

Studiul genei *CG18135* de la *Drosophila melanogaster* prin utilizarea unor metode avansate de Analiză Genetică, Genomică și Bioinformatică

AOSR-TEAMS II EDIȚIA 2023-2024, Nr. 1 - DOMENIUL ȘTIINȚE BIOLOGICE

Scopul acestui proiect este adnotarea funcțională a genei *CG18135* de la *Drosophila melanogaster* și detalierea rețelor de interacțiuni la care participă această genă/proteină (Figura 1).

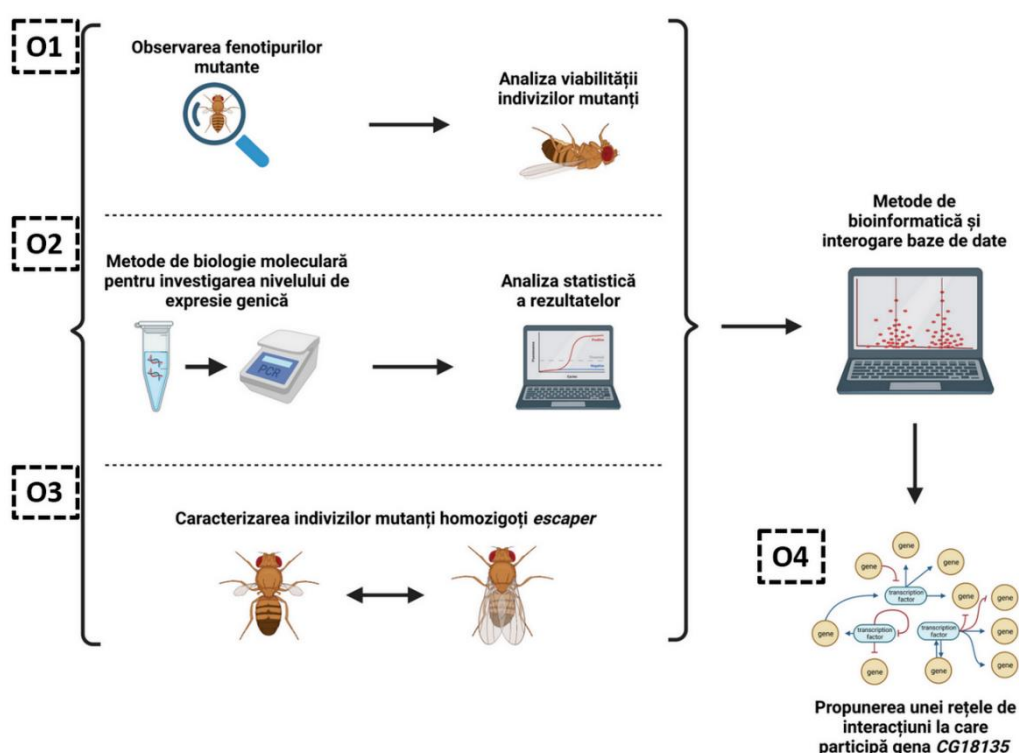


Figura 1. Prezentarea schematică a obiectivelor prezentului proiect de cercetare.

În cadrul investigațiilor realizate pentru raportarea precedentă, am identificat motivul de legare al factorilor de transcriere SP1/KLF în secvența genomică asociată cu inserția *P{lacW}* în gena *CG18135*. Din familia factorilor de transcriere SP1/KLF fac parte proteinele Sp1, buttonhead și cabut, care participă în variate procese de dezvoltare sau metabolism energetic. Implicațiile teoretice ale acestor investigații sunt prezentate în detaliu în articolul „A novel insertional allele of the *CG18135* gene is associated with severe mutant phenotypes in *Drosophila melanogaster*” publicat în revista *Frontiers in Genetics* (Ratiu și colab., 2024).

Precum a fost prezentat în raportul intermediar 2 (din decembrie 2023), omologul genei *CG18135* la eucariote superioare este *GPCPD1* (*glycerophosphocholine phosphodiesterase 1*), iar participarea în metabolismul energetic este documentată. În cadrul acestui raport intermediar am investigat contextul interacțiunilor biochimice în care este implicată proteina corespunzătoare genei *GPCPD1*. Astfel, am identificat o proteină de interes, CHAT, situată în avalul proteinei *GPCPD1* în calea biochimică a metabolismului glicerofosfolipidelor (din baza de date KEGG). Figura 1 prezintă această cale biochimică și localizarea proteinelor *GPCPD1* și CHAT.

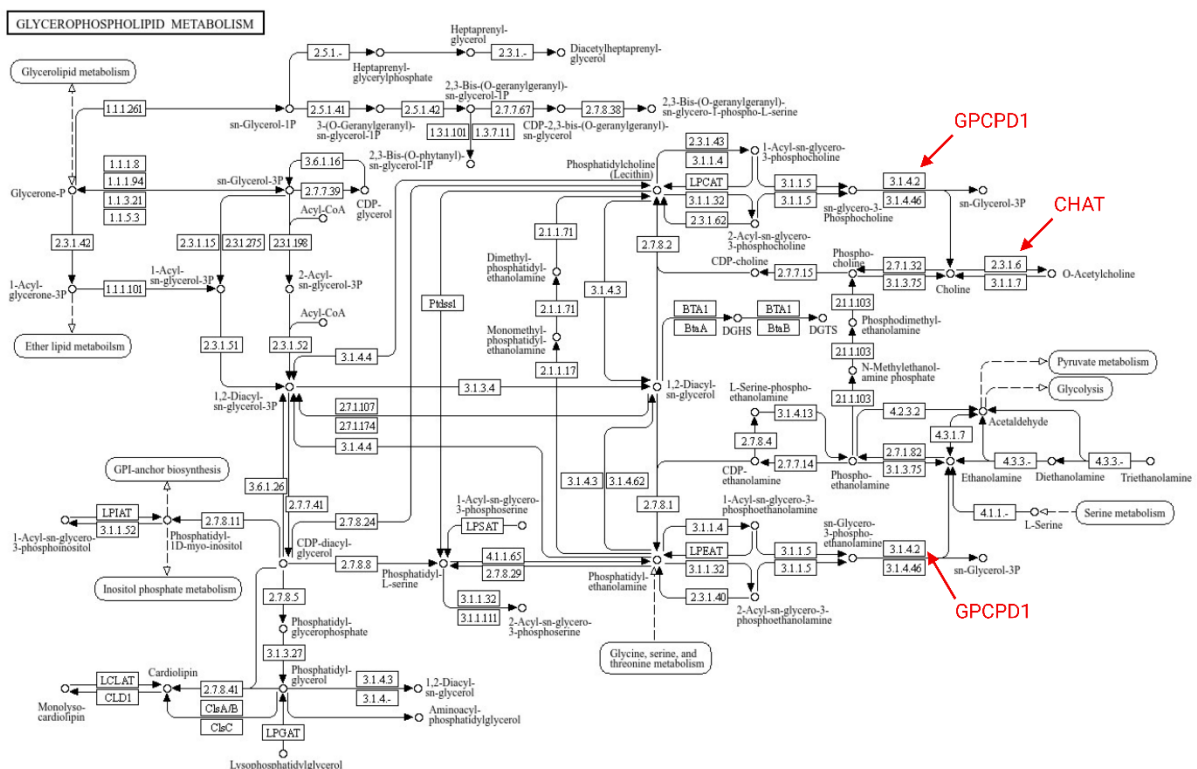


Figura 1. Calea biochimică a metabolismului glicerofosfolipidelor. Calea biochimică a fost obținută prin coroborarea informațiilor disponibile în baza de date KEGG, iar editările au fost create cu BioRender.com.

Adițional, pentru evaluarea nivelului de expresie genică am ales gena *Myo10A* de la *Drosophila melanogaster*. Funcția proteinei corespunzătoare a fost anterior corelată cu

proteina asociată genei *CG18135* în legarea fasciculelor de miozină (detalii prezentate în raportul intermediar 2). De asemenea, am ales gena *CG11619* datorită similarității de secvență crescută cu gena *CG18135* (29,74%) și a vecinătății cu aceasta (Figura 2).

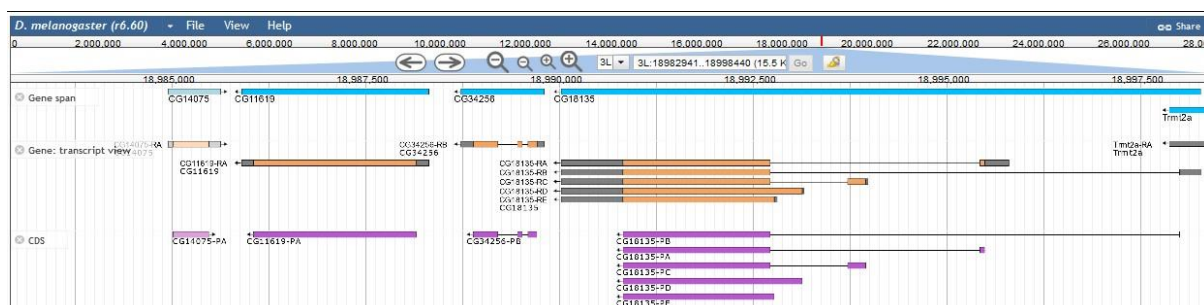


Figura 2. Contextul genomic al genelor *CG18135* și *CG11619* la *D. melanogaster*. Captură de ecran din JBrowse (FlyBase.com).

Pentru toate genele de interes alese, au fost creați primeri specifici cu scopul de a evalua nivelul de expresie genică. Gena *RpL32* a fost considerată pentru normalizarea datelor Ct țintă, iar primerii corespunzători au fost identificați în literatură (Fiumera și colab., 2005). Detaliile corespunzătoare acestora sunt prezentate în Tabelul 1.

Tabelul 1. Detaliile primerilor selectați pentru qRT-PCR. Abrevierile prezentate denotă: HK = *housekeeping*, GOI = *gene of interest*, F = *forward*, R = *reverse*, nt = nucleotidă.

Gena	Tip genă	Orientare primer	Secvența	Temperatura de topire	Procent CG	Lungime amplicon
<i>RpL32</i>	HK	F	CGGCTTCAAGGGACAGTATC	62°C	55%	196 nt
		R	GACAATCTCCTTGCGCTTCT	60°C	50%	
<i>CG18135</i>	GOI	F	ATCTTTCACATCACGCTGCC	55.8°C	50%	201 nt
		R	GACATAGGGAAGCCTCAGCC	57.5°C	60%	
<i>CHAT</i>	GOI	F	CGGAGGATAACTGGGCCTAT	53.8°C	55%	191 nt
		R	CCCACTGTCCAGCATCTCGC	57.9°C	65%	
<i>CG11619</i>	GOI	F	GCCAGTTTCCATCACTACTGG	54.4°C	52.4%	190 nt
		R	CGGGAACGTAGTCCTTGGTC	51.1°C	60%	
<i>Myo10A</i>	GOI	F	TGATGAAGCAGACGACGGC	53.2°C	57.9%	199 nt
		R	CGGAGATTGGTCAGACAAACG	54.4°C	52.4%	

Consecutiv testelor de reversie am reușit să obținem două descendențe distincte, simbolizate **II** și, respectiv, **H1**, în care coexistă diferite genotipuri determinate de cromozomi 3, mai precis Excizie/ $CG18135^{P\{lacW\}CG18135}$, $CG18135^{P\{lacW\}CG18135}$ /TM6B, Excizie/TM6B, Excizie/Excizie. Prin Excizie (Ex) simbolizăm cromozomul 3 din care a fost excizat transpozonul $P\{lacW\}^{CG18135.MZ4CM3}$. Testarea anterioară a unor genotipuri aparținând II au demonstrat că indivizii descendenți contin întradevar transpozonul $P\{lacW\}$ specific, însă nu am realizat o astfel de demonstrație în cazul H1. Însă, mai important, nu aveam nicio informație la nivel de secvență nucleotidică referitoare la consecințele efective pe care le-au avut transpozițiile $P\{lacW\}$ care au generat fenomenele de reversie. Din punct de vedere genetic, o alelă de reversie completează cu alela inserțională originală și indivizii care conțin cele două alele sunt viabili și fertili. La nivelul moleculei de ADN, excizia transpozonului poate fi precisă, refăcându-se arhitectura cromozomială anterioara inserției, sau imprecisă, ceea ce înseamnă că în urma transpoziției pot fi introduse mutații de mici dimensiuni, de tipul delețiilor sau inserțiilor. În plus, noi am evidențiat două fenomene de reversie distincte, care ar putea avea consecințe diferite asupra secvenței ADN din preajma situsului de inserție original.

Astfel, pentru câteva categorii de indivizi aparținând celor două descendențe, II și H1, am realizat o serie de teste de biologie moleculară care să permită cu o rezoluție maximă, la nivel de secvență ADN, analiza $P\{lacW\}^{CG18135.MZ4CM3}$ din situsul caracteristic și, respectiv, a structurii celor două alele de reversie, CG-II Ex și CG-H1 Ex.

Contextul genomic al inserției $P\{lacW\}$ în gena $CG18135$ este prezentat în Figura 3, împreună cu adnotările specifice ale primerilor utilizați pentru reacțiile PCR.

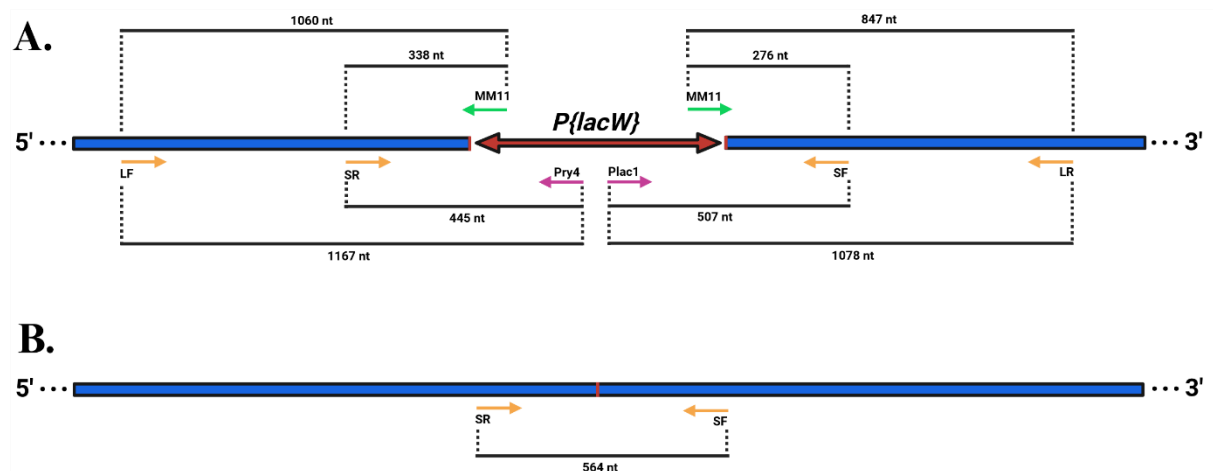


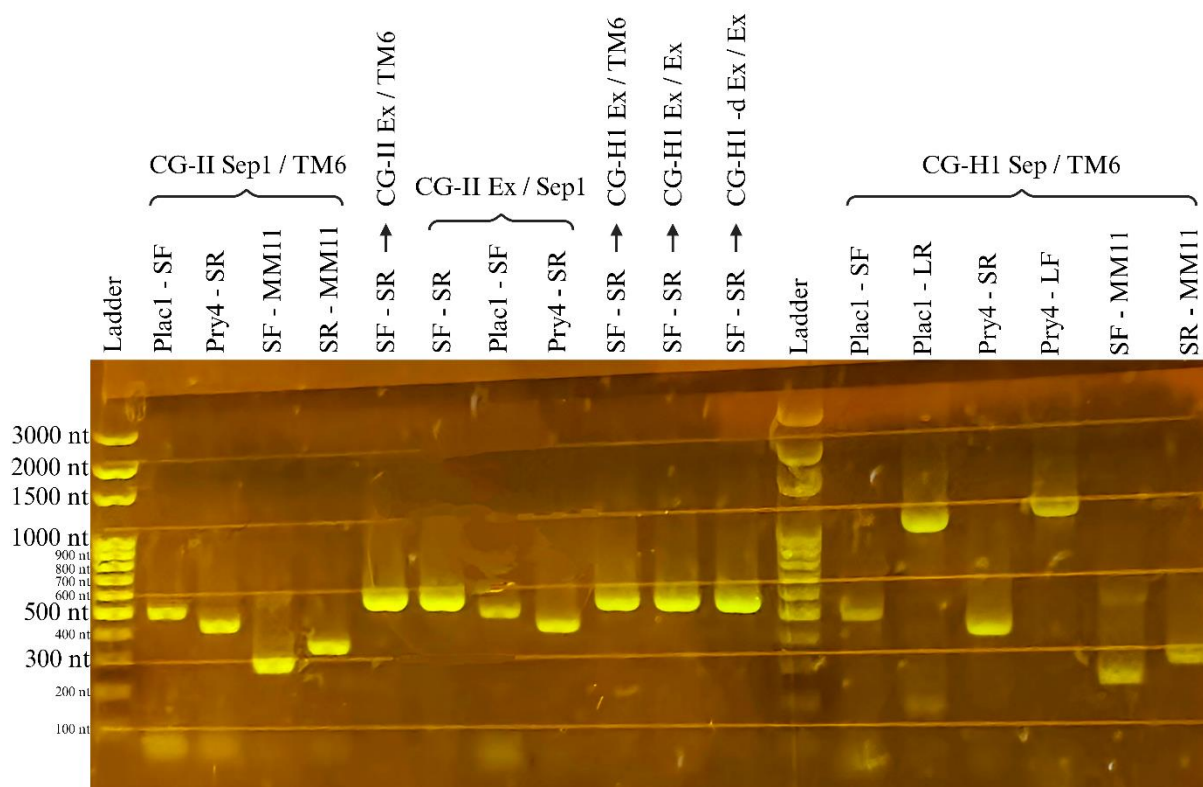
Figura 3. Contextul genomic al genei $CG18135$ în care există inserția $P\{lacW\}$ (A.) și, respectiv, în care transpozonul lipsește (B.). Ampliconii creați în baza primerilor sunt adnotați corespunzător și au prezentate dimensiunile așteptate. Figură creată cu BioRender.com.

Indivizi din liniile de interes au fost utilizați pentru extracție ADN cu *kit*-ul Wizard Genomic DNA Purification (Promega). Eluția finală a fost realizată în volum de 70 μL , iar citirea concentrațiilor obținute a fost realizată cu spectofotometrul ASP-2680 (ATCGene). Detaliile corespunzătoare extracției ADN sunt prezentate în Tabelul 2, unde cu Sep1 indicăm cromozomul 3 care conține inserția $P\{lacW\}^{CG18135.MZ4CM3}$.

Tabelul 2. Detaliile corespunzătoare extracției ADN pentru genotipurile selectate.

Genotip <i>D. melanogaster</i>	Concentrație ADN
CG-II Sep1 / TM6	860.8 ng/ μL
CG-II Ex / TM6	1170 ng/ μL
CG-II Ex / Sep1	1561.8 ng/ μL
CG-H1 Sep / TM6	490.3 ng/ μL
CG-H1 Ex / TM6	397.4 ng/ μL
CG-H1 Ex / Ex	88.3 ng/ μL
CG-H1 -d Ex / Ex	732 ng/ μL
CG18135 Sep1H / GFP	935 ng/ μL

Fiecare reacție PCR a fost realizată într-un volum de 40 μ L cu 1 μ L de ADN pentru fiecare probă, cu excepția probei corespunzătoare genotipului CG-H1 Ex / Ex, unde au fost



utilizați 4 μ L ADN. De asemenea, concentrațiile per reacție PCR au fost: 1X buffer, 3 mM $MgCl_2$, 0.25 mM dNTPs, 2U *Taq-polymerase*, 0.5 μ M primer *forward*, 0.5 μ M primer *reverse* și apă ultrapură. Reacțiile PCR au fost rulate după programul: 95°C pentru 5 minute, 30 de cicluri cu 95°C pentru 30 secunde, 58°C pentru 30 secunde și 72°C pentru 1 minut, 72°C pentru 10 minute și 4°C pentru 1 minut. Pentru vizualizarea ampliconilor obținuți în reacția PCR a fost preparat un gel de agaroză 1.5% în care au fost adăugați 2 μ L SYBR Safe DNA *gel stain* (Invitrogen). Pentru a evalua dimensiunea ampliconilor obținuți a fost utilizat standardul 100bp *DNA Ladder Ready to Load* (Solis Biodyne). Rezultatele migrării gelului de agaroză pentru fiecare probă experimentală sunt prezentate în Figurile 3 și 4.

Figura 3. Rezultatele gelului de agaroză migrat pentru ampliconii corespunzător genotipurilor de interes.

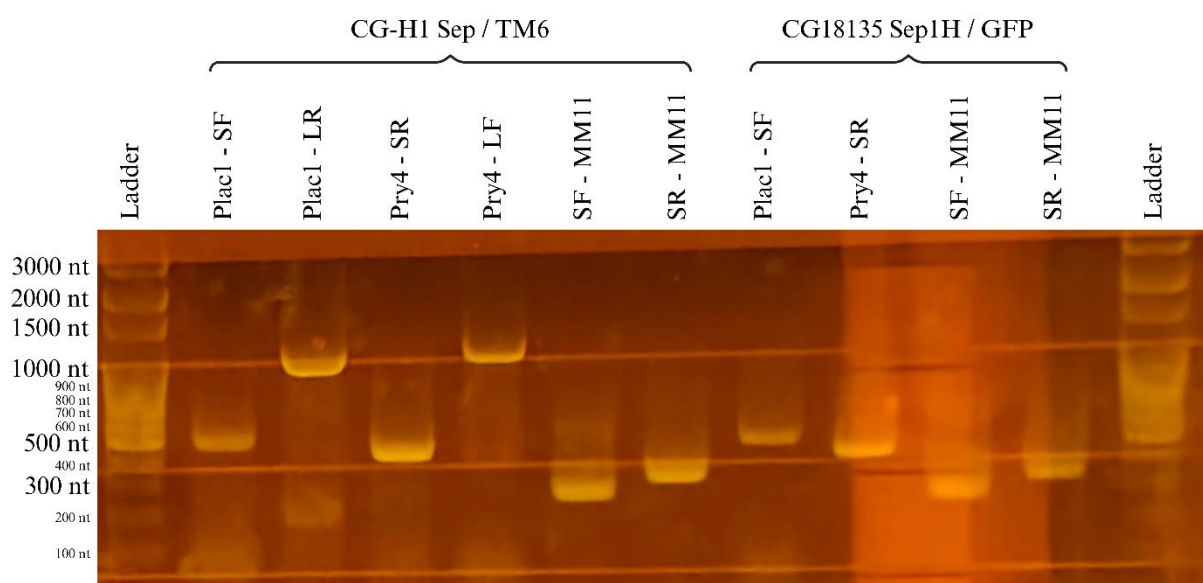


Figura 4. Rezultatele gelului de agaroză migrat pentru ampliconii corespunzători genotipurilor de interes sau genotipului control (linia CG18135 Sep1H / GFP).

Ampliconii considerați de interes au fost selectați pentru purificare din produsul PCR cu *kit*-ul dedicat Genomic DNA Clean & Concentrator (Zymo Research), iar concentrațiile ADN sunt prezentate în Tabelul 3.

Tabelul 3. Detaliile corespunzătoare purificării ampliconilor ADN pentru genotipurile selectate.

Ampliconii de interes au fost secvențiați cu tehnologia Sanger de către firma Eurofins (Eurofins Sanger Sequencing Services). Pentru fiecare probă de interes a fost pregătită o soluție cu 40 ng amplicon în 5 μ L peste care au fost adăugați 5 μ L de primer de concentrație 5 pmoli / μ L. Pentru fiecare amplicon au fost considerate două secvențieri, una cu primer *forward* și cealaltă cu primer *reverse*.

Rezultatele de secvențiere au inclus diferite categorii de fișiere tipice, precum AB1 și FASTA, dar și în format PDF, după cum exemplificăm în Figura 5.

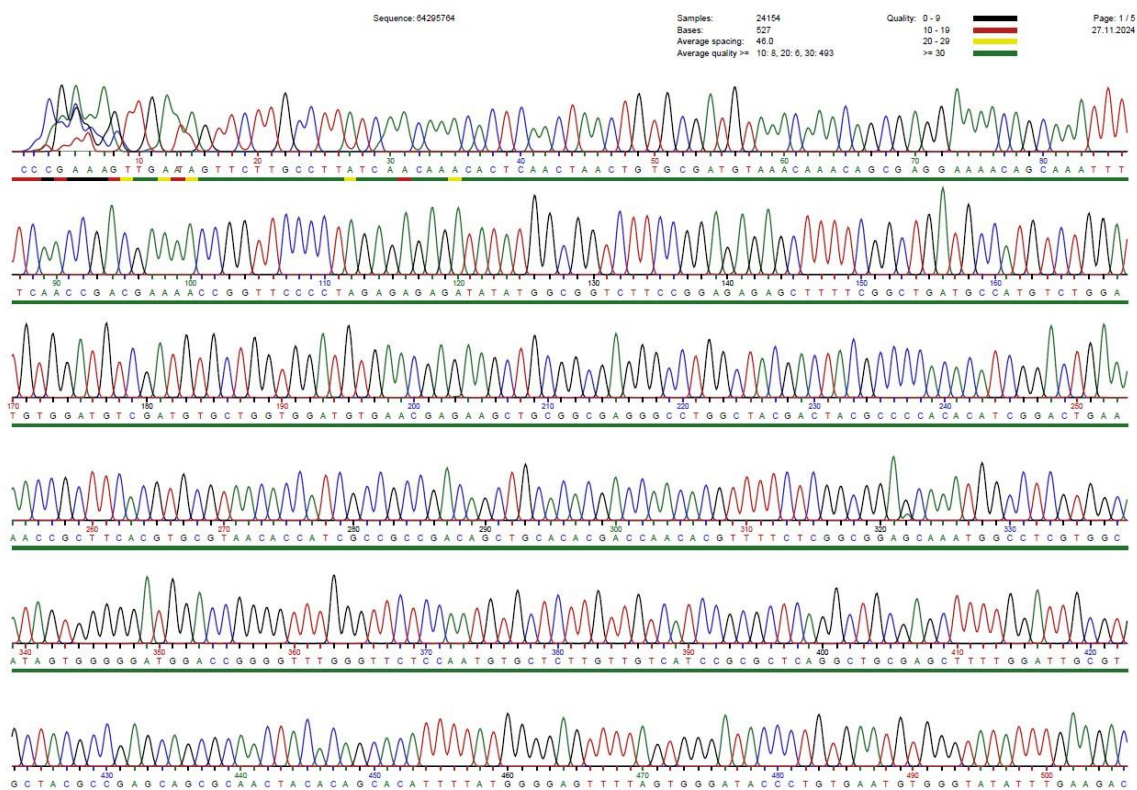


Figura 5. Rezultat de secvențiere Sanger pentru ampliconul corespunzător genotipului CG-H1 Ex / Ex obținut cu perechea de primeri SF – SR.

Pentru analiza inițială a datelor de secvențiere am vrut să evidențiem comparativ vecinătățile inserției $P\{lacW\}^{CG18135.MZ4CM3}$ din genotipul CG-II Sep1 / TM6 și, respectiv, secvența alelei de reversie CG-II Ex din genotipul CG-II Ex / Sep1. Pentru aceasta, prima dată am aliniat comparativ secvențierile care au utilizat primerul SF din ampliconii SF – Plac1 și SF – SR, respectiv. În plus, am afișat și secțiunea SF – inserție $P\{lacW\}$ a ampliconului SF – SR obținut *in silico* pentru secvența de referință a genomului *D. melanogaster* (Figura 6).



Figura 6. Regiunea genomică învecinată situsului de inserție pentru $P\{lacW\}^{CG18135.MZ4CM3}$ în cromozomul care conține transpozonul (64295696), secvența de referință (ref_SF_ins) și cromozomul care conține reversia (64295726, marcat). Săgețile indică regiunile din cadrul electroforegramei de secvențiere în care sunt evidențiate diferențele dintre secvența cromozomului revertant și celelalte două, din referință și cromozom cu inserție (BioEdit și FastStone Image Viewer).

Spre surprinderea noastră, în cromozomul din care a fost excizat transpozonul $P\{lacW\}$, în preajma situsului în care a fost inserat anterior, am identificat două mutații, mai precis deleția AC (g18996807-18996808) și substituția T>C (g18996775). Mutațiile sunt specificate în funcție de catena de referință a cromozomului și au fost confirmate și cu secvențierea din direcția opusă. În regiunea genomică care se învecinează cu celălalt capăt al transpozonului nu am putut evidenția diferențe între referință și cele două categorii de cromozomi 3. O analiză similară, efectuată pentru ampliconi H1 a evidențiat exact aceeași situație. Acest rezultat este intrigant și pentru a avea un răspuns definitiv va trebui să analizăm suplimentar datele de secvențiere și să facem o serie de investigații *in cogitatio*.

În ceea ce privește inserția transpozonului în cromozomii inserționali caracteristici descendenților II și H1, datele de secvențiere au confirmat prezența $P\{lacW\}^{CG18135.MZ4CM3}$ și, respectiv, prezența în copie dublă și structura octetului caracteristic, generat de inserție.

În cadrul acestei raportări sunt prezentate rezultatele unor experimente de genetică și de biologie moleculară, dar și analize bioinformatică care au rolul de a completa obiectivul 4 al proiectului.

Datele obținute în cadrul acestui proiect de cercetare cu ajutorul tehnicii qRT-PCR au fost analizate utilizând o procedură originală, care este prezentată în cadrul unui articol care a

fost acceptat pentru publicare în revista BioTechniques, având titlul „qDATA - an R application implementing a practical framework for analyzing quantitative Real-Time PCR data”. Taxele pentru publicarea acestui articol vor totaliza 2075,36€ = 10328,86RON (2 decembrie 2024).

În afara acestor cheltuieli, costurile estimative pentru obținerea și întreținerea descendențelor tință, II și H1, experimentele de biologie moleculară, procurarea primerilor și secvențiere au fost de 2000 €.

În cadrul acestui proiect au fost publicate două articole cotate ISI, unul în secțiunea *Genetics of Common and Rare Diseases* al revistei *Frontiers in Genetics*, intitulat „A novel insertional allele of the *CG18135* gene is associated with severe mutant phenotypes in *Drosophila melanogaster*”, și unul acceptat și în curs de publicare în revista *BioTechniques*, intitulat „qDATA – an R application implementing a practical framework for analyzing quantitative Real-Time PCR data”.

Bibliografie

Ratiu, A. C., Ionascu, A., & Ecovoiu, A. A. (2024). A novel insertional allele of the *CG18135* gene is associated with severe mutant phenotypes in *Drosophila melanogaster*. *Frontiers in Genetics*, 15, 1355368.

Fiumera AC, Dumont BL, Clark AG. Sperm competitive ability in *Drosophila melanogaster* associated with variation in male reproductive proteins. *Genetics*. 2005;169(1):243-57.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. and Ishiguro-Watanabe, M.; KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 51, D587-D592 (2023).

4 decembrie 2024

Conf. dr. Rațiu Attila Cristian

Drd. Ionașcu Adrian