevano sekvenco DNK in identifikacijo genetskih sprememb, ki so povezane z različnimi boleznimi. Zaradi razprostrtosti genetskih sprememb po celotnem genomu lahko predvidevamo, da ima vsak posamezni genom, kot je tudi prvi referenčni genom, genetske spremembe, ki so ali tudi niso prisotne pri večini ljudi (46,47). Tako ni nujno, da je referenčni genom ustrezen standard (47).

Prav tako vemo, da se humane genetske spremembe razlikujejo med etničnimi skupinami. Zato se raziskovalci v zadnjem času trudijo, da bi sistematično analizirali različne humane genome in sestavili referenčne genome različnih človeških populacij (48). Različni referenčni genomi bodo omogočili večjo občutljivost odkrivanja humanih genetskih sprememb (47).

Pomen 3. generacije sekvenciranja; tehnologije dolgih odčitkov

Tehnologije dolgih odčitkov (angl. *long-read sequencing*) so nadgradnja tehnologij NGS. Omogočajo direktno sekvenciranje zelo dolgih sekvenc DNK (več kot 10.000 bp). Dodatna prednost teh metod je možnost sekvenciranja ene same molekule DNK, zaradi česar predhodno klonalno pomnoževanje ni potrebno, globina sekvenciranja oz. pokritost pa je zato bolj enakomerna (43). V raziskavah so ugotovili, da lahko zaradi naštetih prednosti lažje prepoznamo oz. identificiramo strukturne spremembe, predvsem tiste, ki so dolge 50–2000 bp (7). S temi metodami zaznamo skoraj 2,5-krat več strukturnih sprememb kot z metodami NGS. Z uporabo predhodnih tehnik bi rutinsko lahko spregledali vsaj 48 % delecij in 83 % insercij (49). S tehnologijami dolgih odčitkov pa so raziskovalci uspeli prepoznati tudi kompleksne regije, npr. VNTR, segmentne duplikacije (večje od 1000 bp) in sekvenco centromer (7). Največja pomanjkljivost teh metod je dejstvo, da so relativno nove, generirajo napake in imajo dokaj nizko zmogljivost (nizka hitrost) s posledično visoko ceno. Zaradi vsega naštetega jih za zdaj ne uporabljamo v kliničnem diagnosticiranju. Se pa tehnologije 3. generacije pospešeno razvijajo in obetajo napredek v zaznavanju strukturnih sprememb, tudi v kliničnem okolju (43).

Obeti za prihodnost

Nezanemarljiv delež humanih genetskih sprememb, povezanih z razvojem genetskih bolezni, še ni odkrit oz. določen. Čeprav so se z razvojem tehnologije v zadnjem desetletju naše možnosti za njihovo identifikacijo izredno povečale, še vedno obstajajo omejitve in možnosti za izboljšanje pristopov.

Zato moramo natančneje opredeliti večje število humanih referenčnih genomov čim več različnih človeških populacij. Tehnologije dolgih odčitkov (skupaj s tehnologijami kratkih odčitkov, ki odpravljajo napake) omogočajo izdelavo velikega števila referenčnih genomov. Ocenjujejo, da bi v primeru poznavanja sekvence približno 300 različnih referenčnih genomov podvojili število znanih strukturnih sprememb in odkrili večino pogostih strukturnih sprememb (s frekvenco manj pogostega alela, večjo od 1 %). Tako bi poznavanje strukture teh sprememb olajšalo njihovo določanje v genomih, ki so bili pridobljeni s tehnologijami kratkih odčitkov (47,50).

Številne patogene genetske različice se nahajajo v nekodirajočih regijah genoma. Za določanje teh genetskih sprememb je ključno sekvenciranje celotnega genoma (WGS), ki ga danes v klinični praksi uporabljamo redko. Širšo implementacijo WGS v klinično prakso namreč upočasnjujejo visoki stroški izvedbe, nizka pokritost sekvenciranja in velika količina ustvarjenih podatkov, ki so problematični za obdelavo (51,52), kar naj bi se v bližnji prihodnosti spremenilo.

Poleg WGS je pomembna tudi določitev faznih genomov. To pomeni, da se genetske spremembe dodeli bodisi materinim ali očetovim kromosomom. Oba starševska haplotipa namreč sekvenciramo, s čimer dobimo bolj popolno informacijo o genomu posameznika. S takim načinom lahko ugotavljamo vzorce dedovanja, alelnospecifično izražanje (raven izražanja posameznega alela pri heterozigotih), analiziramo sestavljene heterozigote in identificiramo patogene genet-

ske različice. S faznimi genomi, ki so sekvencirani s tehnologijami dolgih odčitkov, lahko odkrijemo 30 % več strukturnih sprememb kot z uporabo enakih tehnologij, ki ne določajo faznih genomov (14,43). Danes je neposredno faziranje genoma še nekoliko oteženo, saj je s tehnologijami kratkih odčitkov, ki so dostopne na trgu, težko sestaviti sekvence posameznih homolognih kromosomov.

Genetske spremembe lahko identificiramo tudi s pomočjo sestavljanja genoma vsakega posameznika *de novo*. V prihodnosti naj bi bil to vodilni način identifikacije genetskih sprememb v klinični praksi. Z naprednimi/novimi tehnologijami dolgih odčitkov bi najprej sekvencirali oba haplotipa preiskovanca (očetovega in materinega) in tako sestavili fazni genom. Nato bi oba haplotipa primerjali z drugimi referenčnimi genomi. To bi bilo še posebej primerno za bolezni, ki se razvijajo v starosti, saj v teh primerih starševska DNK navadno ni dostopna. Tako bi lahko pospešili odkrivanje osebnih oz. družinskih patogenih genetskih sprememb (7,53).

Zaključek

Prepoznavanje genetskih sprememb, ki so vzrok genetskih bolezni ali vplivajo na klinično odločanje pri posameznem bolniku, postaja vsakdanja klinična praksa. To posebej velja na področju opredelitve prirojenih bolezni v otroški populaciji, saj omogoča hitro in dokončno opredelitev vzroka bolezni, kar je v večini ključno za učinkovito in pravočasno klinično ukrepanje. Pričakujemo, da bo napredek na področju genomike vodil v nekoliko drugačen pristop pri kliničnem prepoznavanju genetskih bolezni. Ko bo cena tehnologij dolgih odčitkov (3. generacije sekvenciranja) dovolj dostopna in ko bodo bolj zmožljive, hitrejšje in zanesljive, bomo lahko sekvencirali, fazirali in sestavljali genom vsakega posameznega preiskovanca. To je osnova za učinkovito odkrivanje posamezniko-

vih genetskih sprememb in identifikacijo morebitnih patoloških sprememb. V Sloveniji je uvajanje tehnologije sekvenciranja naslednje generacije potekalo hitro po njeni uveljavitvi in je tako na voljo že od leta 2014. Izkušnje pri tem pa bodo seveda izrednega pomena za nadgradnjo obstoječih diagnostičnih pristopov, in sicer s sekvenciranjem dolgih odčitkov, ki ga v Sloveniji raziskovalno na področju humane genetike že uporabljamo.

Literatura

1. Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, et al. Mapping and sequencing of structural variation from eight human genomes. *Nature*. 2008; 453(7191): 56–64. Available from: <https://pubmed.ncbi.nlm.nih.gov/18451855>.
2. Serre D, Pääbo S. Evidence for gradients of human genetic diversity within and among continents. *Genome Res*. Cold Spring Harbor Laboratory Press; 2004; 14(9): 1679–85. Available from: <https://pubmed.ncbi.nlm.nih.gov/15342553>.
3. Rahim NG, Harismendy O, Topol EJ, Frazer KA. Genetic determinants of phenotypic diversity in humans. *Genome Biol*. BioMed Central; 2008 24; 9(4): 215. Available from: <https://pubmed.ncbi.nlm.nih.gov/18439327>.
4. Griffiths AJF, Wessler SR, Carroll SB, Doebley J. *An Introduction to Genetic Analysis*. 11th ed. New York: W. H. Freeman and Company; 2015.
5. Davis CP. Types and List of Examples of Genetic (Hereditary) Diseases [Internet]. eMedicineHealth. [cited 2020 Apr 23]. Available from: https://www.emedicinehealth.com/types_and_list_of_genetic_diseases/article_em.htm
6. Cui Y, Li G, Li S, Wu R. Designs for Linkage Analysis and Association Studies of Complex Diseases BT - *Statistical Methods in Molecular Biology*. In: Bang H, Zhou XK, van Epps HL, Mazumdar M, editors. Totowa, NJ: Humana Press; 2010. p. 219–42. Available from: https://doi.org/10.1007/978-1-60761-580-4_6
7. Eichler EE. Genetic Variation, Comparative Genomics, and the Diagnosis of Disease. *N Engl J Med* 2019; 381(1): 64–74. Available from: <https://pubmed.ncbi.nlm.nih.gov/31269367>
8. Patton KT, Thibodeau GA. *Genetics and Heredity: The Human Genome*. Anat Physiol. 8th ed. London: Elsevier Health Sciences; 2014. p. 1133–5.
9. Shabalina SA, Spiridonov NA. The mammalian transcriptome and the function of non-coding DNA sequences. *Genome Biol*. 2004; 5(4): 105. Available from: <https://doi.org/10.1186/gb-2004-5-4-105>.
10. Jackson M, Marks L, May GHW, Wilson JB. The genetic basis of disease. *Essays Biochem*. Portland Press Ltd.; 2018 Dec 2; 62(5): 643–723. Available from: <https://pubmed.ncbi.nlm.nih.gov/30509934>.
11. Bhattacharyya N, Khanra K. Mutation - Types of point mutations. *A Text B. Mol Biol*. First edition. New Delhi: EduPedia Publications; 2016. p. 218–22.
12. Help Me Understand Genetics (Genomic Research): What are single nucleotide polymorphisms (SNPs)? *Natl Libr Med*. (US); Genet. Home Ref. 2020 [cited 2020 May 25]. Available from: <https://ghr.nlm.nih.gov/primer/genomicresearch/snp>.
13. Strachan T, Read A. *Human Molecular Genetics*. 5th ed. London: Garland Science; 2018.
14. Huddleston J, Chaisson MJ, Steinberg KM, Warren W, Hoekzema K, Gordon D, et al. Discovery and genotyping of structural variation from long-read haploid genome sequence data. *Genome Res*. 2016/11/28. Cold Spring Harbor Laboratory Press; 2017; 27(5): 677–85. Available from: <https://pubmed.ncbi.nlm.nih.gov/27895111>.
15. Hu J, Ng PC. Predicting the effects of frame-shifting indels. *Genome Biol*. [Internet]. BioMed Central; 2012; 13(2): R9–R9. Available from: <https://pubmed.ncbi.nlm.nih.gov/22322200>.
16. Stankiewicz P, Lupski JR. Structural Variation in the Human Genome and its Role in Disease. *Annu Rev Med*. Annual Reviews; 2010; 61(1): 437–55. Available from: <https://doi.org/10.1146/annurev-med-100708-204735>.
17. Mlinaric-Rascan I, Karas Kuzelicki N, Ostanek B, Jagodic Mlinaric K. *Farmakogenomika*. Ljubljana: Fakulteta za farmacijo; 2010.
18. Ivics Z, Izsvák Z. Repetitive elements and genome instability. *Semin Cancer Biol*. 2010; 20(4): 197–9. Available from: <http://www.sciencedirect.com/science/article/pii/S1044579X10000623>.
19. Duitama J, Zablotskaya A, Gemayel R, Jansen A, Belet S, Vermeesch JR, et al. Large-scale analysis of tandem repeat variability in the human genome. *Nucleic Acids Res*. 2014/03/20. Oxford University Press; 2014; 42(9): 5728–41. Available from: <https://pubmed.ncbi.nlm.nih.gov/24682812>.
20. Kolarski M, Ahmetovic B, Beres M, Topic R, Nikic V, Kavecian I, et al. Genetic Counseling and Prenatal Diagnosis of Triploidy During the Second Trimester of Pregnancy. *Med Arch*. (Sarajevo, Bosnia Herzegovina). AVICENA, d.o.o., Sarajevo; 2017; 71(2): 144–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/28790549>.
21. Miller OJ, Therman E. Abnormal Phenotypes Due to Autosomal Aneuploidy or Polyploidy. *Hum. Chromosom*. New York, NY: Springer New York; 2001. p. 175–85. Available from: https://doi.org/10.1007/978-1-4613-0139-4_12.
22. O'Connor C. Chromosomal Abnormalities: Aneuploidies. *Nat Educ*. 2008; 1(1): 172. Available from: <https://www.nature.com/scitable/topicpage/chromosomal-abnormalities-aneuploidies-290/>.
23. Help Me Understand Genetics (Mutations and Health): How can gene mutations affect health and development? *Natl Libr Med* (US); Genet Home Ref 2020 [cited 2020 Apr 25]. Available from: <https://ghr.nlm.nih.gov/primer/mutationsanddisorders/mutationscausedisease>.
24. Stenson PD, Ball E V, Mort M, Phillips AD, Shiel JA, Thomas NST, et al. *Human Gene Mutation Database (HGMD®): 2003 update*. Hum Mutat. John Wiley & Sons, Ltd; 2003 Jun 1; 21(6): 577–81. Available from: <https://doi.org/10.1002/humu.10212>.

25. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. *Nature*. 2015; 526(7571): 75–81. Available from: <https://pubmed.ncbi.nlm.nih.gov/26432246>.
26. Chiang C, Scott AJ, Davis JR, Tsang EK, Li X, Kim Y, et al. The impact of structural variation on human gene expression. *Nat Genet* 2017; 49(5): 692–9. Available from: <https://doi.org/10.1038/ng.3834>.
27. Hurler ME, Dermitzakis ET, Tyler-Smith C. The functional impact of structural variation in humans. *Trends Genet* 2008; 24(5): 238–45. Available from: <https://pubmed.ncbi.nlm.nih.gov/18378036>.
28. Geršak K. Genetsko svetovanje in prenatalna diagnostika. In: Takač I, Geršak K, editors. *Ginekolo. Perinatol.* 1. izd. Maribor: Medicinska fakulteta; 2016. p. 447–54.
29. Mahdih N, Rabbani B. An overview of mutation detection methods in genetic disorders. *Iran J Pediatr. Tehran University of Medical Sciences* 2013; 23(4): 375–88. Available from: <https://pubmed.ncbi.nlm.nih.gov/24427490>
30. Lovrecic L, Peterlin B. Implementation of molecular karyotyping in clinical genetics. *Zdr Vestn.* 2013; 82(10): 669–76. Available from: <https://vestnik.sz.d.si/index.php/ZdravVest/article/view/960>
31. Schouten JP, McElgunn CJ, Waaijer R, Zwijsenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002 15; 30(12): e57–e57. Available from: <https://doi.org/10.1093/nar/gnf056>
32. Jenko Bizjan B, Katsila T, Tesovnik T, Šket R, Debeljak M, Matsoukas MT, et al. Challenges in identifying large germline structural variants for clinical use by long read sequencing. *Comput Struct Biotechnol J* 2020; 18: 83–92. Available from: <http://www.sciencedirect.com/science/article/pii/S2001037019303678>
33. Cerne D, Ostanek B, Lukac-Bajalo J, Marc J, Hvala I. *Biomedicinska analitika I.* Ljubljana: Fakulteta za farmacijo; 2012.
34. Katsanis SH, Katsanis N. Molecular genetic testing and the future of clinical genomics. *Nat Rev Genet* 2013 Jun; 14(6): 415–26. Available from: <https://pubmed.ncbi.nlm.nih.gov/23681062>
35. Giani AM, Gallo GR, Gianfranceschi L, Formenti G. Long walk to genomics: History and current approaches to genome sequencing and assembly. *Comput Struct Biotechnol J* 2020; 18: 9–19. Available from: <http://www.sciencedirect.com/science/article/pii/S2001037019303277>
36. Eggerding FA. Oligonucleotide Ligation Assay. In: Walker JM, Rapley R, editors. *Med Biomethods Handb Totowa, NJ: Humana Press*; 2005. p. 293–303. Available from: <https://doi.org/10.1385/1-59259-870-6:293>
37. Tucker EJ, Huynh BL. Genotyping by High-Resolution Melting Analysis BT - Crop Breeding: Methods and Protocols. In: Fleury D, Whitford R, editors. New York, NY: Springer New York; 2014. p. 59–66. Available from: https://doi.org/10.1007/978-1-4939-0446-4_5
38. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl. Acad. Sci. U. S. A.* 1977; 74(12): 5463–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/271968>
39. Knief C. Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Front Plant Sci* 2014. p. 216. Available from: <https://www.frontiersin.org/article/10.3389/fpls.2014.00216>
40. Basturea GN. *MATER METHODS: Somatic Mutations* [Internet]. Labome.com. 2018 [cited 2020 Apr 29]. p. 8:2673. Available from: <https://www.labome.com/method/Somatic-Mutations.html>
41. Progar V, Petrovič U. Vpliv parametrov sekvenciranja naslednje generacije na zanesljivost rezultatov v metagenomskih študijah. *Inform. Medica Slov* 2013; 18(1/2): 1–8.
42. Sims D, Sudbery I, Illott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet* 2014; 15(2): 121–32. Available from: <https://doi.org/10.1038/nrg3642>
43. Mantere T, Kersten S, Hoischen A. Long-Read Sequencing Emerging in Medical Genetics. *Front Genet* 2019. p. 426. Available from: <https://www.frontiersin.org/article/10.3389/fgene.2019.00426>
44. Guan P, Sung W-K. Structural variation detection using next-generation sequencing data: A comparative technical review. *Methods* [Internet]. 2016;102:36–49. Available from: <http://www.sciencedirect.com/science/article/pii/S1046202316300184>
45. Chaisson MJP, Huddleston J, Dennis MY, Sudmant PH, Malig M, Hormozdiari F, et al. Resolving the complexity of the human genome using single-molecule sequencing. *Nature*. 2015 29; 517(7536): 608–11. Available from: <https://pubmed.ncbi.nlm.nih.gov/25383537>
46. Pan B, Kusko R, Xiao W, Zheng Y, Liu Z, Xiao C, et al. Similarities and differences between variants called with human reference genome HG19 or HG38. *BMC Bioinformatics* 2019; 20(2): 101. Available from: <https://doi.org/10.1186/s12859-019-2620-0>
47. Ballouz S, Dobin A, Gillis JA. Is it time to change the reference genome? *Genome Biol.* [Internet]. 2019;20(1):159. Available from: <https://doi.org/10.1186/s13059-019-1774-4>
48. Li R, Li Y, Zheng H, Luo R, Zhu H, Li Q, et al. Building the sequence map of the human pan-genome. *Nat Biotechnol* 2010; 28(1): 57–63. Available from: <https://doi.org/10.1038/nbt.1596>
49. Chaisson MJP, Sanders AD, Zhao X, Malhotra A, Porubsky D, Rausch T, et al. Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nat. Commun* 2019; 10(1): 1784. Available from: <https://doi.org/10.1038/s41467-018-08148-z>
50. Audano PA, Sulovari A, Graves-Lindsay TA, Cantsilieris S, Sorensen M, Welch AE, et al. Characterizing the Major Structural Variant Alleles of the Human Genome. *Cell* 2019 24; 176(3): 663–675.e19. Available from: <https://pubmed.ncbi.nlm.nih.gov/30661756>
51. Pereira MA. Application of Next-Generation Sequencing in the Era of Precision Medicine. In: Malta FSV, editor. *Rijeka: IntechOpen*; 2017. p. Ch. 13. Available from: <https://doi.org/10.5772/intechopen.69337>
52. Gloss BS, Dinger ME. Realizing the significance of noncoding functionality in clinical genomics. *Exp Mol Med. Nature Publishing Group UK*; 2018 ;50(8): 97. Available from: <https://pubmed.ncbi.nlm.nih.gov/30089779>
53. Tian S, Yan H, Klee EW, Kalmbach M, Slager SL. Comparative analysis of de novo assemblers for variation discovery in personal genomes. *Brief Bioinform. Oxford University Press* 2018. 28; 19(5): 893–904. Available from: <https://pubmed.ncbi.nlm.nih.gov/28407084>

Lara Slavec, mag. lab. biomed.

Univerzitetni klinični center Ljubljana, Ginekološka klinika, Ljubljana, Slovenija in Univerza v Ljubljani, Fakulteta za farmacijo, Katedra za klinično biokemijo, Ljubljana, Slovenija.

prof. dr. Ksenija Geršak, dr. med., svet.

Univerzitetni klinični center Ljubljana, Ginekološka klinika, Ljubljana, Slovenija in Univerza v Ljubljani, Medicinska fakulteta, Katedra za ginekologijo, Ljubljana, Slovenija.

doc. dr. Nataša Karas Kuželički, mag. farm.

Univerza v Ljubljani, Fakulteta za farmacijo, Katedra za klinično biokemijo, Ljubljana, Slovenija.

izr. prof. dr. Katarina Trebušak Podkrajšek, spec. lab. med. genet., spec. med. biokem. (kontaktna oseba / contact person)

Univerzitetni klinični center Ljubljana, Pediatrična klinika, Inštitut za specialno laboratorijsko diagnostiko, Vrazov trg 1, 1000 Ljubljana, Slovenija in Univerza v Ljubljani, Medicinska fakulteta, Inštitut za biokemijo in molekularno genetiko, Vrazov trg 2, 1000 Ljubljana, Slovenija. tel.: +386 1 543 7669 e-naslov: katarina.trebusakpodkrajsek@mf.uni-lj.si

prispelo / received: 16. 7. 2020
sprejeto / accepted: 2. 10. 2020

Trebušak Podkrajšek K, et al. Humane genetske spremembe in njihovo določanje: trenutno stanje in obeti za prihodnost. *Slov Pediatr* 2020; 27(4): 163–171. <https://doi.org/10.38031/slovpediatr-2020-4-01>.