

## DEVELOPMENT OF DNA MARKER SELECTION TOOLS IN AUSTRALIA'S MAJOR PLANTATION EUCALYPTS

Dr Simon Southerton Gondwana Genomics Pty Ltd FWPA Webinar, 9<sup>th</sup> Sept 2015

**CSIRO** 



### **Gondwana Genomics**

Delivers marker-assisted selection services to the eucalypt plantation industry Exclusive license to marker technology developed in concernent over last 10 years Gondwana employs all former CSIRO forestry marker scientists Collaborates with CSIRO in marker research Continually developing new markers for new traits





## Outline

Introduction to marker-assisted selection (MAS) Goals of Blue Gum Genomics project (2010-2014) Major project results Marker validation Application of MAS in breeding programs





## Major benefits of marker-assisted selection (MAS)

### DRAMATICALLY **SHORTEN BREEDING CYCLE** BY SELECTION IN SEEDLINGS **ACCURATE SELECTION** OF BEST **PARENTS** FOR CROSSES **ACCURATE SELECTION** OF ELITE **OFFSPRING INCREASE SELECTION INTENSITY** BY SCREENING 1000'S OF TREES SELECT ON **MULTIPLE TRAITS** AT THE SAME TIME





## Markers are used widely in crops

MAS routinely used in crop and animal breeding Example – MAS helps deliver 1% gain per year (~\$50M) to Aust. wheat industry

#### Marker-selected traits

- Stress tolerance
  - salt tolerance
  - aluminium tolerance
- Disease resistance
  - rust resistance
  - nematode resistance
  - virus resistance
- Flour traits
  - flour quality
  - flour colour
  - gluten strength
  - starchiness







#### However, it has been difficult to identify useful markers in trees





### Marker-assisted selection & tree breeding



- During domestication crops have lost much of their variation
- Trees are highly diverse (undomesticated)

CSIRC

- Large gains still to come from conventional tree breeding
- Markers can accelerate yield gains in trees while maintaining genetic diversity



### **Breeding is targeted at genes**



to build and maintain the tree

**CSIRO** 

**Genes** are short stretches of DNA in the chromosomes



### Markers identify good and bad genes

Each tree has two copies of each gene (or allele), one from each parent Sometimes one of these genes is better than the other For example, one allele may give higher growth



#### Markers can identify better genes for selection during breeding





### Markers can accelerate breeding

#### **TRADITIONAL BREEDING**



#### **MARKER-ASSISTED BREEDING**



**CSIRO** 

#### Markers allow shorter breeding cycles

- No need to measure growth and wood properties
- Breeding is quicker and focused on best trees



## **Capturing gains from MAS**



## **GOALS – BLUE GUM GENOMICS**





## **Blue Gum Genomics Project - Goals**

Discovery of pulp yield markers in 100 genes in *E. nitens*.

Hottest 100, March 2007

Discovery and application of molecular markers in *E. nitens* and *E. globulus* to accelerate and intensify selection for high value wood and growth traits in plantation trees.

BGG September 2010





## **Commercial traits are complex traits**

- Genetics + environment control trait variation
- Traits like KPY, density and growth are complex, and influenced by variation in many dozens of genes

#### Example: KPY

- Probably several thousand genes involved in wood development
- We think variation in roughly 100 genes controls most variation in pulp yield





## Marker discovery in humans

- Association studies began in humans after human genome sequenced
- Now being used widely for discovering SNPs causing diseases (e.g. heart disease, dementia, cancer etc...)





Genome-wide association studies are discovering SNPs controlling many diseases



## **Pioneering research in domestic blue gums**

CSIRO pioneered the use of association studies to discover markers Eucalypts ideally suited to association genetics (high diversity, low linkage disequilibrium)



#### Published

1<sup>st</sup> association study in trees - Thumma et al. 2005 *Genetics* 1<sup>st</sup> functional validation of perfect marker - Thumma et al. 2009 *Genetics* 

Copyright © 2005 by the Genetics Society of America DOI: 10.1534/genetics.105.042028

> Polymorphisms in Cinnamoyl CoA Reductase (CCR) Are Associated With Variation in Microfibril Angle in Eucalyptus spp.

Bala R. Thumma, 0.1.1 Maureen F. Nolan, 0 Robert Evans: and Gavin F. Moran 0.1

\*CSIRO Forestry and Forest Products, Canberna 2600, Australia, <sup>+</sup>CRC for Sustainable Production Forestry, Ganberna 2600, Australia and <sup>+</sup>CSIRO Forestry and Forest Products, Clayton, Victoria 3168, Australia Manuscript received February 20, 2005

Accepted for publication July 5, 2005

Coperight 0 2009 by the Genetics Society of America DOI: 10.11554/arrantes.200.100201

> Identification of a Cis-Acting Regulatory Polymorphism in a Eucalypt COBRA-Like Gene Affecting Cellulose Content

Bala R. Thumma,<sup>+,1</sup> Bronwyn A. Matheson,<sup>+</sup> Deqiang Zhang,<sup>+,2</sup> Christian Meeske,<sup>+</sup> Roger Meder,<sup>+</sup> Geoff M. Downes<sup>+</sup> and Simon G. Southerton<sup>+</sup>

\*Commenuralth Scientific and Industrial Research Organization (CSIRO) Plant Industry, Conferent, Australian Capital Territory 2601, Australia and <sup>1</sup>CSIRO Sciationable Ecosystems, Hohert, Tamensia 2001, Australia Manuscript received June 25, 2000 Accepted for publication August 31, 2009



#### **Perfect markers**

The marker occurs in a gene that influences the trait and it is usually the direct cause of variation in the trait.



## **BGG - Key deliverables**

- **New populations** for future association studies to replace aging provenance trials.
- **Predicted pulp yield, cellulose content and other traits** measured for four populations of *E. nitens* and four populations of *E. globulus* growing in contrasting environments.
- **Solid wood traits** (checking, MFA, density etc) measured in two large populations of *E. nitens* growing in contrasting environments.
- Marker-assisted selection service for selecting *E. nitens* and *E. globulus* trees with substantial improvements in pulp yield, growth and solid wood properties.





## **KEY RESULTS – BLUE GUM GENOMICS**





## **New association populations**

#### HVP

Total of 289 *E. globulus* native seedlots, comprising both Otways (144) and Gippsland (145) provenances planted in adjoining trials in June 2013 on two Gippsland sites.

A trial of *E. nitens* seedlots established spring 2014.

#### **Forestry Tasmania**

*E. nitens* (220 seedlots - 6,600 seedlings) and *E. globulus* (240 seedlots - 7,200 seedlings) planted in north-eastern Tasmania in 2014





## **Candidate genes**

HT transcriptome sequencing of xylem from high and low pulp yield trees

Differentially expressed genes enriched for cytoskeleton & cell wall genes

SNPs selected from among 2000 prioritised genes



3 low pulp bulks

3 high pulp bulks

GENE CATEGORY	TOTAL GENES	UP REGULATED	ENRICHMENT
cytoskeleton organization	92	41	3.70
microtubule-based process	79	35	3.68
cell wall biogenesis	55	28	4.22
carbohydrate metabolic process	507	116	1.89
cell wall organization or biogenesis	131	46	2.91
cellular cell wall organization or biogenesis	88	35	3.30
actin filament-based process	57	26	3.78
plant-type cell wall biogenesis	46	23	4.15
actin cytoskeleton organization	47	23	4.06
plant-cell wall organization or biogenesis	75	30	3.32



Thavamanikumar, Southerton & Thumma (2014) PLoS ONE







CSIRO

## Growth and wood trait data

#### E. nitens

Trial location	Provenance(s) in trial	Families	Heritability												
			DBH	КРҮ	Cellulose	Density									
Meunna (H100)	Cen. Vic.	420	0.34	0.51	0.51	-									
Tarraleah (H100)	Cen. Vic.	150	0.29	0.50	-	0.44									
Loudwater (H100)	Cen. Vic.	140	-	0.39	0.45	0.60									
Florentine	Cen. Vic.	420	na	na	na	na									
Geeveston	Cen. Vic.	125	-	-	-	-									
Mt Worth (2 trials)	<i>Cen. Vic.</i> & Tallaganda	115 & 85	-	0.13	0.05	0.07									

#### E. globulus

Trial location	Provenance(s) in trial	Families	Herital	Heritability								
			DBH	KPY	Cellulose	Density						
Latrobe TAS (2 trials)	Otways & Flinders	150 & 99	-	0.45	0.40	0.35						
West Ridgley TAS	Otways	150	0.1	0.35	0.32	0.49						
Busselton WA	Otways	169	0.14	0.35	0.34	0.41						
Busselton WA	Flinders	104	0.89	0.74	0.71	0.80						
Busselton WA	SE Tasmania	101	0.76	0.62	0.60	0.75						
Busselton WA	Gippsland	86	0.31	0.93	0.86	0.75						
Mt Barker WA	King Island	75	-	-	-	-						





## Solid wood data – E. nitens

Trial location	Provenance(s) in trial	Families	Heritability								
			MFA	MOE	Density						
Tarraleah TAS	Central Victorian	150	0.43	0.95	0.76						
Geeveston TAS	Central Victorian	150	0.3	0.51	0.64						

Silviscan analysis of approximately 500 trees in both trials

Additional 420 trees from 420 families with silviscan data used from Meunna TAS trial (from previous FWPA supported research – Hottest 100)





### Marker discovery using association genetics



- Candidate SNP markers selected from within candidate genes
- SNPs genotyped in four populations of about 500 trees for each species
- Analysis of variance used to identify markers significantly associated with trait
- Meta-analysis used to identify markers that are stable across the four trials





## **Summary of BGG marker results**

Markers	Pulp Yield	Growth	Stiffness (MFA)
E. globulus	62	182	
E. nitens	68	205	97





## **Summary of BGG marker results**

#### KPY

- 60-70 markers for KPY identified in both *E. globulus* and *E. nitens*
- Stable across 4 trials containing approximately 1600 trees
- E. globulus SNPs stable between Otways and Flinders Island races
- Few SNPs (<5%) shared between the two species

#### Growth (DBH)

- About 200 markers for growth identified in both *E. globulus* and *E. nitens*
- Stable across 4 trials containing approximately 1600 trees
- Roughly 20% of SNPs associated with KPY are associated with growth in *E. nitens*
- Majority of these SNPs are positively correlated for both traits
- Few SNPs (<5%) shared between the two species

#### Solid wood traits

CSIR

- 97 markers associated with MFA in *E. nitens* 
  - Many SNPs in genes that have a role in cytoskeleton development (actin, tubulin)



## **TESTING THE MARKERS**





## **Testing marker predictions**

Based on DNA alone we predict the trait in trees that have already been phenotyped

**Predictive ability** 

The correlation (r) between our marker predictions (MBVs) and phenotypic measurements in a modest number of trees



#### Accuracy

Accuracy of phenotypic selection

Accuracy of marker-based selection

v heritability (h<sup>2</sup>) or h

Predictive ability (r) V heritability (h<sup>2</sup>) or h

> Gondwana GENOMICS

### Marker predictions in E. globulus

Predicting pulp yield in 71 *E. globulus* clones (5-6 ramets)





Laboratory Pulp Yield

**CSIRC** 







## Pulp yield predictions in *E. nitens*

Testing pulp yield markers in 64 E. nitens seed orchard trees



Gondwon

Percent of markers used in prediction



Marker predictive ability is limited by accuracy of phenotype

### Growth predictions in *E. nitens*

Testing growth markers in 64 *E. nitens* seed orchard trees



Number of markers used in prediction



Marker predictive ability is limited by accuracy of phenotype

Molecular accuracy





# APPLICATION OF MARKER-ASSISTED SELECTION





## **Blue gum breeding cycle**

#### CURRENT

		Bree	ding			٦	Seed orchards								
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	



CSIRC

#### Marker-assisted selection

Select better parents Select better progeny Select better seed orchard trees



### **Applying MAS in seed orchards**



### **Applying MAS to controlled crosses**



### **Designing better crosses**

You can target complementary crosses to pyramid more good alleles.



### Marker genotypes clearly reveal better trees





### Marker data provided by Gondwana

- Genotype data on each marker
- · Good and bad allele identified
- Molecular breeding value (MBV) for each tree for each trait
- Coefficient of inbreeding (CI) for each tree
- Combining ability matrix for all trees in breeding program for tailoring crosses

							Density SNPs N													MBV	CI															
											1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23 24	25 n		
							Growth SNPs MBV CL												2.8	-0.01																
						+	_														2.6	0.09														
						$\vdash$		1	2	3	4	5	6 m	/	8	9		11 1	12	13 :	14	15	16	17 :	18 :	19	20 3	21 .	12 2	3 24	25	n	36	-0.0	2.3	-0.02
Т												Pu	ln V	ر امار	1 5 1	IPs												м	ΒV	CI			2.0	-0.0	2.0	0.03
+	_							_				T M	141	Ter		113															_	AVA	2.6	0.0	1.7	0.01
+	_	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	n				-	A/B	2.2	-0.0	2 1.5	0.00
┢	1	A/A	A/A	A/B	A/A	A/A	A/B	8/8	A/A	A/B	A/B	A/A	A/A	A/B	8/8	A/A	A/B	A/B	A/A	A/A	A/B	A/A	A/A	A/B	A/B	8/8	A/A	2	2.8	-0.01	-	A/B	2.0	0.0	1.2	0.25
┢	2	A/B	A/B	A/B	A/B	A/B	A/A	A/B	A/A	A/A	A/A	A/B	A/B	A/A	A/B	A/A	A/A	A/B	A/A	A/B	A/A	A/B	A/A	A/A	A/A	A/B	A/A	2	2.6	0.09	_	A/A	1.6	0.0	0.8	0.02
+	3	A/B	A/B	A/A	A/A	A/B	A/B	A/A	A/A	8/8	A/B	A/8	A/B	A/B	A/A	A/A	A/B	A/A	A/A	A/B	A/B	8/8	A/A	8/8	A/B	A/A	A/A	2	.3	-0.02		A/A	1.5	0.0	0.5	-0.01
	4	A/B	A/A	8/8	A/A	A/A	A/B	A/A	A/B	A/B	A/A	A/B	A/A	A/B	A/A	A/B	A/B	A/A	A/B	A/A	A/B	A/A	A/B	A/B	A/A	A/A	A/B	2	2.0	0.03	_	A/B	1.1	0.2	-0.2	0.00
; L	5	A/B	A/A	8/8	8/B	A/B	A/A	A/B	A/A	A/B	A/B	A/B	A/B	A/A	A/B	A/A	A/A	A/B	A/A	A/B	A/A	A/B	A/A	A/B	A/B	A/B	A/A	1	.7	0.01	_	A/A	0.8	0.0	-0.8	-0.01
z L	6	A/A	A/B	A/B	A/B	A/B	A/B	8/8	A/B	A/A	A/B	A/A	A/B	A/B	8/8	A/B	A/B	A/A	A/B	A/B	A/B	8/8	A/B	A/A	A/B	8/8	A/B	1	5	0.00		A/B	0.4	-0.0	1 -0.9	0.03
۳Ľ	7	A/B	A/B	8/8	A/B	A/A	A/A	8/8	A/B	8/8	A/A	A/B	A/A	A/A	8/8	A/B	A/A	B/B	A/B	A/A	A/A	8/8	A/B	8/8	A/A	8/8	A/B	1	.2	0.25		A/B	-0.3	0.0	-1.1	-0.02
-L	8	A/B	B/B	A/A	A/B	A/B	A/B	A/B	A/A	A/B	8/8	A/B	A/B	B/B	A/B	A/A	8/B	A/B	A/A	A/B	B/B	A/B	A/A	A/B	B/B	A/B	A/A	6	.8	0.02		A/A	-0.7	-0.0	1	
	9	B/B	B/B	A/A	A/A	A/B	A/A	A/B	A/B	A/A	A/B	A/A	A/B	A/A	A/B	A/B	A/A	A/B	A/A	A/B	A/A	A/B	A/B	A/A	A/A	A/B	A/B	0	.5	-0.01		A/B	-0.8	0.0	1	
	10	A/A	B/B	8/8	A/B	A/A	A/B	A/A	A/B	A/B	A/B	A/B	A/A	A/B	A/A	A/A	A/B	A/A	A/B	A/A	A/B	A/A	A/B	A/B	A/B	A/A	A/B	-	0.2	0.00		A/B	-1.2	-0.0	2	
	11	A/B	A/B	A/B	A/B	A/B	8/8	A/B	A/A	A/B	A/A	A/B	A/B	8/B	A/B	A/B	8/B	A/B	A/B	A/B	B/B	A/B	A/A	A/B	B/B	A/B	A/A	-	).8	-0.01						
	12	A/B	B/B	A/B	A/A	A/A	8/8	A/B	8/8	A/A	A/B	A/A	A/A	8/8	A/B	A/A	8/B	A/B	A/A	A/A	B/B	A/B	8/8	A/A	8/8	A/B	A/B	-	).9	0.03						
	13	B/B	A/A	A/B	A/B	8/B	A/B	A/A	A/B	8/8	8/8	A/8	B/B	A/B	A/A	8/8	A/B	A/A	A/B	B/B	A/B	A/A	A/B	B/B	A/B	A/A	A/B	1	L.1	-0.02						
эİI	κŪ	7																															Go	ndu	000	

### **Benefits of marker-assisted selection**

#### SHORTEN BREEDING CYCLE BY SELECTION IN SEEDLINGS ACCURATE SELECTION OF BEST PARENTS FOR CROSSES ACCURATE SELECTION OF ELITE OFFSPRING HIGHER SELECTION INTENSITY BY SCREENING 1000'S OF TREES SELECT ON MULTIPLE TRAITS AT THE SAME TIME

POTENTIALLY 4 TO 6 TIMES THE CURRENT RATE OF GENETIC GAIN

**OTHER BENEFITS** 

ELIMINATE LABELLING ERRORS IN BREEDING PROGRAMS ELIMINATE INBREEDING MAINTAIN HIGH GENETIC DIVERSITY FOR FUTURE BREEDING COST SAVINGS ON TRAIT MEASUREMENTS



## **Summary**

- New E. nitens (EN) and E. globulus (EGG) association populations established
- Predicted pulp yield, cellulose content and other traits in EGG and EN
- Solid wood traits in EN
- Large numbers of markers controlling growth and pulp yield identified in *EN* and *EGG*
- Large numbers of markers controlling wood stiffness (MFA) identified in EN
- Markers demonstrated to predict accurately in different populations growing in different environments
- Marker-assisted selection service developed with the formation of Gondwana Genomics spin off





# Acknowledgements

CSIRO Tricia Stewart Cate Smith Bala Thumma Conductor Rob Evans Saravanan Thavamanikumar Conductor Jeremy Brawner Conductor David Spencer Australian Tree Seed Centre

#### Ben Bradshaw

Dean Williams

Stephen Elms Ross Gillies

**Kelsey Joyce** 

Chris Lafferty

Forest&Wood Products Australia

Andrew Lyons

Ian Ravenwood

#### PRIVATE FORESTS TASMANI.





Contact <u>simon@gondwanagenomics.com.au</u> Mobile: 0477 700 643













1 Wilf Crane Crescent, Yarralumla ACT 2600 | www.gondwanagenomics.com.au | +612 6109 6111