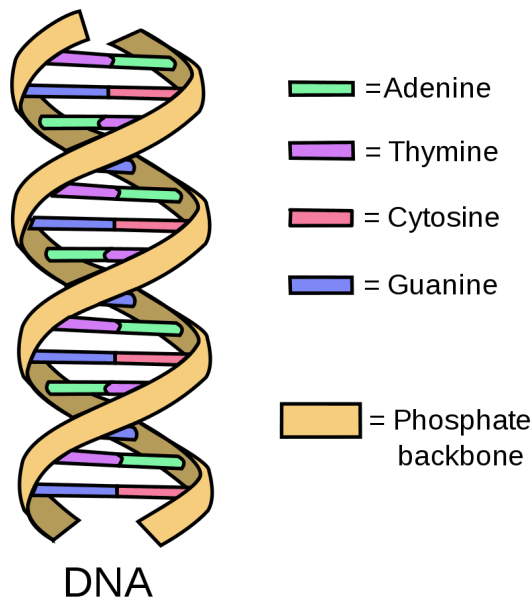


1. Úvod

V tejto kapitole sa zameriame na opis základných biologických pojmov. Popíšeme si DNA a jej vlastnosti, typy mutácií, ktoré sa môžu vyskytnúť v genetickej informácii človeka a spôsoby, akými vieme genetickú informáciu spracovať. Taktiež sa v tejto kapitole zameriame spôsoby odhadovania veku a na stanovenie základných problémov, ktoré pri odhadovaní vznikajú.

1.1. DNA

Deoxyribonukleová kyselina, skrátene DNA, sa nachádza v jadre bunky a tvorí genetickú informáciu človeka. Nesie v sebe genetické inštrukcie pre vývoj, fungovanie, rast a rozmnožovanie. DNA sa skladá z dvoch polynukleotidových reťazcov, ktoré sa otáčajú okolo seba a vytvárajú pravotočivú dvojzávitnicu. Je to polymér, ktorého základnou stavebnou jednotkou je nukleotid [1,2]. Nukleotid sa skladá z jednej zo štyroch dusíkatých báz a to z adenínu (A), guanínu (G), tymínu (T) alebo z cytozínu (C). Taktiež sa skladá z cukru nazývaného deoxyribóza a z fosfátovej skupiny. Nukleotidy sú spojené pomocou kovalentných väzieb, ktoré vznikajú medzi cukrom jedného nukleotidu a fosfátom ďalšieho, čím vytvárajú cukrovo-fosfátovú kostru. Dusíkaté bázy dvoch samostatných polynukleotidových vlákien sa spájajú podľa pravidiel párovania báz. Podľa týchto pravidiel sa adenín spája s tymínom a cytozín s guanínom. Názorný príklad DNA môžeme vidieť na obrázku Obr. 1.

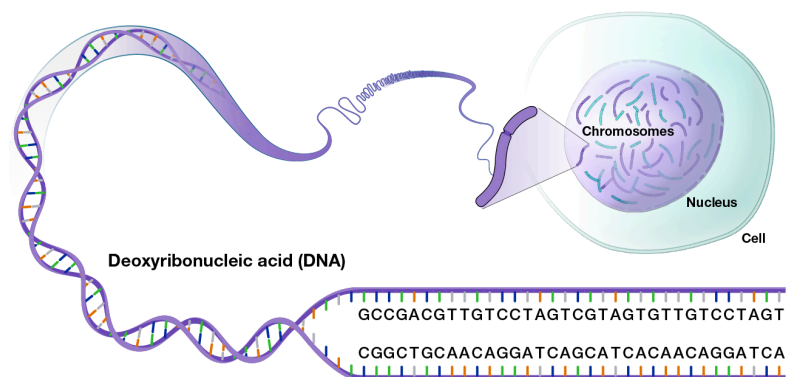


Obr. 1. Zobrazenie deoxyribonukleovej kyseliny a dusíkatých báz. Zdroj: ashg.org

Obe vlákna DNA v sebe uchovávajú biologickú informáciu určenú poradím nukleotidov DNA vlákna sa používajú ako predloha pri vytváraní RNA vlákien. Tento proces, pri ktorom sa pomocou DNA vytvorí RNA, sa nazýva transkripcia. Pri prepise DNA na RNA sa každá dusíkatá báza prepíše na komplementárnu dusíkatú bázu, okrem tymínu, ktorý sa prepíše na uracil (U). Ribonukleová kyselina, skrátene RNA, je na rozdiel od DNA jednovláknová a popisuje postupnosť aminokyselín v proteínoch. Táto vlastnosť RNA sa využíva pri procese nazývanom translácia [3]. U ľudí, DNA sa združuje do dlhých štruktúr, ktoré nazývame chromozómy.

Z informatického hľadiska sa na dusíkové bázy pozeráme ako na prvky abecedy. Tie nám vytvárajú ľubovoľné konečné neprázde reťazce, ktoré nazývame DNA segmenty. Z týchto reťazcov sú zložené konečné neprázde slová nazývané chromozómy. Genóm je kompletná sada DNA inštrukcií, ktorá sa nachádza v každej bunke človeka. Pod pojmom genóm budeme pre účely tejto práce označovať konečný jazyk, ktorý je zložený z týchto neprázdnych konečných slov. Tento jazyk bude

reprezentovať úplnú genetickú informáciu jedinca. Formu tohto jazyka môžeme vidieť na obrázku Obr. 2.



Obr. 2. Zobrazenie formálneho jazyka na ľudskej bunke. Zdroj: genome.gov

1.2. Mutácie a ich varianty

Pod pojmom mutácie označujeme náhle a náhodné zmeny v genetickom materiáli. Tieto zmeny sa prejavujú v zmene primárnej štruktúry nukleovej kyseliny, čiže nastane zmena v sekvencii nukleotidov. Mutácie menia genotyp daného jedinca ale fenotypovo sa nemusia prejaviť. Ich vznik je eliminovaný pôsobením na to určenými enzýmami. Bez nich by nastal nekontrolovaný vznik mutácií a zánik daného jedinca v krátkom čase. Mutácie sú však z evolučného hľadiska veľmi dôležité a tie, ktoré sa preukážu ako prospešné, prenesú sa na ďalšiu generáciu.

Existuje veľa typov, podľa ktorých vieme mutácie rozdeliť. Z hľadiska tejto práce si zdefinujeme tri rozdelenia.

Definícia 1.2.1 Gametické mutácie: postihujú genetický materiál pohlavných buniek, vznikajú počas spermatogenézy, čiže tvorby spermii, alebo oogenézy, čiže tvorby vajčiek.

Definícia 1.2.2 Somatické mutácie: vznikajú počas života jedinca, postihujú telové bunky a neprenášajú sa na ďalšie generácie. Môžu byť príčinou vzniku rakoviny.

Definícia 1.2.3 Zárodočné mutácie: vyskytujú sa v organizme od narodenia a sú dôsledkom gametických mutácií.

Gametické, somatické a zárodočné mutácie sú spontánne mutácie, ktoré vznikajú chybou troch procesov. Môžu vzniknúť chybou replikačného systému, kedy DNA-polymeráza zaradí pri replikácii nesprávny nukleotid, ktorý porušuje komplementaritu báz, rekombinčného systému, kedy delenie pohlavných buniek prebehne nepresne a reparačného systému, kedy enzýmy, ktoré majú na starosti opravu poškodenej DNA urobia chybu. Mutácie v tele daného jedinca môžu taktiež vzniknúť na základe vonkajších, napríklad environmentálnych, vplyvov.

Genóm je kompletná sada DNA inštrukcií, ktorá sa nachádza v tele každej bunky. Jedna kópia genómu obsahuje okolo troch miliárd nukleotidov, ktoré sú rozmiestnené po 23 chromozómoch. Väčšina ľudských buniek sú diploidné bunky, čiže obsahujú dve kópie ľudského genómu. Kópie sa získajú pri spojení dvoch pohlavných buniek. Jedna od matky a jedna od otca.

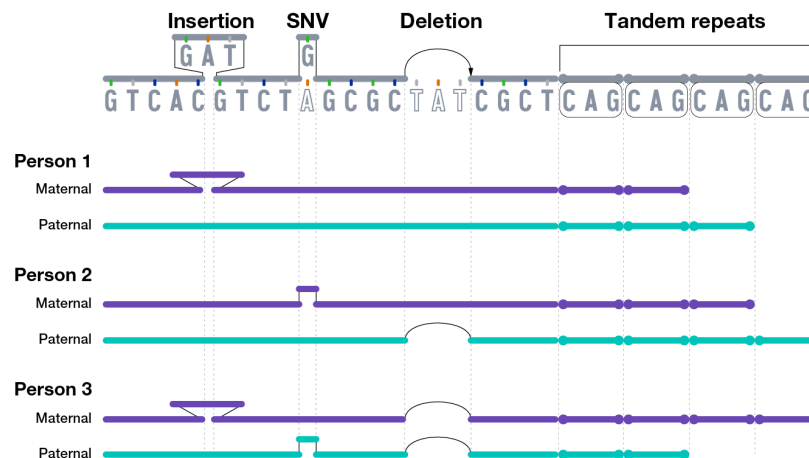
Ľudské genómy sú z väčšej časti identické. Ak si zoberieme dvoch náhodných ľudí z populácie, ich genómy sú približne na 99,9% identické. Toto tvrdenie vychádza zo zistenia, že približne raz za 1300 nukleotidov sa vyskytuje v ľudskom genóme zmena oproti referenčnému genómu. V tomto tvrdení sú však zahrnuté len jednonukleotidové zmeny. V realite, ak zahrnieme do tohto porovnania aj viacnukleotidové zmeny, dvaja náhodne vybraní ľudia z populácie majú genóm na 99,6% rovnaký. Tieto rozdiely v ľudskom genóme nazývame genomické varianty.

1.2.1 Krátke varianty

O krátkych variantoch hovoríme, ak dĺžka tohto variantu je menšia ako 50 nukleotidov. Najmenším a zároveň najbežnejším typom krátkych variantov sú

jednonukleotidové varianty (SNV). SNV nám ukazuje rozdiel v jednom nukleotide. Pri porovnaní dvoch genómov od dvoch náhodných jedincov na tej istej pozícii, môže mať prvý jedinec na danej pozícii tymín (T) a druhý cytozín (C). Podtypom SNV je **jednonukleotidový polymorfizmus (SNP, čítaj snip)**. Aby sme SNV brali ako SNP, daná báza sa musí nachádzať na danej pozícii minimálne u 1% populácie.

Ďalšou skupinou malých genomických variantov je **inzercia** a **delécia**, skrátene **indels**. Pri inzercii sa do genómu pridá malý reťazec nukleotidov. Pri delécii je malý reťazec nukleotidov vynechaný. Tieto varianty sú menej časté, ako SNV ale vedia mať väčší dopad na zdravie jedinca. Jedným z najčastejších typov inzercie a delécie sú **tandemové opakovania**, čo sú krátke reťazce nukleotidov, ktoré sa opakujú niekoľkokrát za sebou a ich výskyt je v bežnej populácii veľmi rozdielny. Počet týchto opakovaných nukleotidových úsekov sa u jednotlivých chromozómov líši. Zatiaľ, čo výskyt týchto tandemových opakovaní vieme u jedného chromozómu počítať v desiatkach, u druhého chromozómu opakovania rátame na stovky. Na nasledujúcom obrázku môžeme vidieť grafické znázornenie krátkych variantov.

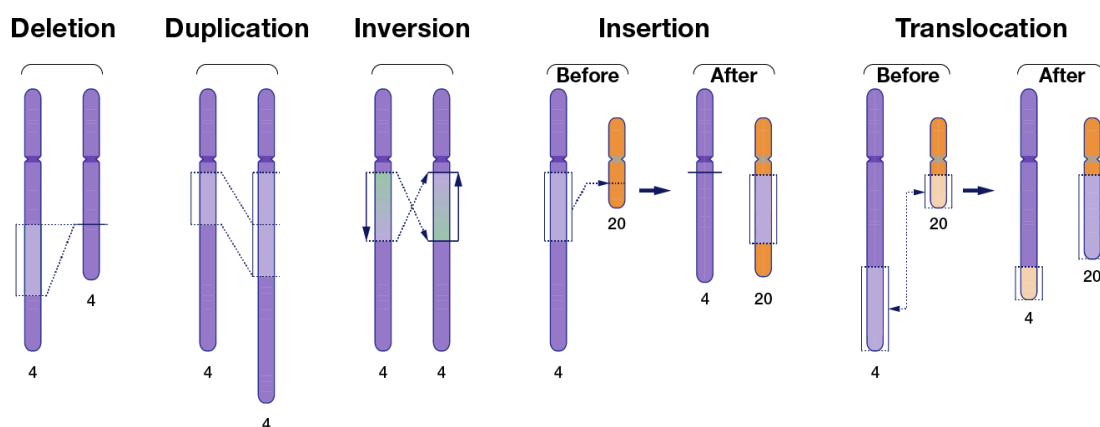


Obr. 4 Grafické znázornenie krátkych variantov na genómch troch ľudí. Zdroj:

genome.gov

1.2.2 Štrukturálne varianty

Štrukturálne varianty sú varianty, ktorých dĺžka sa pohybuje od minimálne od 50, niekedy v tisícoch nukleotidov. Tieto varianty boli do genómu vložené (inzercia), boli z neho vynechané (delécia), invertované alebo presunuté z jednej časti genómu do druhej. Na nasledujúcom obrázku môžeme vidieť grafické znázornenie týchto variantov.



Obr. 5 Grafické znázornenie štrukturálnych variantov. Zdroj: genome.gov

Tandemové opakovania, ktoré obsahujú viac ako 50 nukleotidov nazývame štrukturálnymi variantami. Ak štrukturálne varianty spôsobia rozdielnu celkovú dĺžku genómu oproti referenčnému genómu, nazývame tento stav **variabilita počtu kópií** (CNV). CNV je stav, v ktorom počet opakovaní špecifickej časti DNA segmentu je rôzny medzi jednotlivcami. Počet opakovaní týchto kópií môže byť krátky, ale môže obsahovať aj niekoľko tisíc báz. CNV odlišujeme od ostatných štrukturálnych variantov, ako sú napríklad inverzia a translokácia, lebo pri neskôr spomenutých variantoch nedochádza ku zmene celkovej dĺžky genómu.

1.3. Bioinformatické metódy spracovania ľudského genómu

Na nájdenie genómových variantov sa používajú rôzne metódy. Niektoré z týchto metód sú zamerané len na nájdenie už známych variantov. Patria tu napríklad metódy na odhalenie genetického ochorenia, ktoré má vplyv na zdravie a kvalitu života daného jedinca. Pomocou konkrétneho DNA testu vieme odhaliť, či daný človek má hľadaný

genomický variant. Ďalšími analýzami vieme zistiť prezenciu známych genómových variantov v DNA daného človeka, napríklad pomocou testu microarray vieme zistiť prítomnosť tisícov SNV naraz. Existujú však oveľa lepšie metódy na odhalenie známych aj neznámych genomických variantov. Medzi tieto metódy patrí aj sekvenovanie.