

Az extrém hosszú élettartam genetikai vizsgálata teljes genom szekvencia alapján kutyákban

37. OTDK Biológia Szekció

Készítette:

Tfirst Dóra

Eötvös Loránd Tudományegyetem

Biológus mesterképzés, első év

Témavezető:

Dr. Jónás Dávid

Tartalomjegyzék

Rövidítések listája	3
Ábrajegyzék.....	4
Táblázatjegyzék.....	5
Összefoglaló	6
Bevezetés.....	7
Célkitűzések	9
Módszerek	9
Eredmények.....	13
Diszkusszió.....	18
Korlátok.....	21
Köszönetnyilvánítás	22
Irodalom	23

Rövidítések listája

bp: bázispár

DNS: dezoxiribonukleinsav

Gbp: gigabázispár

GWAS: teljes genom asszociációs vizsgálat

indel: inzerció és delécio variáns

PCR: polimeráz láncreakció

RNS: ribonukleinsav

SNP: egyetlen nukleotidot érintő polimorfizmus

WGS: teljes genom szekvenálás

Ábrajegyzék

1. ábra: A hatóanyagok, gyógyszerek kifejlesztésének vázlata a bioinformatikai elemzésektől indulva: teljes genom asszociációs vizsgálat (GWAS), teljes exom szekvenálás (WES) és teljes genom szekvenálás (WGS). Zhang és mtsai. (2020) nyomán.	8
2. ábra: Kördiagramm a tíz extrém idős kutyában jelenlévő SNP-k annotációjának megoszlásáról (McLaren és mtsai, 2016).	16
3. ábra: Kördiagramm a tíz kutyában jelenlévő indel variánsok hatásairól (McLaren és mtsai, 2016).....	16
4. ábra: Az összes kutyában jelenlévő SNP variánsok által érintett géekre lefuttatott génontológiai elemzés néhány eredménye. A narancssárga színek a celluláris komponens, a világos kék szín a molekuláris funkció, a zöld szín a biológiai folyamatok elemzés eredményét jelöli.	17

Táblázatjegyzék

1. táblázat: A kutatásba bevont extrém idős kutyák adatai.	10
2. táblázat: A leolvasások számainak, átlagos hosszainak és illesztésének adatai.	13
3. táblázat: Autoszómákon, X és Y kromoszómákon és scaffoldokon talált SNP- k és indelek száma	14
4. táblázat: Az extrém idős kutyák variánsai a szűrés után. Az új variánsok azon variánsokat jelenti, melyeket korábban nem publikáltak.	15

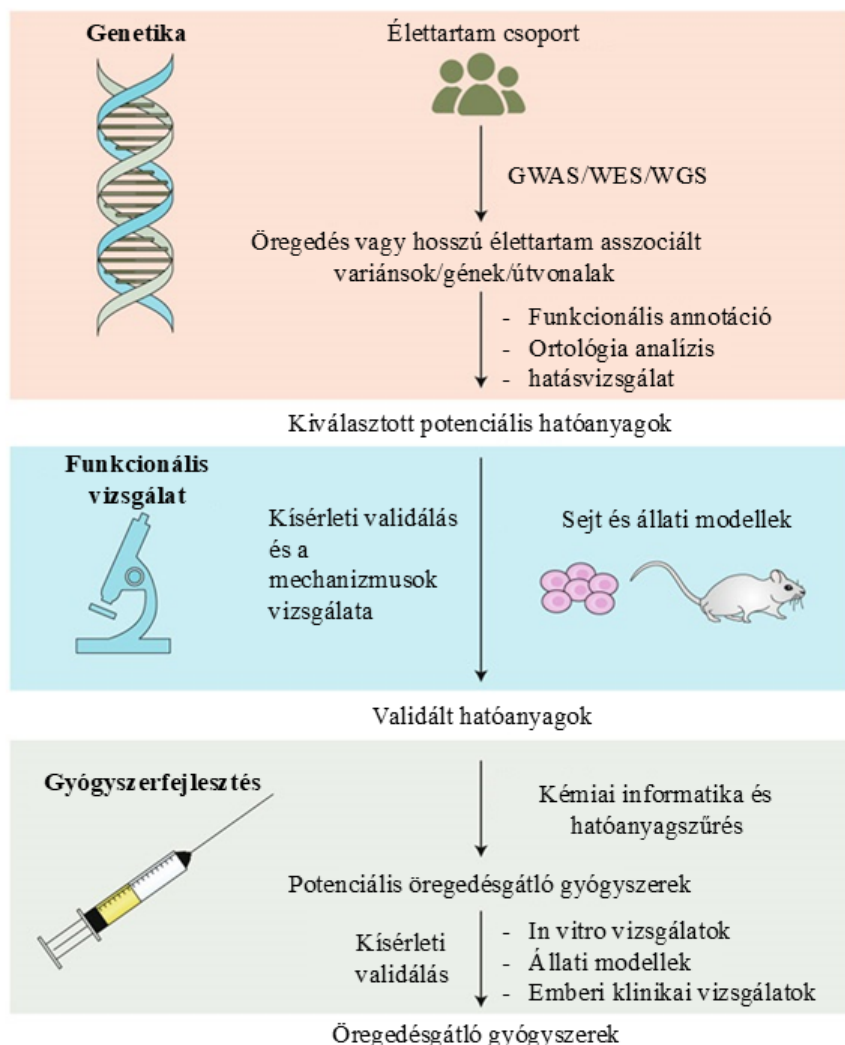
Összefoglaló

Idős korban megnő egyes betegségek kockázata, melyek jelentősen rontják az egyén életminőségét. A világunkban megfigyelhető társadalmi előregedés miatt egyre nagyobb hangsúlyt kap az időskori betegségek megelőzése, gyógyítása és az egészséges öregedés elősegítése. Az öregedést környezeti és genetikai faktorok egyaránt befolyásolják, így szükséges az öregedés genetikai hátterének tanulmányozása. Az extrém idős egyedek feltehetően több olyan genetikai faktorral rendelkeznek, melyek lehetővé teszik számukra a kivételesen magas életkor elérését és védelmet nyújtanak a betegségekkel szemben. A kutatásban tizenegy extrém idős (19-31 éves), krónikus betegséggel nem diagnosztizált kutya (*Canis familiaris*) teljes genom szekvenciájának bioinformatikai elemzését végeztük el, hogy feltárjuk az egészséges öregedéssel kapcsolatba hozható SNP és inzerció-delécio (indel) variánsokat. Összesen 11.185.341 SNP és 5.921.131 indel mutációt találtunk, melyekből 744.445 SNP és 491.104 indel minden extrém idős kutyában jelen volt, valamint 2.747.639 SNP és 1.558.625 indel a vizsgált kutyák legalább 70%-ában megtalálható volt. Az átlagos korú kutyákban is jelenlévő variánsok kiszűrése után 129.167 SNP és 51.431 indel maradt minden extrém idős kutyában, valamint 388.918 SNP és 120.665 indel legalább 70%-ukban. A variánsok többsége introni vagy intragénikus régiókban helyezkedett el, a vártak megfelelően. A nem kódoló régiókat érintő variánsok nem változtatják meg közvetlenül a fehérjék aminosav sorrendjét. Az exoni szekvenciákat érintő variánsokat tekintve az indel mutációk az esetek 84%-ában frameshift mutációt okoztak, illetve a vizsgált SNP-k eredményeztek missense mutációt, míg az indelek nem. Meghatároztuk a variánsok által érintett gének számát, valamint a gének génontológiai elemzését is elvégeztük: erős feldúsulást találtunk a DNS-javítás pozitív szabályozásában, a sejtvázzal, a vezikuláris transzporttal és az idegrendszerrel kapcsolatos celluláris komponensekben és folyamatokban. Az idegrendszert érintő génontológiák feldúsulása kapcsolatban lehet a neurodegeneratív betegségek kialakulásának megelőzésével. Továbbá a génontológiai elemzésben alulreprezentált immunológiai folyamatok alátámasztják, hogy az egészséges öregedésben fontos szerepet játszanak az immunrendszeri mechanizmusok, melyek kulcsfontosságúak a betegségek hatékony megelőzésében és leküzdésében.

Bevezetés

Az öregedés egy fiziológiai, funkcionális hanyatlással járó komplex biológiai folyamat (Wang és mtsai, 2022; McKenzie és mtsai, 2022; Ruple és mtsai, 2022), mely jelentős kockázati tényezője számos betegségnek, mint például neurodegeneratív betegségeknek (pl. Alzheimer-kór és Parkinson-kór), szív- és érrendszeri betegségeknek, krónikus gyulladásos betegségeknek (pl. rheumatoid arthritis), ízületi- és mozgásszervi betegségeknek (pl. osteoarthritis és szarkopénia), a krónikus obstruktív tüdőbetegségnek (COPD) és a ráknak egyaránt (Franceschi és mtsai, 2018; Partridge és mtsai, 2018). Az egész világunkra jellemző előregedés következményeként megnövekedett az időskori betegségek előfordulása, így a társadalmi- és egészségügyi kihívások leküzdése céljából különösen fontos, hogy az idős emberek életminőségét javítsuk, élettartamuk egészséges időszakát meghosszabbítsuk az egészséges öregedésük elősegítésével, támogatásával (Goldman és mtsai, 2013). Az egészséges öregedés magába foglalja a fizikai és mentális betegségek hiányát, a normális napi aktivitást, mobilitást, kognitív funkciókat, szociális viselkedést, az általános jólét megőrzését (Ruple és mtsai, 2022).

Az öregedést számos faktort, köztük környezeti, viselkedési és genetikai egyaránt befolyásolja. Az extrém idős kort megélt egyedek feltehetően több olyan faktorral rendelkeznek, melyek védenek az időskori betegségekkel szemben (Erikson és mtsai, 2016). Több tanulmány is beszámolt már modellállatokban sikeresen elért egészséges élettartam meghosszabbításról. Az öregedés lassítását megcélzó ideális terápiák egyik jellemzője, hogy késői életszakaszban alkalmazva is kifejtik pozitív hatásukat. Egereken végzett rapamycin kezelés mind fiatal, mind középkorban történő alkalmazása megnövelte az egerek élethosszát és javította az egészségi állapotukat (Bitto és mtsai, 2016; Shindyapina és mtsai, 2022; Herrera és mtsai, 2023). *Daphnia magna* planktonikus rákok élettartamát is sikerült meghosszabbítani rapamycin kezeléssel, mely a hatás mechanizmus feltételezett evolúciós konzerváltságának elméletét erősíti (Shindyapina és mtsai, 2022). A rapamycin mellett a metformin is széles körben kutatott hatóanyag (Anisimov és mtsai, 2011; Chen és mtsai, 2021). A megfelelő potenciális hatóanyagok kiválasztásához szükség van bioinformatikai – például teljes genom asszociációs vizsgálat (GWAS) vagy teljes genom szekvenálás (WGS) – elemzésekből nyert adatokra: öregedéssel kapcsolatba hozható variánsokra, génekre, útvonalakra (*l.kép*), melyeket transzlációs modellállatokból is azonosíthatunk, mivel bizonyos öregedési folyamatok evolúciósan konzerváltak.



1. ábra: A hatóanyagok, gyógyszerek kifejlesztésének vázlatja a bioinformatikai elemzésektől indulva: teljes genom asszociációs vizsgálat (GWAS), teljes exom szekvenálás (WES) és teljes genom szekvenálás (WGS). Zhang és mtsai. (2020) nyomán.

A házi kutya (*Canis familiaris*) egy népszerű transzlációs modelljévé vált az emberi öregedésnek. Szemben a laborban nevelt és tartott állatokkal, a kutya közvetlen kapcsolatban van az emberrel, életmódjukat erősen befolyásolhatja gazdájuk életmódja, környezete, anyagi helyzete (Lee és mtsai, 2022; McCoy és mtsai, 2023). A kor előrehaladtával a kutyában spontán kialakulhatnak humán betegségekkel analóg betegségek, mint például ízületi- és mozgásszervi betegségek, neurodegenerációs betegségek – például a humán Alzheimer kórhoz hasonló kutyakognitív diszfunkció (*canine cognitive dysfunction, CCD*) –, szívelégtelenség, szarkopénia vagy a rák (Taormina és mtsai, 2019; Ruple és mtsai, 2022 Demirtas és mtsai, 2023). Így egyes humán betegségek és az öregedés egyaránt tanulmányozható kutyában. Ellenben figyelembe kell venni, hogy a kutyák élethosszát több tényező befolyásolja. Például

a kistestű kutyáknak hosszabb az élettartama a nagytestű társaikhoz képest (Urfer és mtsai, 2020) és a keverék kutyák tovább élnek a fajtiszta kutyákhoz képest (Yordy és mtsai, 2020). A kutyák genomja 2,4 milliárd bázispár (bp) hosszú, körülbelül 77%-a a humán genom méretének (3,1 milliárd bp). A kutya- és a humángenom citozin-guanin aránya hasonló, illetve a kutya majdnem minden fehérjekódoló génjének van humán homológja (Lindblad-Toh és mtsai, 2005; Taormina és mtsai, 2019).

Kutatásom középpontjában az egészséges öregedés – azon belül is az extrém élettartam – mint poligénes jelleg (Erikson és mtsai, 2016) genetikai hátterének feltárása áll. DNS variánsokat detektálva azonosíthatóak meghatározott fenotípusokat – mint az öregedés – kialakító, elősegítő gének és alléljaik. A következőkben bemutatott kutatói munkámban az egészséges öregedéssel összefüggő SNP (egyetlen nukleotidot érintő polimorfizmus) és indel (inzerció és delécia) variánsokat kerestem extrém idős kutyákban, majd azonosítottam az általuk érintett géneket. A detektált variánsok átlagos életkorú kutyák variánsaival is összehasonlításra kerültek (Plassais és mtsai, 2019).

Célkitűzések

A kutatás célja olyan SNP és indel variánsok azonosítása, melyek kizárólag extrém idős kutyákban találhatóak meg, továbbá a mutációk által érintett gének azonosítása, azok biológiai és molekuláris funkciójának meghatározása.

Módszerek

A kutatásba bevont kutyák

A kutatásban tizenegy extrém idős (≥ 19 éves) kutya teljes genom szekvenciájának bioinformatikai elemzését végeztem el. A DNS mintákat az állatok szájüregének nyálkahártyájáról vettük. A kutyák alapadatai az *1. táblázatban* olvashatóak, melyből látszik, hogy a vizsgált mintában volt fajtiszta és keverék kutya, szuka és kan, illetve kis- és nagyméretű kutya egyaránt.

A környezeti hatásokat illetően az összes extrém idős kutya étrendjében egyszerre volt megtalálható a nyers hús, kereskedelembe kapható, feldolgozott élelmiszerek és a kutyával megosztott, eredetileg emberi fogyasztás céljából elkészített főtt ételek. A kutyák többsége

szabadon tartva, vidéki környéken élt, ahol számos lehetőség nyílt a mozgásra, kóborlásra, továbbá a helybeli és idegen emberekkel és állatokkal interakcióba lépni.

1. táblázat: A kutatásba bevont extrém idős kutyák adatai.

Azonosítók	Név	Életkor (év)	Fajta	Ivar	Testméret
<i>old_dog1</i>	Buksi	27	Keverék	Kan	Közepes
<i>old_dog2</i>	Kedves	22	Keverék	Szuka	Közepes
<i>old_dog3</i>	Morzsi	19	Tacskó	Szuka	Kicsi
<i>old_dog4</i>	Razi	25	Keverék	Szuka	Közepes
<i>old_dog5</i>	Bobi	31	Alentejo masztiff	Kan	Közepes
<i>old_dog6</i>	Lujzi	20	Havannai pincs	Szuka	Kicsi
<i>old_dog7</i>	Frakk	21	Rövidszőrű magyar vizsla	Kan	Közepes
<i>old_dog8</i>	Bumi	22	Keverék	Kan	Közepes
<i>old_dog9</i>	Pici	21	Keverék	Szuka	Kicsi
<i>old_dog10</i>	Augie	21	Golden retriever	Szuka	Nagy
<i>old_dog11</i>	Róbert	21	Törpe pinscher	Kan	Kicsi

Teljes genom szekvenálás

A DNS mintagyűjtés 10 minta esetében citológiai mintavevővel történt az állatok szájüregének nyálkahártyájáról, egy minta esetében (*old_dog8*) állatorvos által levett vérből. A DNS izolálását, a könyvtárkészítést (könyvtárkészítő kit: KAPA HyperPlus kit) és a minták szekvenálását az Omega Biosciences (Norcross, GA, Amerikai Egyesült Államok) végezte. A páros végű (paired-end) szekvenálás egy Illumina HiSeq X Ten gépen történt, melyről mintánként ~840 millió pár rövid szekvenciát vártunk. A szekvenciák hossza 150 bp. A 11 mintából kettőt (*old_dog1*, *old_dog2*) 2017-ban egy korábbi kutatáshoz (Jónás és mtsai, 2020) szekvenáltak, míg a fennmaradó kilencet (*old_dog3-11*) 2023-ban.

Referencia genom

Az elemzésben használt referencia genom, a *ROS_Cfam_1.0* (Harrison és mtsai, 2024) és a hozzá tartozó annotáció az Ensembl adatbázisból (Harrison és mtsai, 2024) lett letöltve. A szekvencia egy négy éves labrador retriever kutya genomjából lett rekonstruálva, tartalmazza mind a 38 autoszóma és mindkét ivari kromoszóma (X és Y) szekvenciáit, valamint 376 db *scaffoldot*, mely szekvenciákat még nem sikerült elhelyezni a kromoszómákon.

Teljes genom szekvencia elemzés

Az elemzésem során igyekeztem a használt programok javasolt alapbeállításait alkalmazni, így ezeket a beállításokat az egyes szoftverek említésekor nem emelem ki. Amennyiben ezektől el kellett térnem, azt rövid indoklással jelzem.

Az elemzés a nyers adatok ellenőrzésével kezdődött, melyet a FastQC szoftver (Bittencourt, 2010) alkalmazásával végeztem el, a leolvasásokon esetlegesen előforduló adapter szekvenciák a Cutadapt szoftver (Martin, 2011) segítségével kerültek eltávolításra.

A leolvasások illesztését a referencia genomra a BWA szoftver BWA-MEM algoritmusával (Li és Durbin, 2009) végeztem el. Az „-R” opcióval rendeltem leolvasási csoportokat (*read group*) a fastq fájlokhoz és az „-M” opciót beállítva az output fájl kompatibilissé vált a következő lépésekben használt Picard szoftver (<http://broadinstitute.github.io/picard/>) számára. Az illesztések kromoszómák szerinti szétválogatását a Samtools szoftver (Danecek és mtsai, 2021) *sort* eszközével, a polimeráz láncreakció (PCR) során esetlegesen keletkező műtermék duplikátumainak megjelölését pedig a Picard *MarkDuplicates* eszközével végeztem el.

Az illesztés minőségéről a Picard *CollectWgsMetrics* és a Samtools *flagstat* eszközével állítottam elő összefoglalót. A variánsok helyes detektálásához szükséges megfelelő minőségi pontszámok (Phred pontszámok) újraszámítása a GATK (Van der Auwera és O'Connor, 2020) *BaseRecalibrator* és *ApplyBQSR* eszközökkel történt.

A variánsok hívását a GATK szoftver *HaplotypeCaller* eszközével végeztem. A variánsokból a GATK *SelectVariants* eszközzel válogattam ki a variánsokat, majd a GATK *VariantFiltration* eszközt alkalmazva a következő, ajánlott kritériumértékeket használtam. SNP-re: QD < 2,0; FS > 60,0; SOR > 3,0; MQ < 40,0; MQRankSum < -12,5; ReadPosRankSum < -8,0. Indelekre az alábbi értékeket használtam: QD < 2,0; FS > 200,0; ReadPosRankSum < -20,0; SOR > 10,0. A kiszűrt variánsokat a GATK *SelectVariants* parancs *--exclude-non-variants* és *--exclude-filtered* beállításával távolítottam el.

Az illesztett leolvasásokat és a variánsokat véletlenszerűen kiválasztott genomszakaszokkal vizuálisan is leellenőriztem az IGV genom böngészőben (Robinson és mtsai, 2023).

Az idős kutyák eredményeit 638, feltehetően átlagos korú kutyák adataival (Plassais és mtsai, 2019) hasonlítottam össze. Az összes publikált variáns közül az álpozitív hívásokat a vcftools (Danecek és mtsai, 2011) *--remove-filtered-all* opciójával zártam ki, megtartva kizárólag a „PASS” flaggel rendelkező variánsokat. Az átlagos korú kutyák közös VCF fájlját szétszedve a

megtartott 638 kutyára, a hiányzó genotípusokkal (./.) vagy a homozigóta referencia alléllal (0/0) rendelkező variánsokat a bcftools *view* eszköz -i 'GT[*]='alt''' parancsával távolítottam el a VCF fájlból. Az átlagos korú kutyák variánsai a CanFam3.1. referencia genomon (Hoeppner és mtsai, 2014) lettek meghatározva, így szükséges volt átemelnünk ezeket a variánsokat az általunk is használt, jelenleg hivatalos kutya referencia genomra (ROS_Cfam_1.0). Az átemelést a bcftools *+liftover* eszközzel (Genovese és mtsai, 2024) végeztük el, melyhez szükség volt egy chain fájlra, melyet a témavezetőm készített el.

Az extrém idős kutyák SNP variánsai esetében a bedtools *multiIntersectBed* (Quinlan, 2014) eszközzel határoztam meg, hogy a detektált variánsok összesen hány kutyában vannak jelen. A fejléc megjelenítéséhez a „-header” opció került beállításra. Az eredményekből kiszűrtük azokat a variánsokat, melyek mindegyik kutyában jelen voltak, illetve azokat, melyek legalább hét kutyában. Az átlagos korú kutyákban detektált variánsok ezután kerültek kiszűrésre a bedtools (Quinlan, 2014) *subtractBed* eszközét alkalmazva. A génontológiai elemzéshez szükséges a génlistát a bedtools *intersectBed* (Quinlan, 2014) programot lefuttatva készítettük el, melyben a következő paraméterek kerültek beállításra: *-a* referenciagenom génlistája, *-b* detektált variánsok és *-wa*.

Az indelek detektálására más folyamatot alkalmaztunk, mint az SNP-k esetében, mivel a multiintersectBed feldarabolná az indeleket kisebb darabokra, attól függően, hogy az egyes alszekvenciái az indel mutációknak hány egyedben voltak megtalálhatóak. Először a variáns hívás eredményeként kapott vcf fájlok alapján készítettünk egy konszenzus indel listát a bedtools (Quinlan, 2014) *mergeBed* programját használva. Ezt követően *intersectBed* (Quinlan, 2014) eszközzel minden kutyára egyénileg kiszűrtük azokat a konszenzusos indel variánsokat, melyek melyekkel legalább 30%-ban átfedtek az adott kutyában detektált indelek legalább egyike. A paraméterek a következőképp kerültek megadásra: *-a* konszenzus indelek, *-b* kutyák variánsai, *-wa* és a szűréshez *-F* 0.30. Ezután a multiintersectBed-et futtattam le az output fájlokra, hogy meghatározzam egy indel variáns hány kutyában van jelen. Ezt követően az SNP mutációkhoz hasonlóan szintén kiszűrtem azokat a variánsokat, melyek minden kutyában megvoltak vagy legalább a kutyák 70%-ában. Az átlagos korú kutyákban detektált variánsokat a bedtools *subtractBed* programjával szűrtük ki az extrém idős kutyákban detektált variánsok listájából.

Az adatfeldolgozások után kapott variánsok genomi annotációját a *Variant effect predictor* (McLaren és mtsai, 2016) online programmal nyertem ki. A program inputjához meghatároztam

azokat a megmaradt variánsokat, melyek hosszának 30%-a átfed az eredeti indel variánsok egyikével. Ezt a lépést az *intersectBed* programmal (*-a* eredeti indel lista, *-b* megmaradt indelek, *-wa*, *-F* 0.30) végeztem el.

A mutációk által érintett gének biológiai, molekuláris funkciójára és sejten belüli lokalizációjára génontológiai elemzést a *PANTHER* (Thomas és mtsai, 2003) online programmal futtattam le. Az *intersectBed* eszközzel (*-a* referencia genom génjei, *-b* kutyákban detektált indelek, *-wao*, *-F* 0.30) összegyűjtöttem azokat a géneket, melyeket érintenek a variánsok.

Eredmények

Kutyák adatainak minősége

A legtöbb leolvasás az *old_dog1* (946,5 millió) és *old_dog2* (962,7 millió) egyedektől származik, míg az *old_dog8* kutyától volt a legkevesebb elérhető adat, ettől az egyedről mindössze 414,6 millió leolvasással rendelkezünk. Az *old_dog1* és *old_dog2* egyedek leolvasásai 150 bp hosszúak voltak, adapter szennyeződés nélkül, így ez a leolvasási hossz nem változott az illesztés előtt. Az eredmények alapján a 2023-ban megszekvenált minták (*old_dog3-11*) adapter szennyezettek voltak, így adapter levágást kellett alkalmazni rajtuk. Az adapter levágáson átesett leolvasások átlagos hossza 137,53 bp lett. Az *old_dog2*, *old_dog3* leolvasásai kerültek a legnagyobb arányban illesztésre (99,72% és 99,81%). Minden kutyánál az olvasások legalább 75%-a illesztésre került, kivéve az *old_dog8* egyednél, ahol ez az arány csak 45,7% volt (2. táblázat).

2. táblázat: A leolvasások számainak, átlagos hosszainak és illesztésének adatai.

Azonosító	Elsődleges leolvasások száma		Átlagos leolvasás hossz	Illesztett leolvasások száma	Illesztett leolvasások (%)
	Adapter levágás előtt	Adapter levágás után			
<i>old_dog1</i>	946 527 082	946 527 082	150	743 359 785	78,03
<i>old_dog2</i>	962 650 390	962 650 390	150	967 687 196	99,72
<i>old_dog3</i>	866 862 144	840 123 212	139,2	849 336 233	99,81
<i>old_dog4</i>	876 444 760	842 334 164	139,2	776 005 602	90,73
<i>old_dog5</i>	886 816 384	861 684 510	136,9	778 166 774	89,45
<i>old_dog6</i>	856 710 548	827 734 618	138,5	653 285 797	78,16
<i>old_dog7</i>	853 259 970	826 650 622	138,8	737 875 737	88,34
<i>old_dog8</i>	432 049 566	414 569 754	139,5	189 860 339	45,73

<i>old_dog9</i>	861 008 764	834 798 642	134,5	686 246 369	81,61
<i>old_dog10</i>	870 093 542	842 564 406	136,3	638 109 087	75,38
<i>old_dog11</i>	880 952 894	851 941 168	134,9	802 945 913	93,15
Átlag	870 093 542	793 600 121,78	139,80	711 170 802,91	84

A genom átlagos lefedettsége az *old_dog1* és az *old_dog2* kutyákban volt a legnagyobb (34,1x és 45,5x), míg az *old_dog8* egyedben a legkisebb (0,66x). A kiemelkedően alacsony genom lefedettsége és illesztés utáni adatminősége miatt az *old_dog8* egyedet kizártam az elemzésből.

Azonosított variánsok az extrém idős kutyákban

Összesen 11.185.341 SNP és 5.921.131 indel mutációt találtunk, melyekből 744.445 SNP és 469.864 indel minden extrém idős kutyában (10/10) jelen volt, valamint 2.747.639 SNP és 1.307.459 indel a vizsgált kutyák legalább 70%-ában (7/10) megtalálható volt. A várakozásoknak megfelelően az autoszómákon volt a legtöbb detektált mutáció, mivel ezek teszik ki a kutya genomjának ~94,5%-át (3. táblázat). A legtöbb SNP-t és indelt az első kromoszómán, a kutyák legnagyobb kromoszómáján találtam. A legkevesebb SNP-t a 36-os kromoszómán, a legkevesebb indelt az Y kromoszómán azonosítottam. A legtöbb variánst az *old_dog4*, egy szuka, 25 éves keverék kutyában azonosítottam ($n_{\text{SNP}}=4.944.516$, $n_{\text{indel}}=2.127.622$).

3. táblázat: Autoszómákon, X és Y kromoszómákon és scaffoldokon talált SNP-k és indelek száma

Kromoszómák	SNP (db)	Indel (db)
Autoszómák	17.872.325	9.432.613
X kromoszóma	315.272	194.875
Y kromoszóma	15.951	5.315
Scaffold	183.692	63.838

Átlagos korú kutyák

Összesen 20.269.614 SNP variánst emeltünk át a ROS_Cfam_1.0 referencia genomra, 337.643 variánst nem tudtunk elhelyezni az új referencia genomon. Összesen 12.300.815 indel variáns került átemelésre (245.390-et nem sikerült átemelni). Az SNP-k esetében a referencia és az alternatív allél 2.660.671 esetben került felcserélésre, továbbá 39.150 variáns referencia allélja külön hozzá lett adva a genomhoz. Indelek esetében a referencia allél 1.037.351 esetben cserélődött fel az alternatív alléllal, és 1.000.501 variáns referencia allélja lett hozzá adva a referencia genomhoz.

Az extrém idős kutyák szűrt variánsai

Az átlagos korú kutyák variánsainak kiszűrése és az adatfeldolgozások után az extrém idős kutyákban azonosított variánsok közül 129.167 SNP és 51.431 indel maradt, mely mind a tíz idős kutyában megvolt. A legalább hét extrém idős kutyában jelenlévő 2.747.639 SNP-ből 388.918 maradt, míg az indelek 1.307.459-es száma 120.665 darabra csökkent. Az összes kutyában azonosított SNP-k 7.650 gént, az indelek 30.748 gént érintettek, míg a kutyák 70%-ában jelenlévő SNP-k 13.627 gént, az indelek 30.748 gént (4. táblázat).

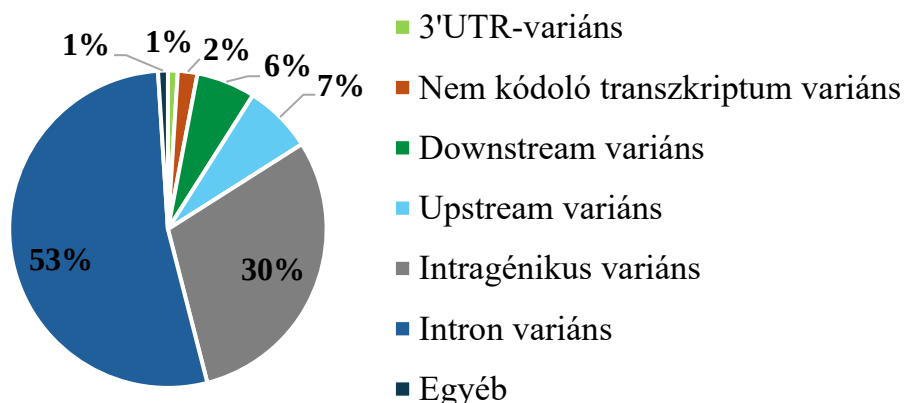
4. táblázat: Az extrém idős kutyák variánsai a szűrés után. Az új variánsok azon variánsokat jelenti, melyeket korábban nem publikáltak.

Variánsok	Számuk (db)	Új variánsok (db)	érintett gének száma (db)	missense mutációk száma (db)
<i>SNP 10/10 kutyában</i>	129.167	73.736	7.650	925
<i>SNP 7/10 kutyában</i>	388.918	188.725	13.627	2.845
<i>Indel 10/10 kutyában</i>	51.431	50.599	30.748	0
<i>Indel 7/10 kutyában</i>	120.665	114.419	30.748	0

A *Variant effect predictor* eredménye alapján a minden extrém idős kutyában jelenlévő SNP-k közül összesen 925 db okoz *missense* mutációt – mely során funkcionálisan eltérő aminosav kerül a kódolt fehérjébe –, a kutyák legalább 70%-ában lévő SNP-k közül pedig 2.845 db. Ezek rendre 302 és 900 gént érintenek

Az SNP-k és az indelek esetében is a detektált variánsok többsége introni (>50%-a) és intragénikus régiókban (26-30%-a) helyezkedik el (2. ábra). A 10/10 kutyában az exoni szekvenciákat érintő SNP-k 54%-a *missense* mutációt, 43%-a szinonim mutációt okoz, míg a 7/10 kutyában 52%-uk *missense* mutációt, 45%-uk szinonim mutációt okoz.

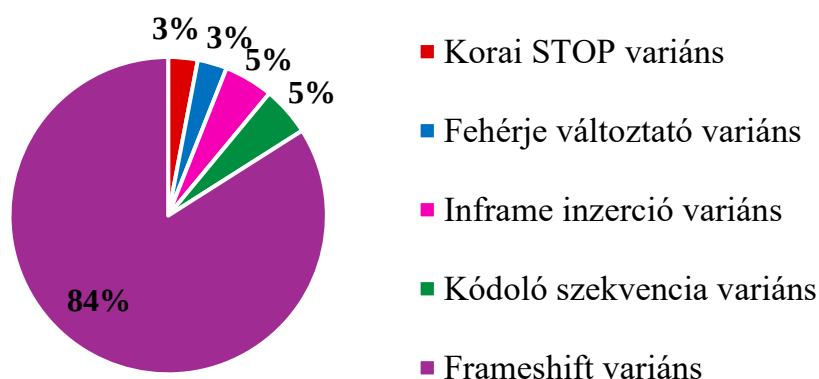
SNP-k által érintett genomi régiók annotációja



2. ábra: Kördiagramm a tíz extrém idős kutyában jelenlévő SNP-k annotációjának megoszlásáról (McLaren és mtsai, 2016).

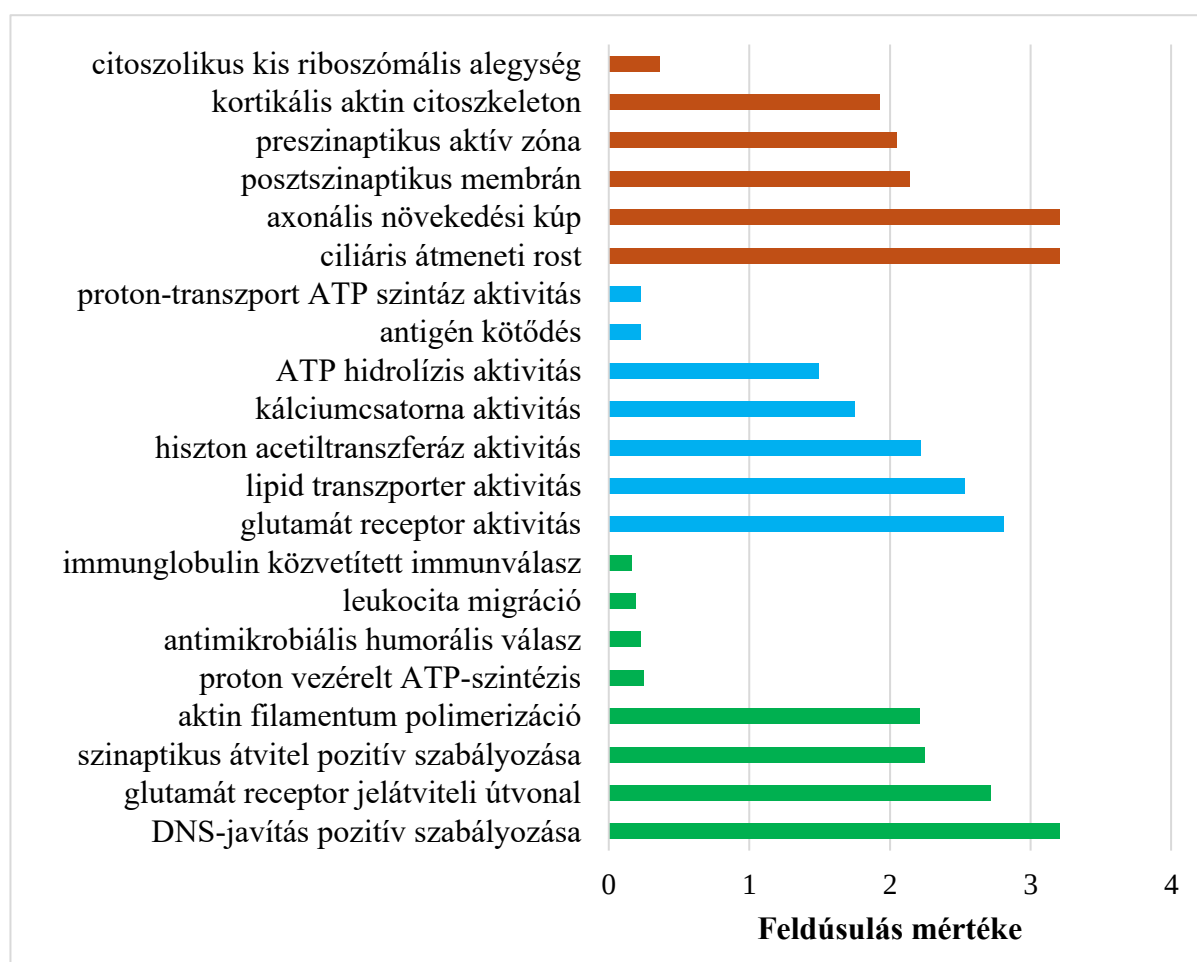
Az exoni szekvenciákat érintő variánsok esetében a 10/10 kutyában jelenlévő indelek 84%-a (115 db, 64 gént érintő) frameshift mutációt, 5%-a inframe inzerció mutációt, 3%-a értelmetlen (korai STOP) mutációt okoz, 3%-ának egyéb fehérje változtató hatása van, 5%-a pedig olyan kódoló szekvencia variáns, mely nem változtatja meg a kódolt fehérjét (3. ábra). A kutyák 70%-ában jelenlévő indelek 69%-a (195 db, 133 gént érintő) frameshift mutációt, 13% inframe inzerció mutációt okoz, 3%-ának az előbbiektől eltérő fehérje változtató hatása van, 14%-a pedig egyéb kódoló szekvencia variáns.

Kódoló genomi régiókat érintő indelek hatásai



3. ábra: Kördiagramm a tíz kutyában jelenlévő indel variánsok hatásairól (McLaren és mtsai, 2016).

A génontológiai elemzés számos szignifikáns eredményt adott mind az SNP-k, mind az indelek esetében. Az összes kutyában jelenlévő SNP-k által érintett gének elemzésében a legerőteljesebb feldúsulás a *DNS-javítás pozitív szabályozásában* volt. További feldúsulást találtunk a transzmembrán transzporter és jelátviteli receptor aktivitásában, a hiszton acetiltranszferáz aktivitásában, a fehérje foszforilációban, a kalcium- és kloridion transzportban (például *kalcium csatorna aktivitás* és *klorid csatorna aktivitás*), az ATP hidrolízisben, több idegrendszeri folyamatban, celluláris komponensben (például *szinaptikus átvitel pozitív szabályozása* és *posztiszinaptikus membrán*) és a sejtvázzal kapcsolatos folyamatokban (például *mikrofilamentum motoros aktivitás*). Ellenben néhány ATP szintézissel kapcsolatos és az immunológiai folyamatok – mint például az *antigén kötődés*, a *baktérium ellen kialakuló védelmi válasz* vagy az *antimikrobiális humorális válasz* – alul voltak reprezentálva (4. ábra).



4. ábra: Az összes kutyában jelenlévő SNP variánsok által érintett génekre lefuttatott génontológiai elemzés néhány eredménye. A narancssárga színek a celluláris komponens, a világos kék szín a molekuláris funkció, a zöld szín a biológiai folyamatok elemzés eredményét jelöli.

A legalább hét kutyában jelenlévő SNP-k által érintett gének több metabolikus folyamatban (például nitrogénvegyületek anyagcsere-folyamatai, szerves anyagcsere-folyamatok, foszfáttartalmú vegyületek anyagcsere-folyamatai) is részt vesznek.

A 10/10 kutyában lévő indelek esetében is a vártnál kevésbé voltak érintettek az immunológiai folyamatokban (például *immunglobulin közvetített immunválasz*) részt vevő gének. Feldúsulást találtunk többek között az RNS transzport és processzálsban (például *tRNS processzáls*, *rRNS processzáls*, *RNS 3'-vég processzáls* és *mRNS splicing*), a transzkripció aktiválásban, kalcium- és kloridion csatorna aktivitásokban, az ATP hidrolízisében, a nukleotid kötődésben (például *egyszálú RNS-kötés* és *mRNA 3'-UTR kötés*), a lipid transzport és processzáls folyamatokban, valamint a *vakuoláris transzportban* és az azzal kapcsolatos folyamatokban (például *makroautofágia* és *vezikulafúzió*). A *vérnyomás negatív szabályozása*, a *szisztémás artériás vérnyomás szabályozása* és a *szöveti homeosztázis* biológiai folyamatok erősen alul voltak reprezentálva az indelek génontológiai elemzésében.

Továbbá a celluláris komponens elemzésben a mind a tíz kutyában jelenlévő SNP-k többek között a csillók, a citoszkeleton, a citoplazma, a sejtkapcsolatok, a riboszóma kis és nagy alegységének felépítésében vesznek részt (4. ábra), az indelek által érintett gének például az intracelluláris organellumok, a sejtmag, a plazmamembrán és a mitokondrium felépítésében.

Diszkusszió

Az extrém idős kutyák adatainak egy része 2017-ben, másik része 2023-ban került megszekvenálásra. Mindkét szekvenálást ugyanaz a cég végezte. Célunk ezzel az is volt, hogy a cég váltásával bekövetkező esetleges változásokat kiszűrjük. A 2023-ban szekvenált teljes genom adatok (old_dog3-11) minősége jelentősen alacsonyabb volt a 2017-ben szekvenált adatokhoz (old_dog1-2) képest. Kevesebb leolvasással rendelkezünk ezektől a kutyáktól és az adapter szennyezés miatt a leolvasásokról le kellett vágnunk az adaptereket. A levágás után az old_dog3-11 leolvasásai bár rövidebbek lettek, átlagos hosszúságuk (137,53 bp) elég hosszú maradt ahhoz, hogy az illesztő programok megtalálják az illesztési helyüket a genomon, így nem befolyásolta szignifikánsan, negatív irányba az illesztés pontosságát, viszont csökkentette a lefedettséget. A kisebb lefedettséghez hozzájárulhatott a szekvenálás során bekövetkezett DNS minták túlzott feldarabolása is, így sok leolvasás azonos információt hordozott a genomról, melyek átfednek az illesztésnél. Az elemzés során megjelöltük a PCR során esetlegesen bekövetkező duplikációkat (ugyanazon leolvasás sokszoros előfordulása), melyek

enélkül a lépés nélkül befolyásolhatják a valódi lefedettség mértékét. A megfelelő lefedettség a variáns hívásnál válik fontossá. A nagyobb genom lefedettség növeli a variánsok azonosításának érzékenységet (Sun és mtsai, 2021). Ezután fontos volt szűrni a kimutatott variánsokat, hogy elkerüljük a téves azonosítást és az álpozitív találatokat.

Jónás és mtsai. (2020) korábbi tanulmányához képest sokkal több mutációt találtam (11,2 millió SNP a korábbi 6.532.384 mutációval szemben, illetve 5,9 millió indel a korábbi 836.296 indellel szemben). Az átfedéseket tekintve 2,7 millió SNP és 1,5 millió indel megvolt az extrém idős kutyák legalább 70%-ban szemben a korábbi ~3 millió SNP-vel és ~0,3 millió indellel (Jónás et al., 2020, 2. táblázat). Ez mindkettő egyértelműen a sokkal nagyobb elemszámnak tudható be. Hasonló trend volt megfigyelhető a korábban nem publikált mutációk esetében is.

A különböző autoszómákon és az ivari kromoszómákon kimutatott variánsok száma a vártnak megfelel. A legtöbb variánst az első kromoszóma hordozta, mely a kutyák legnagyobb kromoszómája, a legkevesebb variánst pedig a sokkal kisebb 36-os és Y kromoszómákon találtam.

Az átlagos korú kutyák adatainak átemelése a *ROS_Cfam_1.0* referencia genomra sikeres volt, mindössze az SNP-k ~1,6%-át (337.643 db) és az indelek ~2%-át (245.390 db) nem tudtunk átemelni. Összesen 638 átlagos korú kutya adatai álltak rendelkezésre, így az átemelés után is nagyon sok variáns maradt. A *bcftools +liftover* volt a legalkalmasabb eszköz az átemelésre, mivel a chain fájl és a genom nem teljes kompatibilitása esetén a program nem dobja el automatikusan a nem illeszkedő variánsokat, hanem képes kicserélni a referencia allélt és az alternatív allélt, ha a két referencia genom genotípusa az adott variáns helyén különbözik, illetve hozzá tudja adni a referencia allélt, ha a chain fájlban az adott helyen „gap” található (Genovese és mtsai, 2024). Fontos megjegyezni, hogy az átlagos korú kutyák adatai között nem szerepelt az Y kromoszóma, mivel a CanFam 3.1-es (tehát a korábbi) referencia genomban az nem volt meg (az egy szuka kutyából készült), ezért az Y kromoszómára eső variánsokat nem tudtam szűrni.

A szűrés után a csak az extrém idős kutyákban előforduló variánsok arányai szintén megfelelnek a vártaknak. SNP-ből több variánst kaptunk, mint indelből, ami annak tudható be, hogy egy inzerció vagy delécio jelentősebb hatással bír, mint egy mindössze egy nukleotidot érintő SNP mutáció. A missense mutációk száma szintén jelentősen magasabb volt a jelen tanulmányban (n=925), mint a korábban publikált munkában (n=67; Jónás és mtsai, 2020). Ez

részben annak tudható be, hogy a korábbi tanulmányban nem csak az általam is használt mutációkat, hanem két másik adatbázisban szereplő mutációkat is kiszűrték.

A variánsok által érintett gének számának meghatározásához használt génlista a fehérjekódoló géneken kívül az RNS-eket kódoló géneket is tartalmazta. Az indelek által érintett gének száma kiemelkedően magas, a *ROS_Cfam_1.0* referencia genom összes génjét (30.748 db) érintik. Ez a szám egy technikai hiba eredménye lehet, így az indelek elemzését újra le kell ellenőriznünk és szükség esetén kijavítanunk.

A variánsok többsége introni vagy intragénikus régiókban helyezkedett el. Ez a várakozásoknak megfelelő, hiszen ezeket jobban tudják tolerálni a genomok a kódoló régiókat érintő mutációknál, így pedig könnyebben felhalmozódnak az evolúció során. Ettől függetlenül, ezeknek a mutációknak hatása lehet a gén kifejeződésére, ha például szabályozó régiót érintenek. Ezzel ellentétben az exoni szekvenciákban keletkező mutációk könnyebben vezetnek nem kívánt negatív hatáshoz, például funkcióvesztéshez. Ha a kódolásban következett be változás egy SNP miatt, akkor az esetek 54%-ában funkcionálisan eltérő aminosav épült be, 43%-ában pedig szinonim a mutáció hatása. A kódoló szekvenciákat érintő indelek az esetek legnagyobb százalékában (84%) egy nagyon komoly mutációt, frameshift mutációt eredményeztek, mely során elcsúszik az adott gén olvasási kerete (Jónás és mtsai, 2020).

A génontológiai elemzés a vártak megfelelően számos biológiai folyamatnál, molekuláris funkcionál és celluláris komponensnél detektált szignifikáns eredményt. Érdemes megemlíteni, hogy az elemzés nem szolgáltat információt arról, hogy az érintett gének az adott folyamatban milyen szerepet töltenek be: serkentenek vagy gátolnak. Az egyik legjelentősebb feldúsulást a DNS javítással kapcsolatos folyamatban találtuk. Az eukarióta sejtekben fejlett DNS hibajavító rendszer működik a genom eredeti információ-tartalmának megtartása érdekében. Az öregedés során mutációk halmozódnak fel a genomban, mely feltételezhetően részben a DNS hibajavítás hanyatlásának a következménye (López-Otín és mtsai, 2013; Niedernhofer és mtsai, 2018). Az SNP és indel variánsok kevesebb immunológiával kapcsolatos gént érintettek, mint ahogy a referencia génlista alapján vártuk. Ezzel szemben az idegrendszer erősen érintett a biológiai folyamatokat, a molekuláris funkciót és a celluláris komponens elemzést tekintve is. A kor előrehaladtával gyakran alakulnak ki idegrendszerrel kapcsolatos betegségek mind emberben, mind kutyában: például a humán Alzheimer-kór és az azzal hasonló kutya kognitív diszfunkció a kutyákban (Ruple és mtsai, 2022).

A kutatás folytatásként tervezzük az átlagos korú kutyák adatbázisát kiterjeszteni mindhárom másik adatbázisra a Jónás et al. (2020) cikk alapján, illetve külön elemezni fogjuk azokat az SNP és indel mutációkat is, melyek az extrém idős kutyák legalább 70%-ában megtalálhatóak voltak. Végezetül, kópiaszám variációkat is tervezünk keresni a teljes genom szekvenciák alapján és azokat összevetjük a most leírt SNP és indel mutációkkal, valamint az általuk érintett géneket külön is vizsgálni fogjuk.

Korlátok

A kutatás legfőbb limitációja az, hogy a letöltött 638 átlagos korú kutya között lehet olyan egyed, amelyik extrém sokáig élt. Ezt nem állt módunkban leellenőrizni, ugyanakkor ennek rendkívül kicsi az esélye.

Fontos megemlíteni, hogy a kutatást modellállaton végeztük, nem pedig emberen. Bár nagyon sok előnyös és közös tulajdonsága van a kutyának az emberrel, befolyásolják olyan tényezők az élettartamát, melyek az emberek élethosszát nem bizonyítottan befolyásolják (például a testméret). Az emberi öregedést rendkívül ritka egyének tanulmányozásával lehetne kivitelezni, akik egészségesek és 100 évig vagy annál tovább élnek (Zhang és mtsai, 2020). Továbbá az egészséges öregedést nagyon nehéz mérni (a betegségeket tudják azonosítani, nem az egészséget), melyhez szükség van egy konkrét definícióra is. Például az egészséges öregedés mást jelent a kutatók számára, mint az idős embereknek: több idős ember gondolja magáról, hogy egészséges öregedést tapasztal, mint amennyit a tudósok megállapítanak (Behr és mtsai, 2023).

Köszönetnyilvánítás

Szeretnék köszönetet mondani témavezetőmnek, Dr. Jónás Dávidnak, hogy kutatási lehetőséget biztosított számomra és akihez bármikor fordulhattam, ha kérdésem volt és aki fontos tanácsokkal látott el az adatelemzések során és értelmezésükkel kapcsolatban. Továbbá szeretném megköszönni a családomnak, hogy támogattak a dolgozat megírásában.

Irodalom

1. Behr LC, Simm A, Kluttig A, Grosskopf A. 60 years of healthy aging: On definitions, biomarkers, scores and challenges. *Ageing Res Rev.* 2023;88:101934. doi:10.1016/j.arr.2023.101934
2. Sun Y, Liu F, Fan C, et al. Characterizing sensitivity and coverage of clinical WGS as a diagnostic test for genetic disorders. *BMC Med Genomics.* 2021;14(1):102. doi:10.1186/s12920-021-00948-5
3. Thomas PD, Kejariwal A, Campbell MJ, et al. PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res.* 2003;31(1):334-341. doi:10.1093/nar/gkg115
4. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol.* 2016;17(1):122. doi:10.1186/s13059-016-0974-4
5. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. *Bioinformatics.* 2011;27(15):2156-2158. doi:10.1093/bioinformatics/btr330
6. Harrison PW, Amode MR, Austine-Orimoloye O, et al. Ensembl 2024. *Nucleic Acids Res.* 2024;52(D1):D891-D899. doi:10.1093/nar/gkad1049
7. Genovese G, Rockweiler NB, Gorman BR, et al. BCFtools/liftover: an accurate and comprehensive tool to convert genetic variants across genome assemblies. *Bioinformatics.* 2024;40(2). doi:10.1093/bioinformatics/btae038
8. Hoepfner MP, Lundquist A, Pirun M, et al. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One.* 2014;9(3):e91172. doi:10.1371/journal.pone.0091172
9. Zhang ZD, Milman S, Lin JR, et al. Genetics of extreme human longevity to guide drug discovery for healthy ageing. *Nat Metab.* 2020;2(8):663-672. doi:10.1038/s42255-020-0247-0
10. Chen Q, Thompson J, Hu Y, Lesnefsky EJ. Chronic metformin treatment decreases cardiac injury during ischemia-reperfusion by attenuating endoplasmic reticulum stress with improved mitochondrial function. *Aging.* 2021;13(6):7828-7845. doi:10.18632/aging.202858
11. Anisimov VN, Bernstein LM, Popovich IG, et al. If started early in life, metformin treatment increases life span and postpones tumors in female SHR mice. *Aging.* 2011;3(2):148-157. doi:10.18632/aging.100273
12. Bitto A, Ito TK, Pineda V V, et al. Transient rapamycin treatment can increase lifespan and healthspan in middle-aged mice. *Elife.* 2016;5:e16351. doi:10.7554/eLife.16351
13. Herrera JJ, Pifer K, Louzon S, et al. Early or Late-Life Treatment With Acarbose or Rapamycin Improves Physical Performance and Affects Cardiac Structure in Aging Mice. *J Gerontol A Biol Sci Med Sci.* 2023;78(3):397-406. doi:10.1093/gerona/glac221

14. Franceschi C, Garagnani P, Morsiani C, et al. The Continuum of Aging and Age-Related Diseases: Common Mechanisms but Different Rates. *Front Med (Lausanne)*. 2018;5:61. doi:10.3389/fmed.2018.00061
15. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;17(1):10. doi:10.14806/ej.17.1.200
16. Quinlan AR. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr Protoc Bioinformatics*. 2014;47. doi:10.1002/0471250953.bi1112s47
17. Robinson JT, Thorvaldsdottir H, Turner D, Mesirov JP. igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). *Bioinformatics*. 2023;39(1):btac830. doi:10.1093/bioinformatics/btac830
18. Van der Auwera GA, O'Connor BD. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra (1st Edition)*. O'Reilly Media, Inc.; 2020.
19. Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021;10(2). doi:10.1093/gigascience/giab008
20. Picard. Accessed June 27, 2023. <http://broadinstitute.github.io/picard/>
21. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760. doi:10.1093/bioinformatics/btp324
22. S Bittencourt a. FastQC: A quality control tool for high throughput sequence data. *Babraham Bioinformatics*. Published online 2010.
23. Harrison PW, Amode MR, Austine-Orimoloye O, et al. Ensembl 2024. *Nucleic Acids Res*. 2024;52(D1):D891-D899. doi:10.1093/nar/gkad1049
24. Demirtas A, Atilgan D, Saral B, et al. Dog owners' recognition of pain-related behavioral changes in their dogs. *Journal of Veterinary Behavior*. 2023;62:39-46. doi:10.1016/j.jveb.2023.02.006
25. McCoy BM, Brassington L, Jin K, et al. Social determinants of health and disease in companion dogs: a cohort study from the Dog Aging Project. *Evol Med Public Health*. 2023;11(1):187-201. doi:10.1093/emph/eoad011
26. Yordy J, Kraus C, Hayward JJ, et al. Body size, inbreeding, and lifespan in domestic dogs. *Conserv Genet*. 2020;21(1):137-148. doi:10.1007/s10592-019-01240-x
27. Lindblad-Toh K, Wade CM, Mikkelsen TS, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*. 2005;438(7069):803-819. doi:10.1038/nature04338
28. Plassais J, Kim J, Davis BW, et al. Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nat Commun*. 2019;10(1):1489. doi:10.1038/s41467-019-09373-w
29. Lee H, Collins D, Creevy KE, Promislow DEL, Dog Aging Project Consortium. Age and Physical Activity Levels in Companion Dogs: Results From the Dog Aging Project. *J Gerontol A Biol Sci Med Sci*. 2022;77(10):1986-1993. doi:10.1093/gerona/glac099

30. McKenzie BA, Chen F, LaCroix-Fralish ML. The phenotype of aging in the dog: how aging impacts the health and well-being of dogs and their caregivers. *J Am Vet Med Assoc.* 2022;260(9):963-970. doi:10.2460/javma.22.02.0088
31. Urfer SR, Kaeberlein M, Promislow DEL, Creevy KE. Lifespan of companion dogs seen in three independent primary care veterinary clinics in the United States. *Canine Med Genet.* 2020;7(1):7. doi:10.1186/s40575-020-00086-8
32. Shindyapina A V., Cho Y, Kaya A, et al. Rapamycin treatment during development extends life span and health span of male mice and *Daphnia magna*. *Sci Adv.* 2022;8(37):eabo5482. doi:10.1126/sciadv.abo5482
33. Wang K, Liu H, Hu Q, et al. Epigenetic regulation of aging: implications for interventions of aging and diseases. *Signal Transduct Target Ther.* 2022;7(1):374. doi:10.1038/s41392-022-01211-8
34. Ruple A, MacLean E, Snyder-Mackler N, Creevy KE, Promislow D. Dog Models of Aging. *Annu Rev Anim Biosci.* 2022;10(1):419-439. doi:10.1146/annurev-animal-051021-080937
35. Taormina G, Ferrante F, Vieni S, Grassi N, Russo A, Mirisola MG. Longevity: Lesson from Model Organisms. *Genes (Basel).* 2019;10(7):518. doi:10.3390/genes10070518
36. Erikson GA, Bodian DL, Rueda M, et al. Whole-Genome Sequencing of a Healthy Aging Cohort. *Cell.* 2016;165(4):1002-1011. doi:10.1016/j.cell.2016.03.022
37. Jónás D, Sándor S, Tátrai K, Egyed B, Kubinyi E. A Preliminary Study to Investigate the Genetic Background of Longevity Based on Whole-Genome Sequence Data of Two Methuselah Dogs. *Front Genet.* 2020;11:315. doi:10.3389/fgene.2020.00315
38. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The Hallmarks of Aging. *Cell.* 2013;153(6):1194-1217. doi:10.1016/j.cell.2013.05.039
39. Niedernhofer LJ, Gurkar AU, Wang Y, Vijg J, Hoeijmakers JHJ, Robbins PD. Nuclear Genomic Instability and Aging. *Annu Rev Biochem.* 2018;87:295-322. doi:10.1146/annurev-biochem-062917-012239