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 : 1.2

**Relevantie van het onderzoeksonderwerp**

Dit onderzoeksonderwerp is van wezenlijk belang, omdat er in de experimentele wetenschap en de geneeskunde op dit moment wordt gestreefd naar alternatieven voor het gebruik van dieren bij diverse laboratoriumexperimenten. Vanwege de belangrijke rol die diverse wormsoorten spelen als modellen in medisch en biologisch onderzoek, is het essentieel om hun genetisch materiaal te bestuderen. Dit project richt zich specifiek op de analyse van het genetisch materiaal van de soort Lumbricidae.(Leung et al. 2008)

Genoomannotatie is het proces dat gericht is op het identificeren van functionele componenten binnen een DNA-sequentie. Dit proces van annotatie biedt inzicht in het genoom door de plaats en functie van genen te specificeren, waaronder genen die eiwitten coderen of andere functies vervullen, evenals de bijbehorende regulerende elementen. De assemblage is altijd gebaseerd op de reads die zijn gegenereerd tijdens het sequentieproces. Het proces van genoomassemblage houdt in dat het originele genoom wordt gereconstrueerd uit kleine stukjes DNA, verkregen door middel van sequenting. Deze reads zorgen ervoor dat het oorspronkelijke genoom meestal meerdere keren wordt gedekt. Bij het analyseren van genomische en metagenomische gegevens, is de gebruikelijke oplossing een verzameling contigs. Een contig is een aaneengeschakelde nucleotidesequentie. Deze contigs kunnen worden samengevoegd tot scaffolds, waarbij scaffolds bestaan uit een reeks contigs met een schatting van de afstanden tussen deze sequenties.(Anon n.d.)

De soort Lumbricus Terresteris hoort tot de fylogenetische familie Annelida, Clitellata, Oligochaeta, Crassiclitellata, Lumbricina, Lumbricidae(Blaxter, Spurgeon, and Kille 2023). De soort ringwormen (Annelida) zijn de oudste evolutionaire groep. De musculatuur lijkt hier sterk op de dwarsgestreptte musculatuur van dieren. De soorten ringwormen (Annelida) zijn de oudste evolutionaire groep. In dit geval lijkt de musculatuur sterk op dwarsgestreepte spieren bij dieren(Giovanni Pilato n.d.) Daarom is deze soort een van de mensrelevante modellen voor laboratoriumonderzoek.

Het annoteren van genoom is nog steeds een proces dat veel tijd kost en verschillende soorten sequentieanalyses samenbrengt. 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. De bekende transcripten en eiwitten zijn opgeslagen in genetische databases zoals NCBI en BLAST.

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.) . In eerste instantie wordt de kwaliteitsselectie uitgevoerd met behulp van een set positieve controles voor het programma. Na het voltooien van deze controles wordt het percentage fout-positieven resuultaten duidelijk zichtbaar.

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

Hoewel er in de bestaande literatuur slechts een beperkt aantal studies is dat de assemblage en annotatie van het genoom van Lumbricus terrestris behandelt, zijn er wel talrijke beschrijvingen van de genoomannotatie van andere organismen. Voor de soorten *C.Elegans* en *Drosophila melanogaster* is er bijvoorbeeld een volledige annotatie. Voor het eerst wordt een gedetailleerde genoomassemblage van de genen van een soort Lumbricus terrestris gepubliceerd op 30 oktober 2023(Blaxter et al. 2023). Door de innovatieve long-read sequencing methoden van Pacific Biosciences hebben de wetenschappers het genoom gesequenced en gepubliceerd. Hoewel het gepubliceerde genoom compleet is op sequentieniveau, is de analyse van de annotaties erg fragmentarisch. Deze genoomassemblage is de eerste die voor het publiek beschikbaar is. Het project legt de focus op het verder onderzoeken van de metadata van genoomassemblages. Tot nu toe zijn er in de literatuur geen uitgebreide en systematische studies gedaan over de annotatie van het Lumbricus terrestris genoom.

Doelstellingen van het onderzoek

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

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

1. het analyseren van genoommodel 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 statisstisch model

Workflows:

References

Anon. n.d. ‘De Novo Assembly Tutorial’. Retrieved 30 November 2024 (https://colauttilab.github.io/NGS/deNovoTutorial.html).

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.

Giovanni Pilato. n.d. ‘The significance of musculature in the origin of the Annelida’. Retrieved 30 November 2024 (http://ouci.dntb.gov.ua/en/works/ldperODl/).

Leung, Maxwell C. K., Phillip L. Williams, Alexandre Benedetto, Catherine Au, Kirsten J. Helmcke, Michael Aschner, and Joel N. Meyer. 2008. ‘Caenorhabditis Elegans: An Emerging Model in Biomedical and Environmental Toxicology’. *Toxicological Sciences* 106(1):5–28. doi: 10.1093/toxsci/kfn121.

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/).