

3. 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 daganatos elváltozások kapcsán is 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 használata gyorsan terjed, mely magával vonja a bioinformatikai módszerek fejlődését is. Ezek következménye, hogy az egyes daganat típusoknál fellelhető leggyakoribb mutált génekről és a kapcsolódó eltérésekről adatbázisok készülnek. Az adatbázisok felhasználásával detektálhatóak és összehasonlíthatóak a kísérletben részt vett minták mutációs tulajdonságai. Az első éves munkám folytatásaként a harmadik félévben is tumor-normál (pontosabban tumor-tumor melletti normál szövet) mintapárok és kezelésen átesett és 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 fájl alapján, mely a vizsgálatok szempontjából kitüntetett régiók koordinátáit tartalmazza. A mi esetünkben a fehérje kódoló exomi régiók adatait foglalta össze.

2 Az aktuális félévben elvégzett kutatások ismertetése

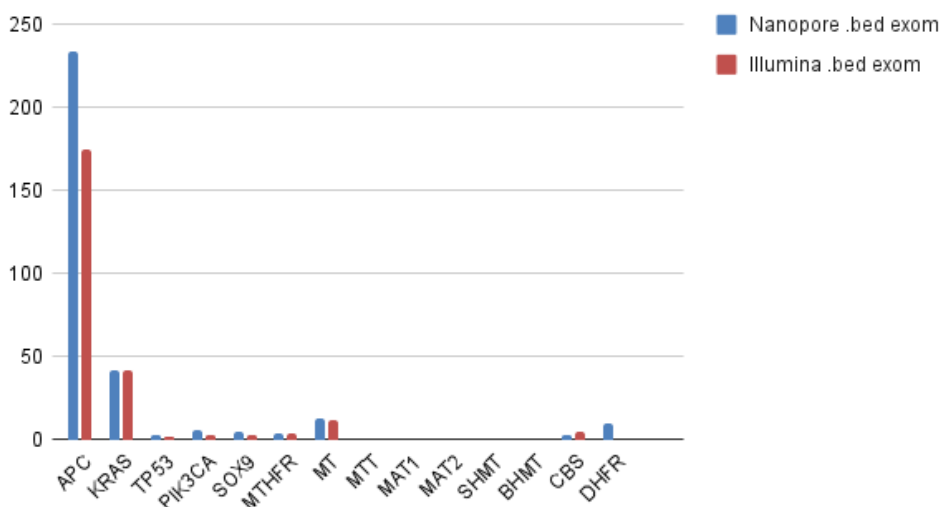
2.1 Minták szekvenálása, bioinformatikai kiértékelése

2.1.1 Mintacsoport exomi és genomi szekvenálása, az adatok összehasonlítása

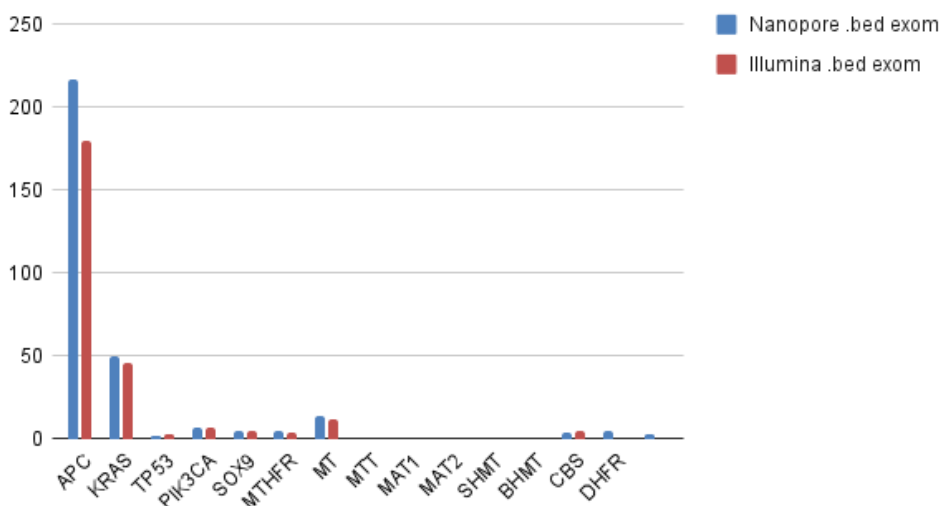
A kiértékelt adatok teljes exomi és teljes genomi szekvenáláson estek át. Előbbi az Illumina NextSeq 500/550, utóbbi a Nanopore PromethION24 szekvenátor használatával. A teljes genomi szekvenálás során egy adott organizmus (esetünkben humán) mintájában a teljes genomi (azaz teljes genetikai információ), míg teljes exom esetén az exomi (azaz a genomban lévő fehérje-kódoló régiók) bázisrendjét határozzuk meg. A szekvenált minták vastagbél tumoros metszetekből származnak, két típusú területről: tumorosból és tumor melletti normálból 5 makrodisszekált mintapár (összesen 10 minta). A vizsgálat célja az volt, hogy összehasonlítsuk a két különböző szekvenálási eljárás adatait az exomi területekre koncentrálva. Főként a bázisarányokra és a mutációs profilokra koncentráltunk.

A kiértékelés menete az előző beszámolóknak ismertett lépések szerint történt. Röviden összefoglalva a következőket végeztem el; rendre Nanopore szekvenálás esetén: *pycoQC* minőségellenőrzés, *Guppy* bázishívás + referencia genomra illesztés, *samtools sort/index/merge* - .bam fájlokon végzett műveletek végrehajtása, *modbam2bed* - metilációs vizsgálatok, *EPI2ME Labs human variation workflow* - variánsazonosítás, *bcftools annotate* - .vcf fájlok annotálása, *bcftools stats* - statisztikai adatok egy részének kinyerése; Illumina szekvenálásnál: BaseSpace felületen fellelhető Dragen (germline és somatic pipeline) csomagok + minőségellenőrzés. Az adatok elemzését általam írt és már meglévő, szabad felhasználású programok segítségével végeztem. A specifikusabb elemzéshez szükség volt a Python (és szükség esetén R) nyelv használatára. Ide tartoznak pl. az adatbázisból kinyert génmutációk jelenlétének detektálása a variánsok között első körben a kromoszómán elfoglalt pozíciók alapján, majd mélyebbre ásva az egyes (bázis)alterációkat és a klinikai relevanciákat figyelembe véve.

Mutációk száma a NAT mintában



Mutációk száma a tumor mintában



Mivel a genomi területek az exomokhoz képest lényegesen kiterjedtebbek, a hozzájuk tartozó adatokat szűkítenem kellett (annak érdekében, hogy ugyanazon területeken vizsgálódjunk). Ilyen esetekben alkalmazhatóak a .bed fájlok, melyekben oszlopokra bontva megadhatjuk, hogy melyik kromoszóma, melyik pozícióját szeretnénk megtartani. A .bed fájl alapján történő szűrést a referenciagenomra illesztés után hajtottam végre mindkét szekvenátor adatain, így minden statisztikai adat ugyanazon régiókra lett meghatározva a két módszer esetén. A kapott eredményeket összehasonlítottam figyelembe véve vastagbélrák tumorokra jellemző gének (*APC*, *KRAS*, *TP53*, *PIK3CA*, *SOX9*), illetve a folsavciklus néhány specifikus génjén (*MTHFR*, *MT*, *MTT*, *MAT1*, *MAT2*, *SHMT*, *BHMT*, *CBS*, *DHFR*) tapasztalt mutációit. Az ábrán látható, hogy a Nanopore szekvenátor magasabb gyakorisággal tudta detektálni az említett mutációkat a legtöbb gén esetén. A részletes eredményeket publikációban szeretnénk közzétenni.

2.1.2 Különböző mérettartományú tumor-normál mintapárok szekvenálása - kiegészítés az előző félévhez képest

- Mikrodisszekált egyedi sejtek szekvenálása: 12 tumor-normál mintapár (összesen 24 minta)gm került összegyűjtésre, a minták lézermikrodisszekcióval lettek eltávolítva friss fagyasztott mintát tartalmazó lemezről. Ebben az esetben Illumina szekvenátorral nyertük ki az exomi DNS szekvenciákat. Az adatok kiértékelése egy külső felületen történt, melynek neve BaseSpace és kifejezetten az Illumina szekvenátorok adatainak kiértékeléséhez készült. Itt történt meg a bázisazonosítástól a variánsazonosításig az összes lépés, majd az azonosított variánsokat (annotálás után) összevettem a referencia adatbázisból kinyert leggyakoribb génmutációkkal. A kapott eredmények jóságát egy korábbi kutatás eredményeire ([1]) alapozva ellenőrizzük, azonban ez már nem készült el az előző beszámoló leadási határidejéig. A kiértékelés megtörtént, mely alapján a következőket diszkutáltuk. Ugyan a gyűjtött minták a metszet különböző területeiről kerültek összegyűjtésre, mind a tumoros, mind a NAT típusúakban, nagyfokú heterogenitást tapasztaltunk. Emellett láttuk, hogy már a tumor melletti normál területeken is megjelentek egyes tumorspecifikus mutációk. Ez a munka poszter formájában bemutatásra került a Magyar Belgyógyász Társaság 2023-as Nagygyűlésén.

2.2 Egyedi sejt izolációs rendszer fejlesztése

Az ismertett időszakban inkább a mintavételt követő biológiai vizsgálatokra, azok bioinformatikai kiértékelésére koncentráltam. A korábbi beszámolókból említett és röviden ismertett 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, azonban az igazoló kísérletek elvégzése még hátravan.

3 Tanulmányi tevékenység az aktuális félévben

A félév során az alábbi kurzust végeztem el:

- A gépi tanulás új eredményei szeminárium (FIZ/3/092)

4 Publikációk az aktuális félévben

2024 januárjában benyújtottam egy összefoglaló cikket "Can long-read sequencing tackle the barriers, which the next-generation could not? A review." címmel a Pathology and Oncology Research nevű újságba. Jelenleg elbírálás alatt áll.

5 Konferenciák az aktuális félévben

Az aktuális félévben posztert mutattam be a Magyar Belgyógyász Társaság 2023-as Nagygyűlésén "Vastagbél-daganatok tumor heterogenitás vizsgálata lézer mikrodisszekált sejtek exom-szekvenálási adatainak bioinformatikai kiértékelésével" címmel.

6 Elnyert ösztöndíjak az aktuális félévben

- Kooperatív Doktori Program 2023: C2270480 azonosítószámú pályázat, melynek címe: "Szöveti eredetű egyedi sejt manipulációs eszköz fejlesztése és daganat-biológiai alkalmazása".

References

- [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