

IDENTIFYING GENOME EDITING TARGETS: Using Metabolism to Rationally Mine Genomics Data

Frank Skraly, Madana M. R. Ambavaram, Venkatesh Bollina, Maria Somleva, Meghna Malik, Jihong Tang, Oliver Peoples, **Kristi D. Snell**
 Metabolix Oilseeds, Inc., Saskatoon, Saskatchewan, Canada, Yield10 Bioscience, Inc., Woburn, MA *Contact info: snell@yield10bio.com

Yield10 develops food and feed crops to produce higher yields with lower inputs of land, water or fertilizer

INTRODUCTION

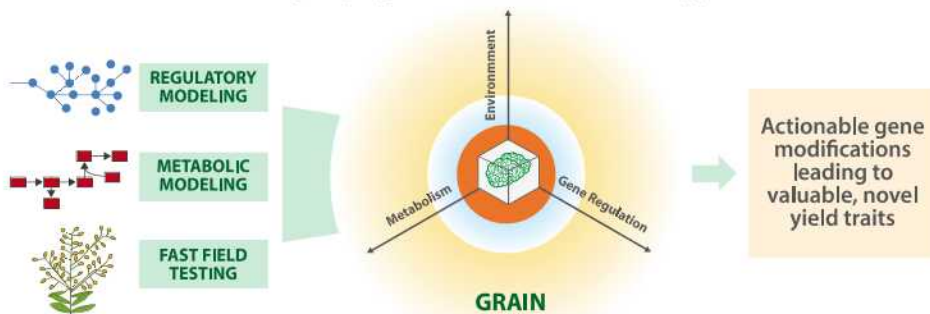
Developing technologies to achieve step change increases in crop yield remains a critical unmet need in agriculture and a key challenge for future global food security. Yield10 is focused on achieving step change increases in the inherent yield of major food and feed crops. *This will likely require multiple gene modifications to increase photosynthesis and efficiently deliver the increased photosynthate to the desired organ, either harvestable seed or biomass depending on the crop.* On some level, a plant phenotype or trait reflects altered metabolism due to genetic variation. Yield10 is leveraging our unique expertise in metabolic engineering to identify genes to achieve traits such as increased crop yield by using modeled metabolic changes to inform the mining of large genomics data sets. This trait agnostic approach provides a unique and powerful method to identify smart gene targets.

Increased Photosynthesis → Plant Central Metabolism → Increased Harvest Index

What genes or gene combinations should be edited to deliver the next generation of agricultural traits?

GRAIN 3.0 (Gene Ranking Artificial Intelligence Network)

Yield10 Bioscience has developed the GRAIN Platform to identify unique gene combinations to increase crop yield.



Each modeling component can produce its own set of gene targets. Combining modeling components provides a unique tool where metabolic information can inform genomics decisions.

METABOLIC MODELING

The metabolic modeling component of GRAIN, based on flux balance analysis, thermodynamics, and kinetics, allows the examination of plant metabolism to determine the importance of individual genes to a trait and to determine theoretical yields upon gene modifications.

Model: Flux Balance Analysis

All reactions in a system

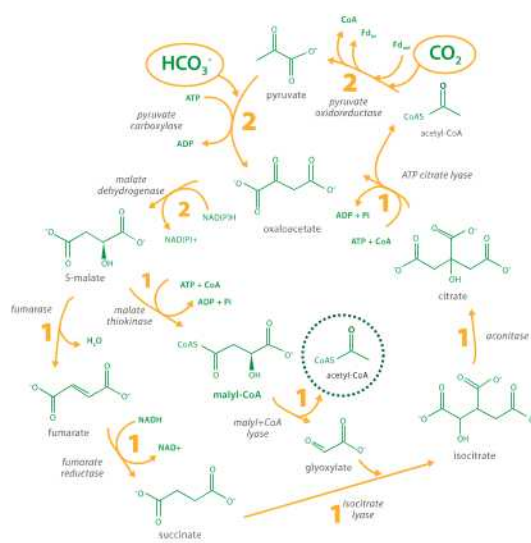
- Enzymatic reactions, transport functions, inputs/output, cofactors (i.e. ATP, ADP, NADPH...)
- Avoid thermodynamically unfavorable solutions

Define problem (directive to minimize or maximize certain flux(es)) e.g. maximize biomass production

Model output

- Sets of ideal fluxes, theoretical yields
- Importance of individual genes to trait
- Pathway comparisons or optimizations
- Alternate pathways to achieve goal

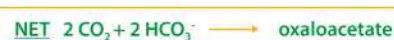
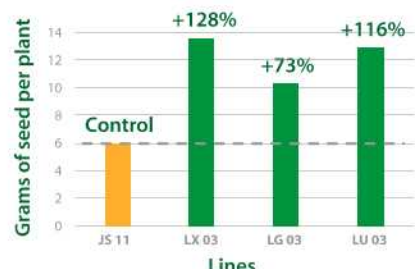
METABOLIC MODELING VALIDATION: PROOF POINTS ACHIEVED



Overexpression of novel carbon fixation pathway in *Camelina sativa* increases seed yield

- Yield10 modeling predicts a **112%** increase in maximum theoretical seed yield with new pathway under photorespiratory conditions
- Experimental results show increase up to **128%**

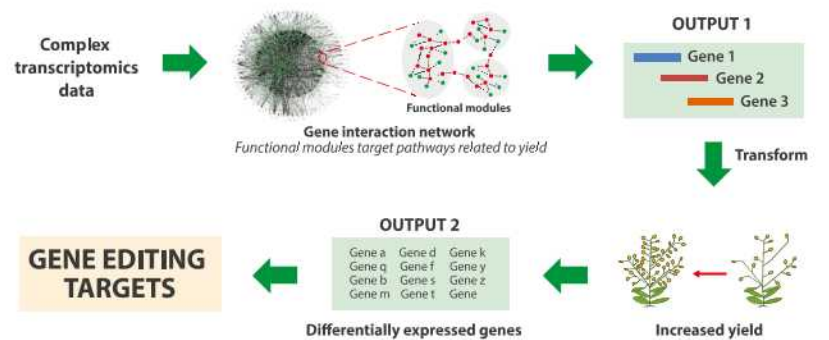
Camelina greenhouse study: Seed yield in best plants



References. Yield10 review article on metabolic engineering: Skraly et al., Metabolic engineering to increase crop yield: From concept to execution, Plant Science, 2018. Novel carbon fixation pathway description: Malik et al., Camelina sativa, an oilseed at the nexus between model system and commercial crop, Plant Cell Rep., 2018.

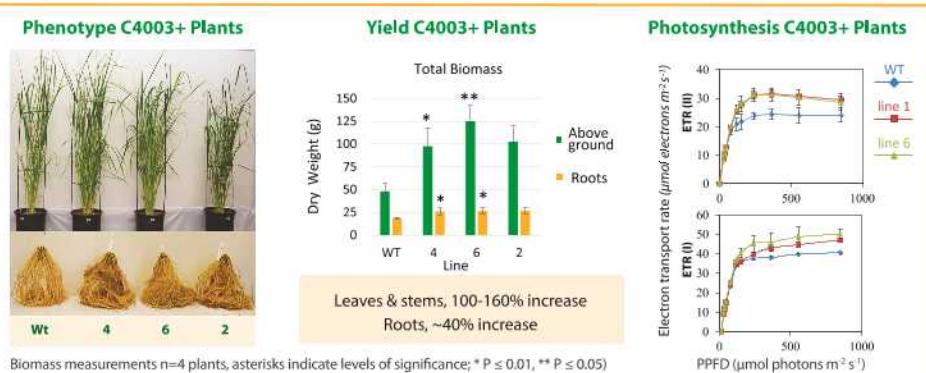
REGULATORY MODELING

The regulatory modeling component of GRAIN, based on transcriptome-based regulatory association networks, converts large amounts of data into actionable gene targets.



REGULATORY MODELING VALIDATION: PROOF POINTS ACHIEVED

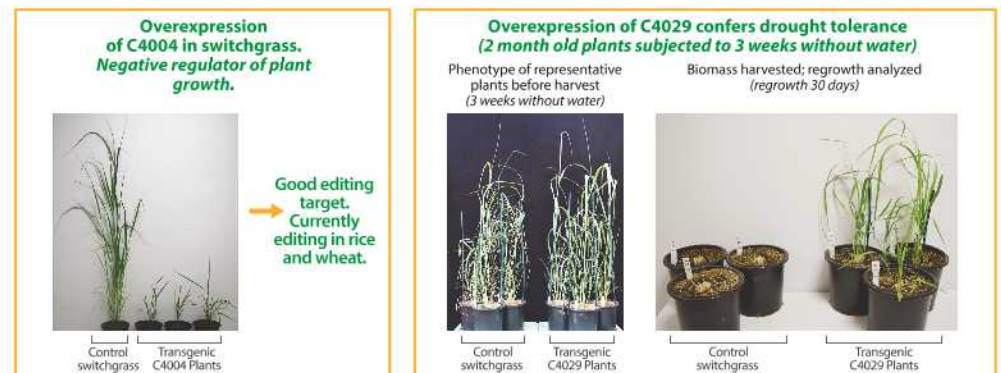
3 Global Regulatory Genes Identified in Switchgrass, Highlight C4003 trait



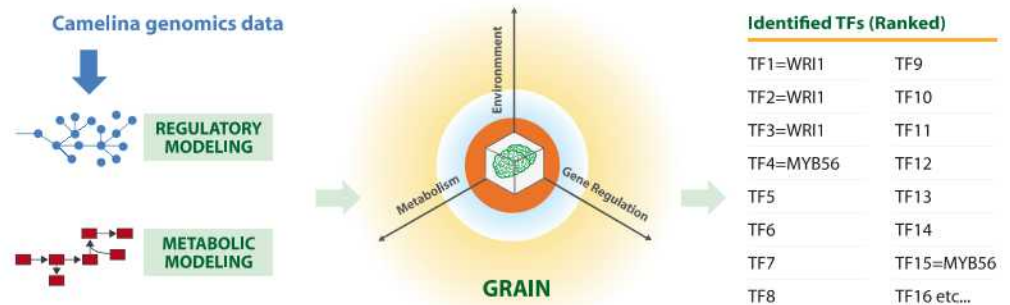
In switchgrass C4003+ plants, biomass yield increased up to 160%, root yield increased up to 40%. Primary photosynthesis difference observed in electron transport rate of photosystem I and II. Experiments in wheat and rice ongoing, early results are promising.

Reference: Ambavaram et al., Novel transcription factors PvBMY1 and PvBMY3 increase biomass yield in greenhouse-grown switchgrass, Plant Science, 2018.

Transcriptomics of High Yielding Switchgrass Plants: Powerful Resource for Gene Discovery



GRAIN 3.0: COMBINED MODELING COMPONENTS TO IDENTIFY TRANSCRIPTION FACTORS (TFs) TO INCREASE OIL CONTENT AND/OR SEED YIELD IN CAMELINA



Identified TFs (Ranked)	
TF1=WRI1	TF9
TF2=WRI1	TF10
TF3=WRI1	TF11
TF4=MYB56	TF12
TF5	TF13
TF6	TF14
TF7	TF15=MYB56
TF8	TF16 etc...

- Ranked list of TFs obtained, known TFs that impact oil and/or seed yield (Wrinkled1, MYB56) observed in top spots of list validating approach
- Many uncharacterized genes identified, IP white space

References. Ma et al., WRINKLED1, A ubiquitous regulator in oil accumulating tissues from Arabidopsis embryos to oil palm mesocarp, PLoS One, 2013; Zhang et al., MYB56 encoding a R2R3 MYB transcription factor regulates seed size in Arabidopsis, 2013, J. Integr. Plant Biol.

CONCLUSION

- The GRAIN platform provides a unique method to identify smart gene targets and has been successfully used to identify traits
- Yield10 is validating traits in a range of crops and is collaborating with partners in corn, soybean, wheat, and sorghum
- The GRAIN platform is trait agnostic and can be tailored to a wide range of crops
- Depending on the gene target, genome editing can be used for gene deletion, promoter replacement, gene insertion, etc.