



Date: Sept 2019
Reference: GCFF TN-029

Technical Note

GCFF has produced a forest phenotyping platform! So what?

Author/s: Heidi Dungey, Max Bombrun, David Pont, Natalie Graham, Jonathan Dash, Toby Stovold.

Corresponding author: heidi.dungey@scionresearch.com

Summary:

One of the GCFF programme's aims was to combine genetics and remote sensing to create a forest phenotyping platform. So how have we done?

A forest phenotyping platform has been produced that uses a big-data machine-learning approach, and this has been used to model forest productivity across two large forests. This platform has been shown to be effective for:

- 1) determining the main drivers of forest productivity across two commercial forests;
- 2) finding outstanding individual trees in the forest and;
- 3) identifying the pedigree (mums and dads) of outstanding trees using DNA-based technologies, informing future deployment and possibly breeding outcomes.

The work has also demonstrated that methods developed for individual-tree delineation can be applied to understand how a tree's growth can be affected by local micro-environmental effects and by local competition effects. The forest phenotyping platform will be invaluable for capturing real-time effective inventory data, and for understanding the value of forest management interventions and the choice of planting stock in determining forest productivity and profitability. Huge potential also exists to adapt the methods developed here for application into tree breeding programmes.

Introduction

At the beginning of the Growing Confidence in Forestry's Future (GCFF) programme in 2013, a 'phenotyping platform' was proposed. The concept was to use remotely sensed data to describe the productivity of the forest at the stand level, and also where possible, for individual trees.

There have been three major outcomes:

1. Quantification of the key drivers of productivity at the stand level over whole forests (Box 1).
2. Individual-tree phenotyping for height works well, to a precision that is perhaps even better than the much slower ground-based measurement (Box 2).
3. A demonstrated ability to reconstruct the pedigree of at least some of the trees in a stand using DNA markers (Box 3).

The phenotyping platform can be used as a tool to assist tree breeders to help maximise productivity and profitability of future forests.

What is phenotyping?

Phenotyping is the accurate and precise physical description of trees, traditionally done using tools such as diameter tapes, sonic velocity tools and through the taking and measuring of wood cores.

In forest trees, phenotyping with high precision has been difficult because trees and forests are large, long-lived and highly variable.



Box 1: Whole-of-forest phenotyping



Figure 1: Whole of forest phenotyping has been undertaken for multiple forests in the North Island of New Zealand to quantify forest productivity and identify outstanding trees.

Step 1: Data

Data was collated across all forest stands, including management (e.g., spacing, thinning), genetics (e.g., GFPlus rating), climate and LiDAR-derived topographical information.

Step 2: Modelling across-forest productivity

Intensive modelling in one forest showed that genetic improvement rating (SeedlotCod) and spring temperature (spring.temp) have the greatest impact on stand productivity. Genetics was also important in a second forest.

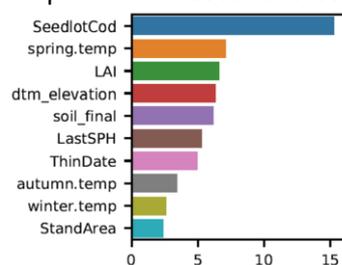


Figure 2: Intensive modelling of main factors contributing to observed variation in site index.

Step 3: Identifying outstanding trees and locating them on the ground

Outstanding trees were located by modelling competition, environment and genetic effects across several stands. On-ground location of trees was undertaken using GPS and a tablet with a pre-loaded individual-tree map of the site.

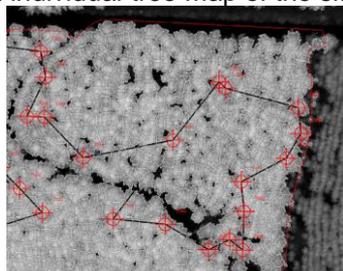


Figure 3: Individual trees were selected for on-ground inspection. Needles or cambium were collected for DNA extraction for parentage analysis.

Step 4: Genotyping to identify parents

DNA was extracted and sent for genotyping. Parentage is confirmed using DNA-marker patterns. (See Box 3).

Box 2: Individual tree phenotyping

Step 1: Finding a tree in the forest

Methods were developed to find individual trees, define the shape of the crown and derive growth and form measurements.

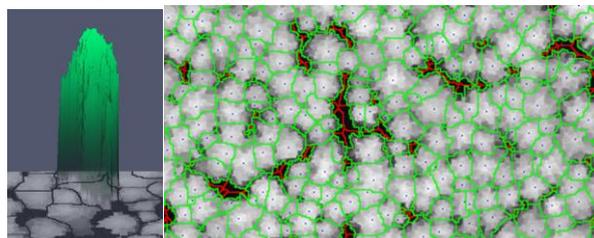


Figure 4: Individual trees can have their canopies delineated (right) and characterised (left) allowing estimation of correlated growth and form traits.

Step 2: Individual-tree LiDAR-derived data is useful

LiDAR-derived individual-tree data were tested for their suitability for use in tree breeding analyses. The best results were obtained for height (H), diameter at breast height (DBH) and stem volume (V). LiDAR-derived height measurements for individual trees were suitable for use in tree breeding.

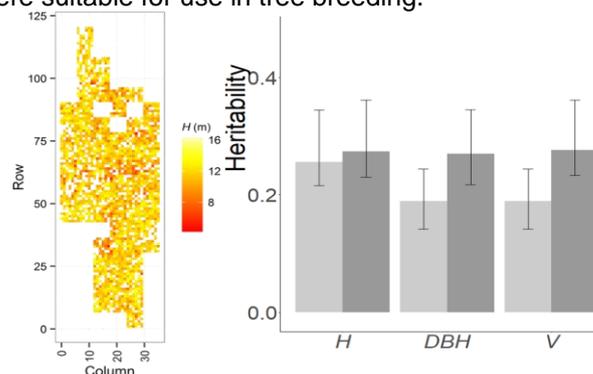


Figure 5: Individual trees in a genetics trial (left) showing trial layout and individual-tree height (yellow = tall; red = short) and heritability estimates (right) from LiDAR-derived traits (right; light grey) and on-ground measurements (dark grey).

Step 3: Delineating individual trees allowed competition to be quantified

Once trees are delineated, their crown dimensions and stocking can be used to understand competition. Accounting for competition effects significantly reduced error (residual) variances when included in the model used in tree breeding analyses. This, in turn, improved the heritabilities estimated.

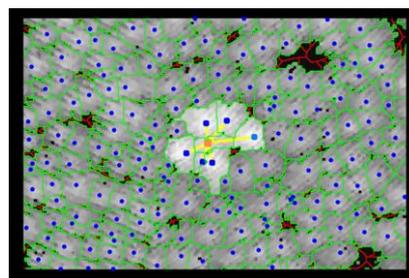


Figure 5: Individual trees in a stand showing key tree relationships and distances used to quantify competition.

Box 3: Reconstructing the pedigree

Step 1: Locate the tree

An individual tree was detected using LiDAR and located on the ground (right).

Step 2: Extraction of DNA

Either fresh pine needles or cambium was obtained from candidate trees, and taken to the laboratory for DNA extraction. Once extracted, DNA was sent for genotyping.

DNA extractions were performed using ~100 mg of tissue and the NucleoSpin® Plant II (Machery-Nagel, Düren, GER) kit, as per the manufacturer's instructions, with the modifications as described in [2]. All DNA was frozen and stored at -20 °C until shipping to Rapid Genomics in Gainesville, Florida.

Step 3: Genotyping

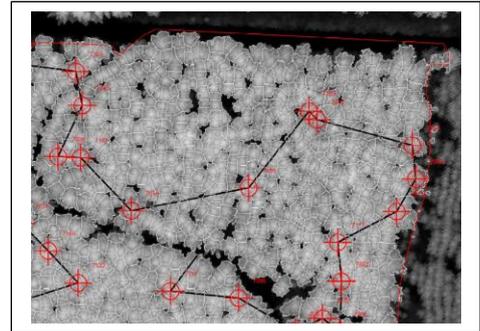
Rapid Genomics performed the genotyping using a method called exome capture genotyping by sequencing. This method can provide a lot of markers (SNPs), but can also have some genotyping errors.

Step 4: Parentage analysis

Assignment of parentage based on DNA marker profiles was performed using two different methods. Parents assigned using DNA were compared with known and/or expected parents of individual trees (right). In the majority of cases, parents were correctly assigned. In some cases, however, pedigree reconstruction identified mismatched parents (documented versus reconstructed pedigrees) or no assignment was possible.

Pedigree reconstruction was demonstrated to be possible, but with caveats. The method works the most effectively where parents are already genotyped and where there is low genotyping error.

The Radiata Pine Breeding Company and Scion have now designed and tested a Radiata Pine SNP Chip, a direct outcome of the Genomic Selection Programme. This SNP Chip should provide a better platform for pedigree reconstruction, with fewer genotyping errors. This pedigree reconstruction platform should be available within the 2020 calendar year.



Documented		Reconstructed	
female parent	male parent	female parent	male parent
A	H	A	H
B	I	B	I
C	J	C	J
D	H	D	H
B	I	H	A
E	K	L	K
F	L	F	J
A	H	A	H
D	H	B	H
A	H	A	H
G	M	G	H
G	M	G	C
G	M	G	B
G	M	G	C
G	M	G	H
N	O	B	H

Colour code:

- Green – match to documented pedigree
- Orange – mismatch to documented pedigree but match between two pedigree reconstruction methods tested
- Red- no assignment or mismatch
- Blank – parent not on file

Previous work under workstream 2.1b of the GCFP programme demonstrated the successful recreation of parentage using DNA in a research trial, FR10/0. Many operational stands have incomplete seedlot records. DNA testing provides a means by which this information could be retrospectively determined. Furthermore, the ability to recreate pedigrees through DNA-based parentage assignments would allow the performance of seedlots to be teased apart further, by identifying which genotypes have contributed towards final stand composition. By identifying the best (or worst) performers, this would indicate which genotypes might be the best (or least recommended) for certain sites.

So what?

The biggest questions from this research might be “So what does it mean to me?” and “How might I use it in my forests?”

Use the modelling approach to quantify areas in my forest management that could be tweaked to improve forest productivity

As a concept, modelling the components that contribute to stand-level productivity have been shown to be effective to identify the key drivers of productivity in at least two large forests of radiata pine. While this research has used a LiDAR-derived summary of forest site index, the modelling concept can be used across any forest. All that is required is the collation of a good set of data.

Identification of productivity drivers are advantageous to forest managers as these drivers can then be managed to improve productivity and profitability. For example, data-driven selection of the right genetics for the worst performing areas of the forests, or discontinuing with the worst performing seedlot, would lift average forest productivity.

Why would I care about determining the location of outstanding individual trees?

Candidate trees that are able to be identified from remotely-sensed data are likely to be extremely useful in the context of a major incursion of a new pest or disease. Identification of trees that are green in severely affected stands can be undertaken remotely, and then tracked down in the forest. Using pedigree reconstruction, the parents can be identified. The same family can then be re-created for breeding and/or deployment purposes.

Who cares about competition effects?

Competition among trees in the forest is normal and natural. For the first time, remotely sensed data has allowed us to examine the extent of competition within a stand and within a genetics trial.

In a tree breeding context, an ability to quantify competition means that tree breeders will be able to capture and quantify more genetic gain.

For forest growers, individual-tree competition effects can help forest managers maximise stand growth and the effectiveness of thinning regimes.

The GCFF programme has allowed us to make a step-change in forest management.

Since the beginning of GCFF, forest management has changed for many, with LiDAR and photogrammetry data capture becoming more and more common. Phenotyping the forest is the beginning of a new big data era for the forest industry and a more data-informed method of management.

Acknowledgements

Funding for this research came from the “Growing Confidence in Forestry’s Future” research programme (C04X1306), which is jointly funded by the Ministry of Business Information and Employment (MBIE) and the Forest Growers Levy Trust, with the support of the NZ Forest Owners Association (FOA) and the NZ Farm Forestry Association (FFA). Development of the exome capture probe panel was funded by the RPBC’s MBIE Genomic Selection Partnership programme (RPBC1301).

Suggested reading

- Dungey, H.S., Dash, J.P., Pont, D., Clinton, P.W., Watt, M.S. and Telfer, E.J. (2018) Phenotyping Whole Forests Will Help to Track Genetic Performance. *Trends in Plant Science*, xx, 1–11.
- Gordon, A.D. and Pont, D. (2015) Inventory estimates of stem volume using nine sampling methods in thinned *Pinus radiata* stands, New Zealand. *New Zealand Journal of Forestry Science*, 45.
- Graham, N., A. Ismael, and T. Stovold, Reconstructing parentage in *Pinus radiata* using exome capture genotyping proof of concept in FR10/0. 2018.
- Graham, N, Ismael, A, and Stovold, T. (2019) DNA fingerprinting to reconstruct parentage from trees in the forest – a second proof of concept. GCFF TN-024
- Kalinowski, S.T., M.L. Taper, and T.C. Marshall, Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 2007. **16**(5): p. 1099–1106.
- Melo, A.T.O. and I. Hale, ‘apparent’: a simple and flexible R package for accurate SNP-based parentage analysis in the absence of guiding information. *BMC Bioinformatics*, 2019. **20**(1): p. 108.
- Pont, D., Kimberley, M., Brownlie, R., Morgenroth, J. and Watt, M.S. (2015) Tree counts from airborne LiDAR. *New Zealand Journal of Forestry*, 60, 38–43.
- Pont, D., Kimberley, M.O., Brownlie, R.K., Sabatia, C.O. and Watt, M.S. (2015) Calibrated tree counting on remotely sensed images of planted forests. *International Journal of Remote Sensing*, 36, 3819–3836.
- Pont, D., Dungey, H., Watt, M.S., Morgenroth, J. and Stovold, G.T. (2016) The use of LiDAR for Phenotyping. *Forest Genetics for Productivity Conference*, Rotorua, New Zealand p. Rotorua, New Zealand. https://www.researchgate.net/publication/304864832_The_use_of_LiDAR_for_Phenotyping
- Telfer, E.J., et al., Extraction of high purity genomic DNA from pine for use in a high-throughput Genotyping Platform. *New Zealand Journal of Forestry Science*, 2013. **43**(3).
- Watt, M.S., Kimberley, M.O., Dash, J.P. and Harrison, D. (2017) Spatial prediction of optimal final stand density for even-aged plantation forests using productivity indices. *Canadian Journal of Forestry Research*, 535, 527–535.