Plan van Aanpak

Project Lumbricus terrestris

A shiny fish with a small white bug

Description automatically generated with medium confidence

Project : 'Lumbricus terrestris'

Naam projectleden: Laïsja Thomasia , Alena Savachenko

Opdrachtgever : Chris van Oevelen

Versie : 1.0

**Relevantie van het onderzoeksonderwerp**

Het onderzoeksonderwerp is van groot belang omdat er in de experimentele wetenschap en geneeskunde momenteel wordt gezocht naar een alternatief voor dieren voor het uitvoeren van diverse experimenten in het laboratorium. Omdat diverse wormsoorten een belangrijk model zijn voor medisch en biologisch onderzoek, is het belangrijk om genetisch materiaal van deze soorten te onderzoeken. Dit project richt zich op het onderzoeken van genetisch materiaal van soort *Lumbricus terrestris* (regemworm*)*.

Genoomannotatie is het proces van het identificeren van functionele elementen in een DNA-sequentie. Annotatie geeft betekenis aan het genoom door aan te geven de locatie en functie van genen (coderend voor eiwitten of andere functies), evenals regelgevende elementen. De assemblage is altijd gebaseerd op de reads die zijn gegenereerd tijdens het sequentieproces. De eerste stap in de novo assemblage van het transcriptoom is kwaliteitscontrole kwaliteitscontrole, die het filteren en snoeien van gelezen op niveau van kwaliteitsniveau en het verwijderen van adapters. Degevens kunnen worden gefilterd om alleen die aflezingen of delen van aflezingen te behouden die basen bevatten met de de beste kwaliteitsindex (Q) bevatten. RNA-sequencing omvat niet alle genen in het genoom, sommige genen komen tot expressie op lage niveaus of alleen in een paar weefsels, in bepaalde stadia van ontogenie of in bepaalde omstandigheden, en ze kunnen volledig gemist worden. Daarnaast zijn veel van de transcripten die tot expressie komen niet volledig gesplitste transcripten. Daarom is onafhankelijke verificatie vereist voordat een tot expressie gebrachte regio als een functioneel gen.

De klasse regenwormen zijn fundamentele drijvende krachten achter biomedisch onderzoekecosystemen, en levert een enorme bijdrage aan de landbouw en natuurlijke ecologische processen. Tot op heden bedraagt het aantal soorten in deze groep bijna 7000 (Dr Julia Baum n.d.). Aangezien de problemen met habitatvernietiging, verlies van soorten, klimaatverandering en dramatische veranderingen in de interacties tussen gemeenschappen acuter zijn dan ooit tevoren. Informatie over de structuur en eigenschappen van genomen van regenwormen is cruciaal voor het vinden van oplossingen voor de fundamentele problemen waarmee de mensheid vandaag de dag wordt geconfronteerd. Ringwormen zijn de oudste evolutionaire groep, die voor het eerst een musculatuur hebben die sterk lijkt op de dwarsgestreepte musculatuur van gewervelde dieren. De ringwormen staan vanwege hun evolutionaire positie aan de oorsprong van het fylogenetische dierengeslacht, waarvan de evolutie leidde tot het ontstaan van hoger georganiseerde leden van het dierenrijk. Kennis van het genoommodel en de bijbehorende moleculair genetische processen van ringwormen zal het gebruik van complexere zoogdieren op het gebied van experimentele wetenschap, met name voor het testen van nieuwe geneesmiddelen, verminderen.

Genoomannotatie blijft een arbeidsintensief proces dat meerdere soorten sequentieanalyses combineert. Gezien de grootte en complexiteit van genomen, is de eerste stap naar volledige genoomassemblage meestal het verkrijgen van sequencinggegevens om ruwe assemblage en voorspelling van genmodellen te verkrijgen.

Het hele annotatieproces bestaat over het algemeen uit de volgende stappen:

1) Het maskeren van sterk repetitieve elementen in de genoomsequentie

2) het gebruik van transcripten en eiwitten van dezelfde of verwante soorten om ab initio te voorspellen

3) gebruik van genzoekalgoritmen om mogelijke genstructuren te identificeren;

4) het combineren van deze gegevens om een eerste reeks genmodellen te creëren;

5) filteren van de resultaten op kwaliteit om de meest waarschijnlijke genmodellen te identificeren die volledige eiwitcoderende regio's.(ncbi.nlm.nih.gov n.d.)

**Een overzicht van publiek beschikbare genomen en annotaties in de soort Lumbricus terrestris**

Er zijn echter maar een paar artikelen in de literatuur die de assemblage en annotatie van het Lumbricus terrestris genoom onderzoeken. Voor het eerst wordt een gedetailleerde genoomassemblage van de genen van een soort Lumbricus terrestris gepubliceerd op 30 oktober 2023(Blaxter, Spurgeon, and Kille 2023). Door de technologieën voor long-read sequencing van Pacific Biosciences hebben de wetenschappers het genoom gesequenced en gepubliceerd. Het gepubliceerde genoom biedt, hoewel het volledig op het niveau van de sequentie is, een uiterst gefragmenteerde analyse van de annotaties. Deze genoomassemblage is de eerste publiekelijk openbaar beschikbare genoom assemblage. Bij dit project ligt de nadruk op het verder onderzoeken van de metadata van genoomassemblages. Tot nu toe zijn er in de literatuur geen uitgebreide en systematische onderzoeken gedaan naar de annotatie van het Lumbricus terrestris genoom.

Doelstellingen van het onderzoek

* het verkrijgen van een annotatie van het volledige genoomannotatie van Lumbricus terrestris en de verbetering ervan met behulp van genoomvoorspelling van transcriptiestartsites
* structurele en functionele genoomanalyse Lumbricus terrestris

- analyse van regulerende relaties tussen genoomelementen Lumbricus terrestris.

Op basis van het gestelde doel werden de volgende doelstellingen geformuleerd:

1. het analyseren van genoommodel en de de relatieve inhoud van sterk repetitieve

Elementen in het genoom van de Lumbricus terrestris. (Augustus soft)

2. het uitvoeren van structurele en functionele annotatie van genen voor de

Lumbricus terrestris en vergelijken met de beschikbare annotaties (BLAST,GO)

3. Retraining Augustus

Pipline:

Fasta-file ref. genoom -> augustus, output gff3 -> functionele annotatie Blast (input: eiwitsequentie augustus)- > visualisatie (ref. genoom, ref. trasncripts)

References

Blaxter, Mark L., David Spurgeon, and Peter Kille. 2023. ‘The Genome Sequence of the Common Earthworm, Lumbricus Terrestris (Linnaeus, 1758)’. *Wellcome Open Research* 8:500. doi: 10.12688/wellcomeopenres.20178.1.

Dr Julia Baum. n.d. ‘Ever Thought about Earthworms?’ *African Wildlife Economy Institute*. Retrieved 25 November 2024 (https://www0.sun.ac.za/awei/articles/ever-thought-about-earthworms).

ncbi.nlm.nih.gov. n.d. ‘The NCBI Eukaryotic Genome Annotation Pipeline’. Retrieved 25 November 2024 (https://www.ncbi.nlm.nih.gov/refseq/annotation\_euk/process/).