

Iron deficiency triggered transcriptome changes in bread wheat

by Diemer Harbers and Douwe Metselaar

Abstract IJzer is een stof die in een aantal processen bij planten belangrijke rollen vervult. Het systeem hier achter is bij *Triticum aestivum* (bread wheat) nog vrij onbekend. RNA-seq kan hier verandering in brengen door het transcriptoom te bestuderen. In het onderzoek van Wang et al. (2020) werd gekeken naar DEG's en is een totaal van 5285 DEG's gevonden in de leaves groep. Daarnaast werden 1906 DEG's gevonden in de roots groep. De beide groepen overlappen 684 DEGs. In dit onderzoek is het onderzoek gereproduceerd en is ook DESeq2 gebruikt als normalisatie methode. Om de resultaten te visualiseren zijn volcano-plots gemaakt en een Venn-Diagram gegenereerd. Het resultaat van dit onderzoek was vergelijkbaar met het onderzoek van Wang et al. (2020). Er werden 5506 DEG's in de leaf groep waargenomen, 684 in de overlappende groepen en 1986 in de root groep. De resultaten hadden gedeeltelijke overeenkomst, mogelijk kunnen de verschillen zijn ontstaan door bijvoorbeeld een andere versie van DESeq2.

Introductie

Een tekort aan ijzer kan de gewasopbrengst sterk beïnvloeden. Daarnaast speelt ijzer een belangrijke rol in het metabolisme en is het nodig voor plantengroei. IJzer fungeert onder andere als cofactor van veel enzymen. Verder is het betrokken bij de elektronentransportketen en fotosynthese. Echter is het achterliggend moleculaire mechanisme betreffende ijzertekorten niet bekend voor tarwe. Met RNA-seq zal geprobeerd worden de invloed van ijzertekorten op veranderingen in het transcriptoom te achterhalen. In bladeren en wortels van de gebruikte cultivar (*Triticum aestivum* cv. Bobwhite S26) is onderzocht wat het effect van ijzertekort op het transcriptoom is. Daarbij zijn twee condities met elkaar vergeleken: ijzerarme - tegen een ijzerrijke omgeving. Dit resulteerde in vier afzonderlijke groepen. Vervolgens zullen tussen deze groepen genen geïdentificeerd worden met een differentiële expressie. Tijdens deze analyse zal gestreefd worden naar het reproduceren van het onderzoek van Wang et al. (2020).

Materiaal en Methode

De gebruikte dataset is afkomstig uit de GEO database (accession: GSE162027) van NCBI uit het artikel Iron deficiency triggered transcriptome changes in bread wheat geschreven door Wang et al. (2020). De dataset bestaat uit een Komma-gescheiden bestand waarin een countmatrix is opgenomen. In het bestand zijn twee groepen opgenomen, de Leaf groep en de Root groep. De groepen bestaan uit beide zes samples, drie treatment en drie controle samples. De ruwe reads waren afkomstig van *Triticum aestivum* (bread wheat) gesequenced met het platform Illumina HiSeq 2000. De data is op kwaliteit gecontroleerd en daarna is een differentiële genexpressie analyse uitgevoerd met behulp van het DESeq2 package. De genen die een adjusted p-value kleiner dan 0,05 hadden, en de threshold groter dan een absolute log₂ fold change van 1 passeerden werden gezien als DEG, oftewel een differentieel tot expressie gebracht gen. De complete analyse van de kwaliteitscontrole, de DESeq2 analyse en het visualiseren van de resultaten tot het maken van een volcano-plot en Venn-diagram is uitgevoerd in RStudio (versie bioinf.nl/rstudio) met R (versie bioinf.nl/rstudio) en verschillende R-plotpakketten en analysepakketten zoals ggplot2, DESeq2 en VennDiagram.

Resultaten

Om een duidelijke visualisatie te hebben tussen de control tegen treatment van de roots en leaves is besloten om twee volcano plots te maken, die worden weergegeven in Figure 1. De genen die differentieel tot expressie komen laten een turquoise kleur zien, wanneer deze genen ook nog significant verschillen worden deze genen hoger in de volcano plot weergegeven. De genen die niet differentieel tot expressie zijn gekomen zijn rood gekleurd.

Aan de volcano plots is te zien dat er verschillen waarneembaar zijn tussen de leaves en roots, zowel bij een lage als een hoge foldchange. Op basis van de plots is moeilijk vast te stellen hoeveel genen dit betreft, dit komt omdat de dichtheid van de punten op sommige plekken in de volcano-plot hoog is. Hiervoor zal een andere visualisatie uitsluiting moeten geven. Dit is gedaan in de vorm van een Venn-Diagram.

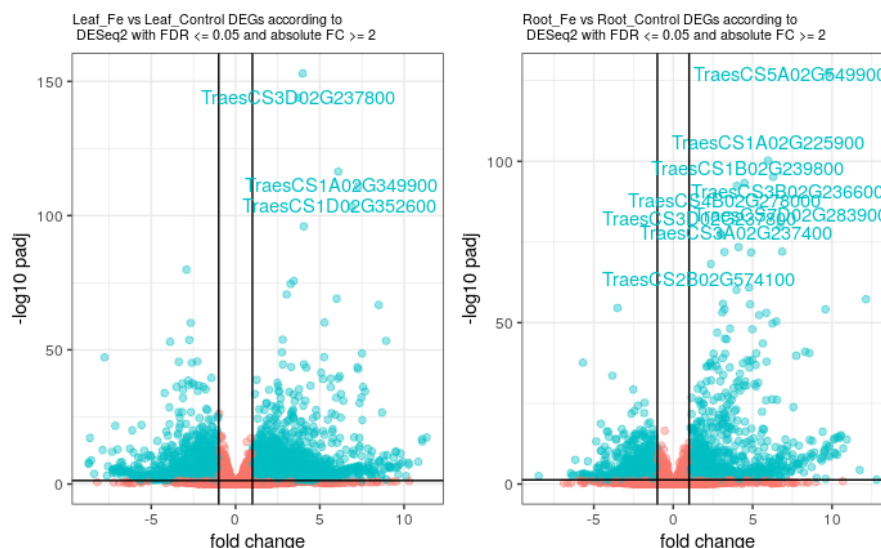


Figure 1: Volcano plots van leaf treatment tegen leaf control en root treatment tegen root control. Op de x-as is de log2 fold change uitgezet tegen de adjusted p-value op de y-as voor elke gen. Genen die up gereguleerd of down gereguleerd zijn, en dus een log2 fold change hoger dan 1 hebben, en een adjusted p-value hoger dan 0,05 worden als groene bolletjes weergegeven. De rode bolletjes zijn genen die niet up of down gereguleerd zijn of een adjusted p-value lager dan 0,05 hebben.

De Venn-Diagram kan uitsluitel geven over het aantal genen dat up en down gereguleerd is in de beide groepen. Daarnaast is het ook interessant om te kijken naar de overlap tussen de beide groepen. In de Venn-Diagram in Figure 2 is het totaal aantal genen die differentieel tot expressie komen in de leaves en roots weergegeven, daarnaast zijn de resultaten uitgesplitst per groep.

In het midden van de Venn-Diagram is de overlap uitgesplitst in een tabel. De leaf groep bestaat in totaal uit 5506 DEG's, 3838 genen zijn up-gereguleerd en 1168 down-gereguleerd. In de root groep is een totaal van 1986 DEG's gevonden, waarvan 840 genen up-gereguleerd zijn en 1146 down-gereguleerd. De 684 overlappende DEG's zijn overeenkomstig tussen beide groepen.

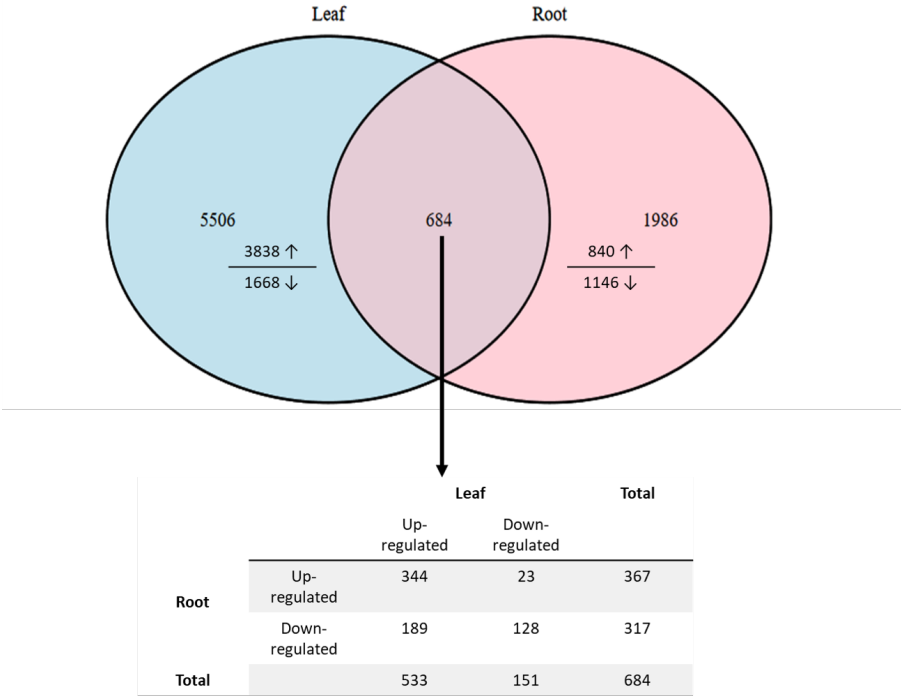


Figure 2: Venn-Diagram van DEG's die gevonden zijn in de dataset geanalyseerd doormiddel van DESeq2 in de groepen leaves en roots. De linker cirkel geeft de DEG's weer van de leaf groep, in de cirkel staan het aantal up en down gereguleerde genen. De rechter cirkel geeft de DEG's weer van de root groep. De overlap is het aantal DEG's wat in beide groepen up en down gereguleerd is. De tabel maakt onderscheid tussen de overlap tussen beide groepen.

Conclusie

Om het effect van ijzertekort op het transcriptoom te onderzoeken is een RNA-sequencing analyse opgezet. Hiervoor zijn twee groepen tarwe onder verschillende concentraties ijzer opgegroeid: een ijzerarme – en een ijzerrijke omgeving. Uit twee weefsels is vervolgens RNA geëxtraheerd en gesequenced. Een count matrix, met de hoeveelheden reads die tegen een gen gemapt zijn, is in R geladen. Vervolgens is met DESeq2 een differentiële genexpressieanalyse uitgevoerd. In totaal zijn 8176 genen als DEG bestempeld. Deze genen passeerden de log2 fold change en adjusted p-value threshold. 684 genen zijn gevonden die in beide weefsels differentieel tot expressie komen.

Discussie

Het doel van deze studie is het verifiëren van de bevindingen van Wang et al. (2020). De auteurs zijn tot een totaal van 7876 DEG's gekomen. In bladeren hebben de auteurs 3694 en 1591 omhoog - en omlaag gereguleerde genen gevonden, respectievelijk, in ons geval zijn er 3838 en 1668 DEG's in hetzelfde weefsel gevonden. Voor de wortelen hebben de auteurs 820 en 1087 DEG's gevonden, in vergelijking tot 840 en 1146 omhoog – en omlaag gereguleerde DEG's in deze analyse. Het aantal gemeenschappelijke DEG's tussen bladeren en wortels zijn in beide studies gelijk: 684. Verschillen tussen deze twee studies kunnen wellicht verklaard worden aan de hand van het gebruik van verschillende versies van DESeq2. In vergelijking tot het originele onderzoek is er geen Gene ontology (GO) enrichment analyse uitgevoerd. Om de differentiële expressie van genen goed te kunnen linken aan pathways en te kijken welke pathways worden verrijkt onder ijzerarme condities, zou deze analyse in de toekomst uitgevoerd kunnen worden.

Bibliography

M. Wang, J. Gong, and N. K. Bhullar. Iron deficiency triggered transcriptome changes in bread wheat. *Computational and Structural Biotechnology Journal*, (18):2709–2722, 2020. doi: <https://doi.org/10.1016/j.csbj.2020.09.009>. [p1, 3]

Diemer Harbers
Hanze University of Applied Sciences
Zernikeplein 11

`d.w.j.harbers@st.hanze.nl`

Douwwe Metselaar
Hanze University of Applied Sciences
Zernikeplein 11

`d.g.metselaar@st.hanze.nl`