

4. féléves beszámoló

Szakállas Nikolett (szakallasn3@student.elte.hu)

Fizika Doktori Iskola,
Statisztikus Fizika, Biológiai Fizika
és Kvantumrendszerek Fizikája PhD program,
Eötvös Loránd Tudományegyetem

Témavezető: Dr. Szabó Bálint

Eötvös Loránd Tudományegyetem,
Természettudományi Kar
Biológiai Fizika Tanszék

Konzulens: Prof. Molnár Béla

Semmelweis Egyetem,
Belgyógyászati és Onkológiai Klinika

Dolgozat címe: Szöveti eredetű egyedi sejt manipulációs eszköz fejlesztése és biológiai alkalmazásai

1. Bevezetés, visszatekintés az előző félévi tevékenységre

Ahogy egyéb rosszindulatú elváltozások kapcsán ismeretes, úgy a vastagbél daganatok kialakulása is egy több lépcsős folyamat, mely során a normál vastagbél hám- vagy kötőszövetek neoplasztikus változásokon mennek keresztül mutációk felhalmozódása mellett. A mutációs mintázatokba a normál és tumoros vastagbél szöveti eredetű minták teljes genomi és teljes exomi szekvenálásával nyerhetünk betekintést. Az elmúlt években a szekvenálási eljárások alkalmazása egyre elterjedtebb, mely a vonatkozó bioinformatikai módszerek fejlődését eredményezi. Ezek következménye, hogy az egyes daganat típusoknál fellelhető leggyakoribb mutált génekről és a kapcsolódó eltérésekről adatbázisok készülnek. Az adatbázisok felhasználásával detektálhatóak és összehasonlíthatóak adott kísérletben vizsgált minták mutációs tulajdonságai.

Az első éves munkám folytatásaként a harmadik és negyedik félévben is tumor-normál (pontosabban tumor-tumor melletti normál szövet) mintapárok, és kezelésen átesett, illetve kontroll minta adatok szekvenálási eredményeinek kiértékelésével foglalkoztam. Az előkészített minták exomi és teljes genomi szekvenáláson estek át. A különböző kiértékelési szempontok alkalmazása mellett bevezetésre került a szekvenálási adatok pozíció tartományokra szűkített kiértékelése egy .bed kiterjesztésű fájl alapján, mely a vizsgálatok szempontjából kitüntetett régiók koordinátáit tartalmazza. A mi esetünkben a fehérje kódoló exomi régiók adatait foglalta össze. Az eddigi vizsgálatainkat elvégeztük egyedi sejt méretű, szöveti eredetű lézermikrodisszekált mintákra is. Jelen beszámolóban ennek kezdeti eredményei is ismertetve lesznek.

2. Az aktuális félévben elvégzett kutatások ismertetése

Az aktuális kutatási félévben csoportok (CRC, CRC-NAT, CRC-AD, IBD és NEG) közötti összehasonlítást végeztem Nanopore szekvenátorral, illetve a szöveti eredetű lézermikrodisszekált közel egyedi sejt (CRC, CRC-NAT és NEG) minták vizsgálata és kiértékelése is elkezdődött. Az eddigi vizsgálatokhoz képest lényeges különbség, hogy az utóbb ismertetett kísérlet során a vizsgálat tárgyai nem vér-, hanem szöveti-metszeti eredetűek. Következésképpen a mintavétel és a mintaelőkészítés is speciális lépéseket igényelt. A mintavételt magam, míg a szekvenálást és az azt megelőző lépéseket molekuláris biológus kollégák végezték.

2.1. Csoportos összehasonlítás

Közel 120 minta került teljes genomi szekvenálásra, melynek eszköze a Nanopore PromethION24 rendszer volt. A vizsgálat célja, hogy összehasonlítás kapjunk az említett szekvenátor és az Illumina NextSeq 500/550 által nyújtott adatokat tekintve, különös figyelmet fordítva az exomi régiókra. A 120 minta kisebb csoportokra osztható, melyek között CRC, CRC-NAT, AD, IBD és NEG esetek fordultak elő. Ezek nagy része egy korábbi tanulmány során átesett teljes exomi szekvenáláson is, az Illumina NextSeq 500/550 rendszeren. [1] Az összevetés célja, hogy olyan CRC specifikus onkológiai profilt alkothassunk azokat a variánsokat is tartalmazva, melyeket az Illumina nem, de a Nanopore képes detektálni. Emellett szeretnénk összefoglalni, hogy milyen a rendszerek hatékonysága, mennyire fednek át az eredményül kapott adataik. Egyszerűbben szólva: ugyanazt tudja-e az átfedő régióban a két rendszer, vagy sem.

A teljes genomi szekvenálási adatok exomra szűkítéséhez azt a .bed fájlt alkalmaztam, melyet a NextSeq 500/550 rendszer is alapul vesz a szekvenálás során. Ez az ún. szűrés egy plusz lépést

jelent a kiértékelés során, de egyben kisebb variáns adatmennyiséget jelent az összehasonlítások terén (hiszen az exom a genom kódoló része, s annak csupán 1%-át teszi ki). A kiértékelés menete a következők szerint történt Nanopore szekvenálás esetén: *pycoQC* minőségellenőrzés a nyers adatokon, majd *Dorado* alapú bázishívás (a módosult bázisokat is beleértve, pl. 5mC és 5hmC) + Hg38 referencia genomra illesztés. Ezután a *samtools sort/index/merge* - .bam fájlok végzett műveletek végrehajtása, a *modkit* metylációs vizsgálatok, végül a *EPI2ME Labs human variation workflow* variánsazonosítása, a *bcftools annotate* .vcf fájlok annotálása történt. A .bed fájl alapú szűrés volt a következő lépés. A *bcftools stats* statisztikai adatok kinyerése is a teljes genomi adatokra történt. Ezután a detektált variánsok összehasonlítását saját programkód alapján végeztem.

Az eredmények még nincsenek publikálva, a közeljövőben várható közlésük tudományos cikk formájában. Előzetten azt tapasztaltam, hogy a két szekvenátor exomi adatai nem fednek át teljesen. A hosszú leolvasási hosszak eredményeképp a Nanopore módszer számukban és jellegükben is több variánst vesz észre és annotál. Érdemes leszögezni ugyan, hogy az Illumina által talált variációk nagy részét a Nanopore is megtalálja, ámbar vannak olyan egyediek is - még ha kis számban is -, melyek csak az Illumina rövid hosszak esetén jelent meg. A tanulmány idővel kiterjeszthető (beteg)csoport specifikus eredmények közzétételére is.

2.2. Egyedi sejt minták

Lézermikrodisszektált szöveti eredetű egyedi sejt mintákon is sikeresen elvégzésre került egy teljes exomi szekvenálás az Illumina NextSeq500/550 rendszerén. CRC, CRC-NAT és NEG vastagbél metszetekről kerültek 30 és 40 μm átmérőjű, kör alakú darabok kivágásra. A vágás előtt és után a lemezek szkennelve lettek, mely fontos a vizsgálatokat követő lokális eltéréseket listázó vizualizációs lépés szempontjából.

12-12 különböző területről, 1 vagy 2 darab kört vágtam ki, majd kollégák segítségével az amplifikációs lépéseket követően megtörtént maga a DNS szekvenálás. A kiértékelést minőségellenőrző lépések közbeiktatásával a Illumina BaseSpace felületen fellelhető Dragen (germline és somatic pipeline) csomagok használatával végeztem. Ezután az eredmények vizsgálata és értelmezése általam írt és már meglévő, szabad felhasználású programokkal történt.

A mintákban fellelhető összes variáns mellett az ATGC adatbázis CRC specifikus top 60 génmutációját, illetve közel 15, a nem hipermutált tumorokra jellemző pl. (*APC*, *KRAS*, *TP53*, *PIK3CA*, *SOX9*) géneken tapasztalt elváltozásokat kerestem. Ezeket sikeresen, elegendően nagy lefedettséggel detektáltam. Ezáltal egyfajta adatbázis született egészséges és vastagbél-daganatos páciensek mintáit tekintve, a variánsokra specifikusan. A következő lépésekben meghatározhatóak a tumorspecifikus mutációk a NEG adatok CRC és CRC-NAT-ból való kiszűrésével. Annak következtében, hogy a kísérlet a közelmúltban zajlott, az eredmények még nem teljesek, de publikálásra kerülnek majd, amint minden releváns összehasonlítás megtörtént.

2.3. Egyedi sejt izolációs rendszer fejlesztése

Az ismertetett időszakban inkább a mintavételt követő biológiai vizsgálatokra, azok bioinformatikai kiértékelésére koncentráltam. A korábbi beszámolómban említett és röviden ismertetett rendszer alapjai azonban átgondolásra kerültek. Ez alapján új kísérleti elrendezés került kialakításra. Az elvégzett mérések alapján azt tapasztaltuk, hogy az egyedi sejtek manipulációjára alkalmas lehet a rendszer, az általam ismertetett egyedi sejt szekvenálás pedig a rendszer relevanciáját alátámasztó sikeres kísérletnek tudható be.

3. Tanulmányi tevékenység az aktuális félévben

A félév során az alábbi kurzust végeztem el:

- Az érzékelés biofizikája II.: Bioakusztika (FIZ/3/045E)

4. Publikációk az aktuális félévben

2024 januárjában benyújtottam egy összefoglaló cikket a Pathology and Oncology Research lap-hoz, mely 2024 májusában jelent meg *Szakállas, N; Barták, BK; Valcz, G; Nagy, ZsB; Takács, I; Molnár, B. Can long-read sequencing tackle the barriers, which the next-generation could not? A review. Pathol. Oncol. Res., 16 May 2024. doi: 10.3389/pore.2024.1611676.* címmel.

5. Konferenciák az aktuális félévben

Májusban részt vettem az Oxford Nanopore Technologies által rendezett Nanopore London Calling 2024 konferencián; júniusban pedig *A legújabb szekvenálási eljárások diagnosztikai alkalmazásai - bioinformatikai megközelítés* című előadást tartottam a Magyar Gasztroenterológus Társaság Nagygyűlésén.

Hivatkozások

[1] A. Kalmar et. al.

Patterns of Somatic Variants in Colorectal Adenoma and Carcinoma Tissue and Matched Plasma Samples from the Hungarian Oncogenome Program
Cancers 2023, 15(3), 907;
doi: 10.3390/cancers15030907