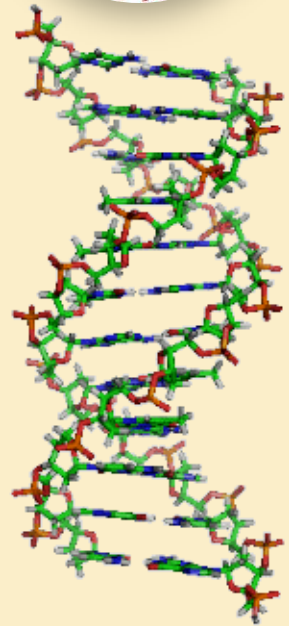


Gene-Associated Single Nucleotide Polymorphism (SNP) in *Cinnamate 4-Hydroxylase (C4H)* and *Cinnamyl Alcohol Dehydrogenase (CAD)* Genes from *Acacia mangium* Superbulk Trees

Tchin Boon Ling (09021555)

Supervisor : Dr Ho Wei Seng

Co-supervisors : Assoc Prof Dr Ismail Jusoh
Dr Pang Shek Ling



Introduction



Improved Trees - Acacia Superbulk



Acacia mangium superbulk

- Genetically improved planting materials
- Produced through plus trees selection over 30 years in a tree improvement program
- Main tree species planted under Licenses of Planted Forest (LPFs) (PERKASA, 2009)
- Up to 2009, 25,380 hectares in Sarawak have been planted with *Acacia* (Ta Ann Holdings Berhad, 2010)



■ Potential uses:

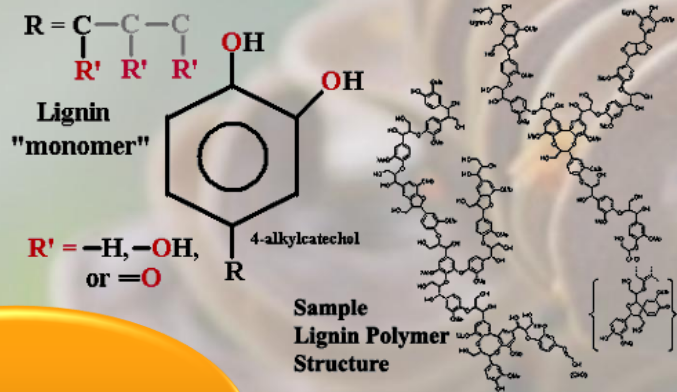
- plywood, veneers and fine furniture
- construction of building
- paper and pulp



- Understanding the structure and composition of *Acacia* wood is vital for better utilization of this wood material



WOOD FORMATION

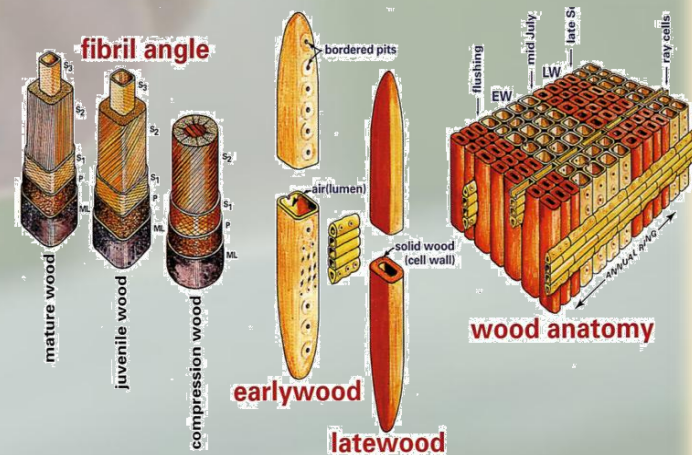
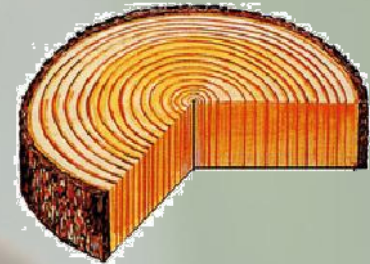


Lignin



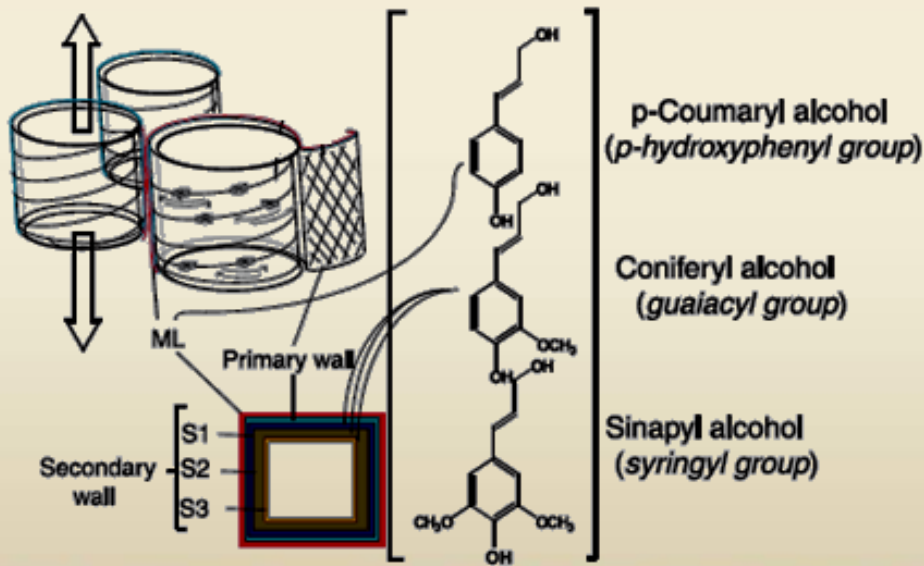
Wood

Carbohydrates
(cellulose & hemicellulose)



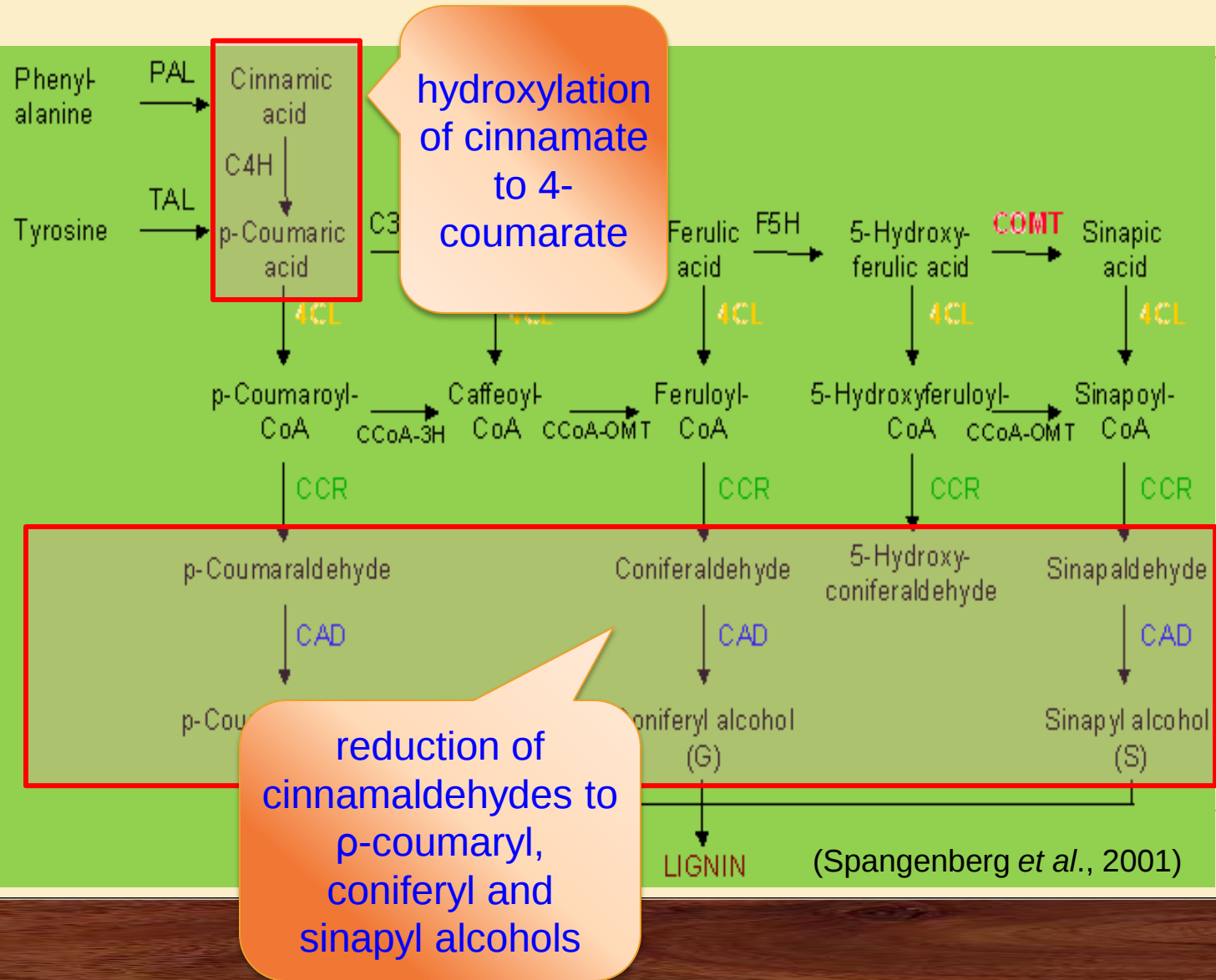
LIGNIN

Lignin



- ❑ 20-30% of the plant biomass
- ❑ Provide mechanical & structural support
- ❑ Protective function against pathogen or decaying fungi

Lignin Biosynthesis Pathway



C4H & CAD

- ✗ Any up- or down-regulation of lignin biosynthesis genes will result in **altered lignin production** (Baucher *et al.*, 2003)
- ✗ Abreu *et al.* (2009) had proposed that the high **β -O-4 (Alkyl Aril Ether) bonds** in lignin of angiosperms may possibly **affect the wood properties**
- ✗ Therefore, extensive study (e.g. **association genetic study**) is needed to be carried out in order to determine the **association** between the lignin genes and the wood properties



Why Association Study?

- Traditional linkage-based approaches is time consuming, laborious and often expensive process of establishing the mapping population (Myles *et al.*, 2009)

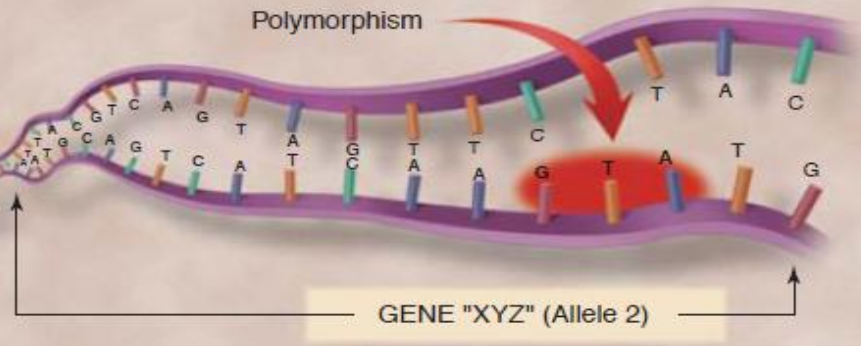
**New
opportunity**

- Candidate-gene-based association study which involves the identification of causative polymorphisms (SNPs) for excellent traits has been proposed as a promising approach to dissect complex traits in forest trees (Gonzalez-Martinez *et al.*, 2006)
- **Advantages of SNPs:** abundance, stable, ubiquity and interspersed in nuclear genome (Fusari *et al.*, 2008)

Susceptible Plant



Polymorphism



The order of individual nucleotides identifies the dominant gene and the differences between the genes. Also, the nucleotides supply the code for protein production resulting in the plant expressing a particular trait. It's the differences in these nucleotide strings that modify or alter the plant's trait expression.

Codes for.....



Altered protein has modified function



Gene and exact polymorphism (QTN) identified



Causal
SNPs in
specific
genes

Quantitative
Trait
Nucleotides
(QTNs) of
quantitative
variation

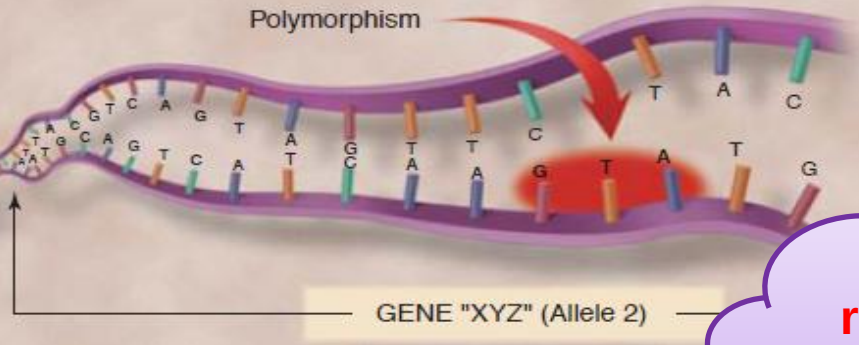
Gene
assisted
selection
(GAS) in
breeding
program

(Fusari et al., 2008)

Susceptible Plant



Polymorphism



GENE "XYZ" (Allele 2)

Codes for.....



Altered protein has modified function

The order of individual nucleotides identifies the dominant gene and the differences between the genes. Also, the nucleotides supply the code for protein production resulting in the plant expressing a particular trait. It's the differences in these nucleotide strings that modify or alter the plant's trait expression.

Maximum resolution and identification of exact allele !!!

Gene and exact polymorphism (QTN) identified

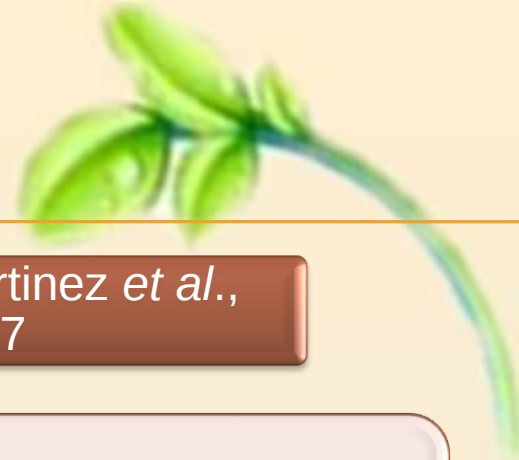


Causal SNPs in specific genes

Quantitative Trait Nucleotides (QTNs) of quantitative variation

- save time & money
- early seedling selection is possible
- better wood quality & quantity being produced
- better economic outcome

Literature Review



Thumma *et al.*, 2005

2 SNPs found in *CCR* gene were significantly associated with microfibril angle from *Eucalyptus nitens*

Yu *et al.*, 2006

Detected the association of the *cad-n1* allele with increased stem growth and wood density in 15 years old loblolly pine

Gonzalez-Martinez *et al.*, 2007

Nonsynonymous substitution for *CAD* was in strong association with earlywood specific gravity and causing lignin modification in *Pinus taeda*

“multigene association genetic study in forest trees has shown the feasibility of candidates genes strategies for dissecting complex adaptive traits, provided that the genes belonging to key pathways and appropriate statistical tools are used”



Objectives



To identify the SNP that might present in *C4H* and *CAD* genes from *Acacia mangium* superbulk trees



To determine the genetic association among SNPs in *C4H* and *CAD* genes and an array of wood properties from *Acacia mangium* superbulk trees



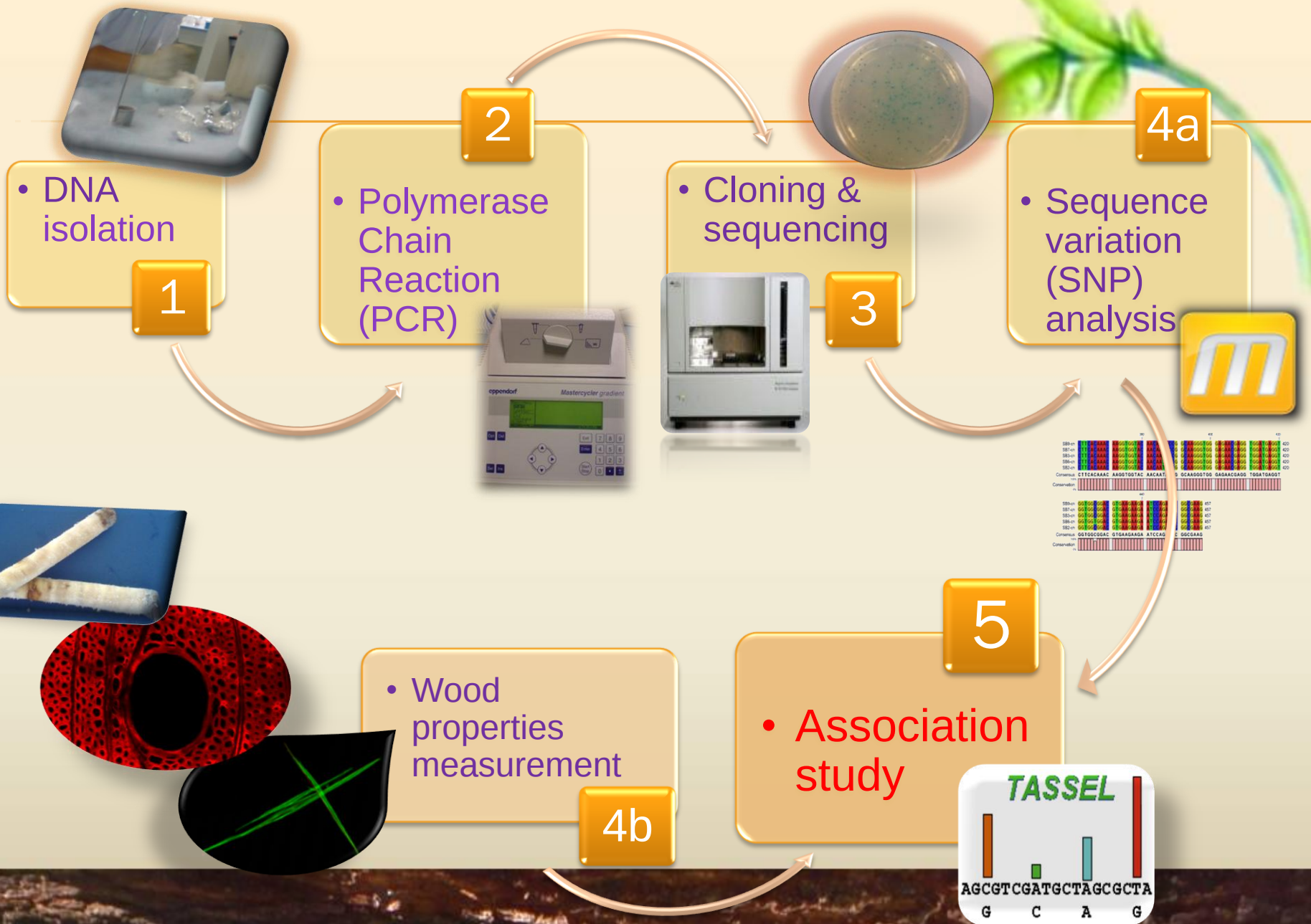
Materials



&

Methods







Results

Results

and

and

Discussion

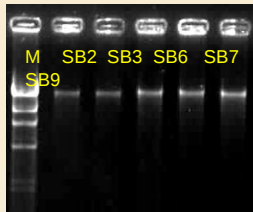
Discussion



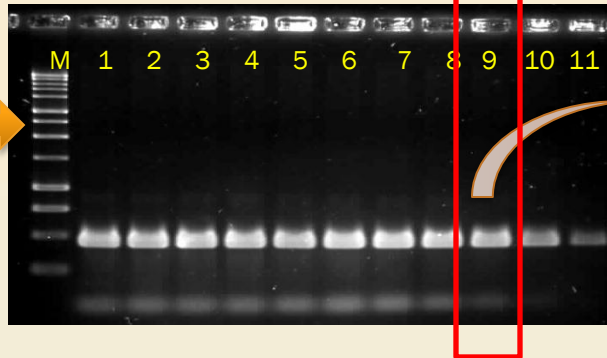
Cloning of Partial Gene Sequence



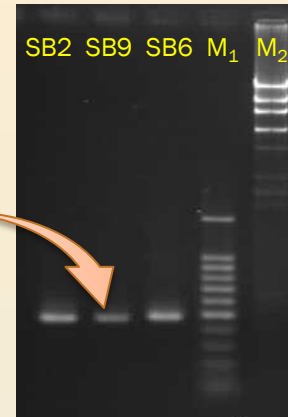
DNA
extraction



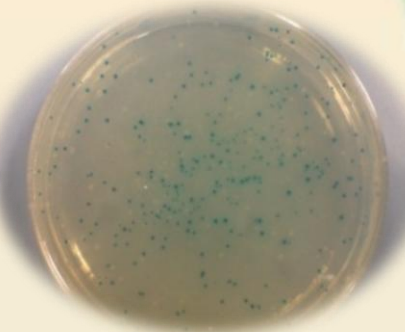
PCR optimization: T_a
63°C



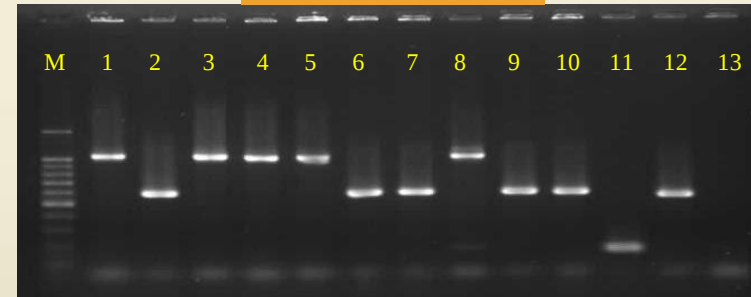
Purified PCR products



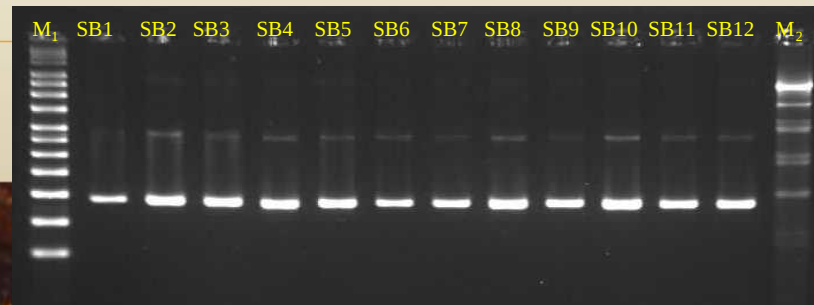
Blue/white screening



Colony PCR

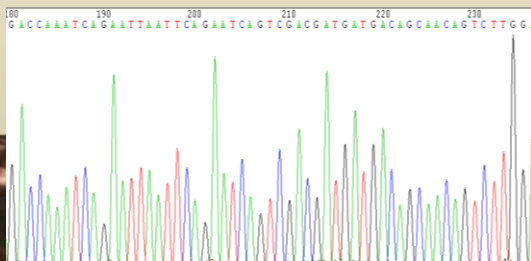


Plasmids DNA



Data
analysis

Sequencing



Data Analysis

(a) Sequences Verification

■ Sequences were analyzed using BLASTn (<http://blast.ncbi.nlm.nih.gov/>) to find out all possible similar nucleotide sequences

Table 1: BLASTn output for partial *C4H* DNA sequence.

Organisms	GenBank Accession No.	Similarity (%)	E-value
<i>Acacia auriculiformis</i> x <i>Acacia mangium</i>	EU275980.1	98	0.0
<i>Sorghum bicolor</i>	AY034143.1	86	4e-8e
<i>Salvia miltiorrhiza</i>	EF377337.1	84	3e-100
<i>Rubus coreanus</i>	EU123531.1	82	1e-98

Table 2: BLASTn output for partial *CAD* DNA sequence.

Organisms	GenBank Accession No.	Similarity (%)	E-value
<i>Acacia auriculiformis</i> x <i>Acacia mangium</i>	EU275982.1	98	0.0
<i>Leucaena leucocephala</i>	EF611250.1	94	0.0
<i>Populus tomentosa</i>	EU760897.1	80	2e-79
<i>Eucalyptus urophylla</i>	GQ387647.1	82	3e-66



INTRON EXON FINDING

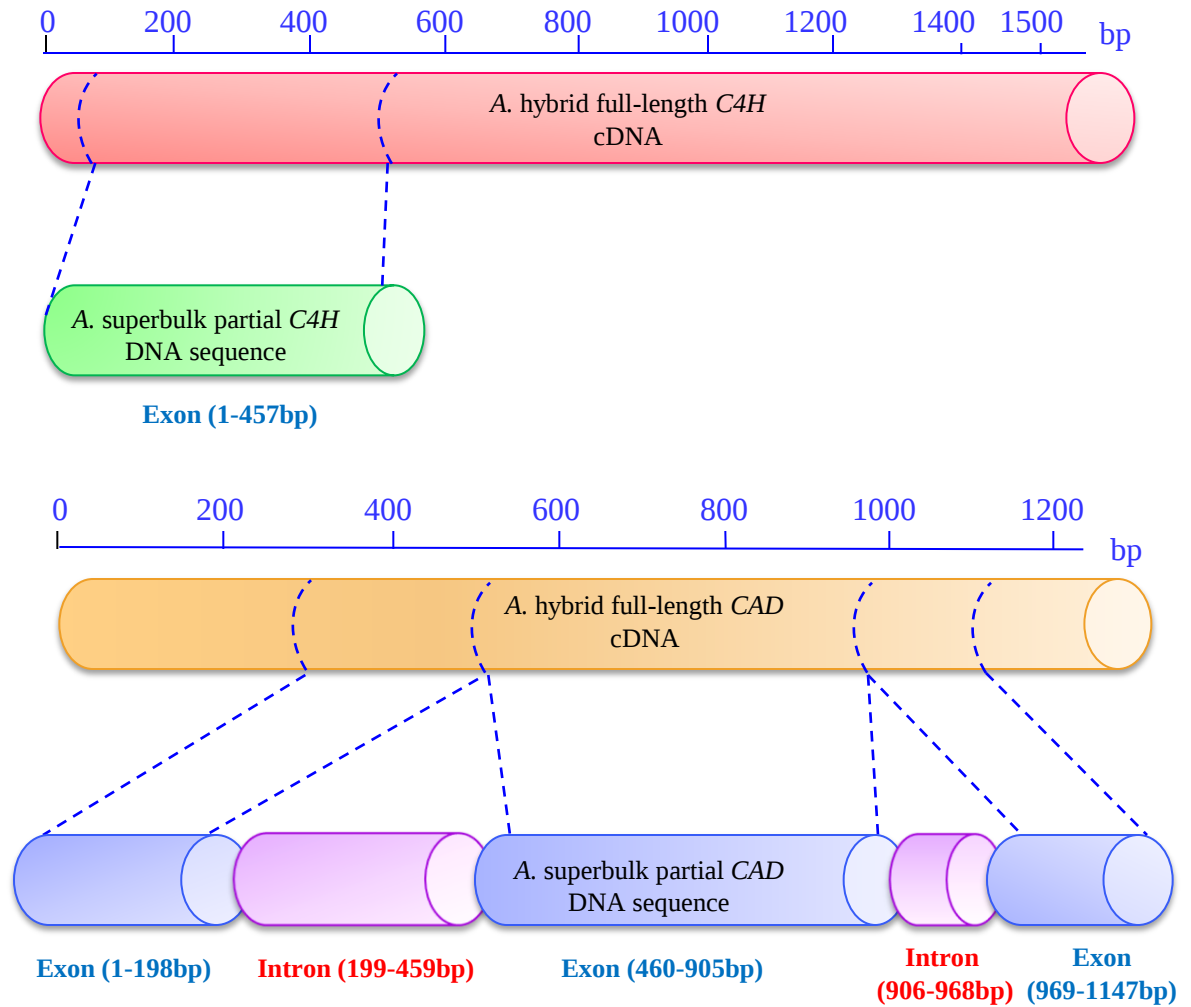
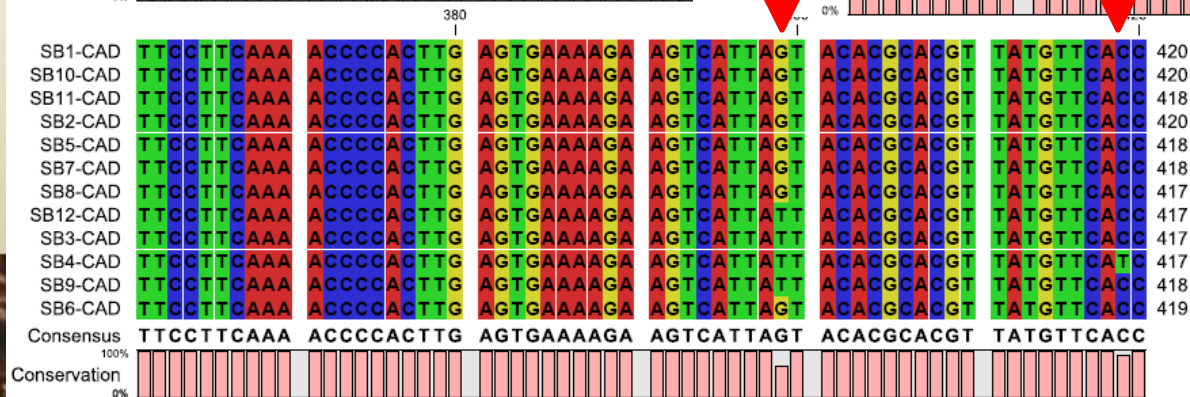
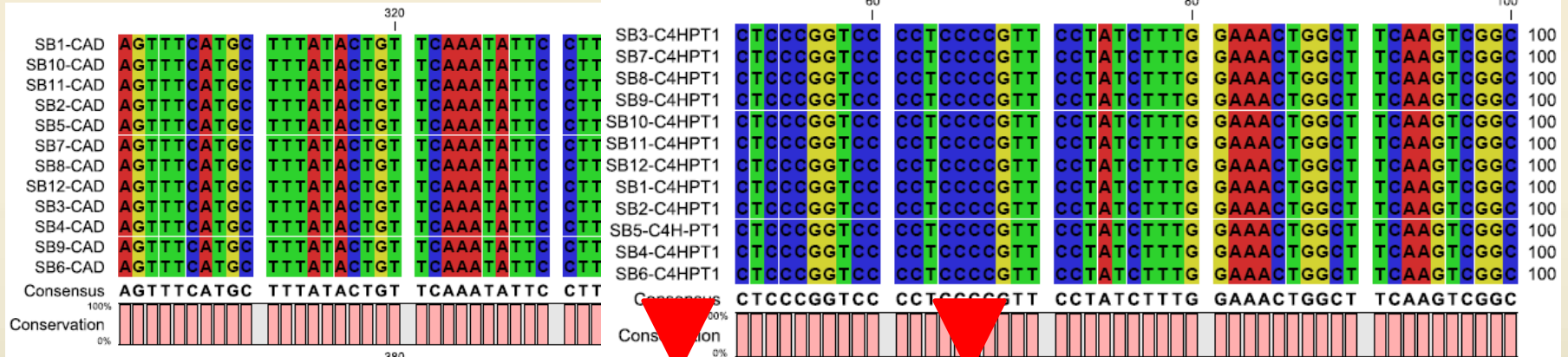
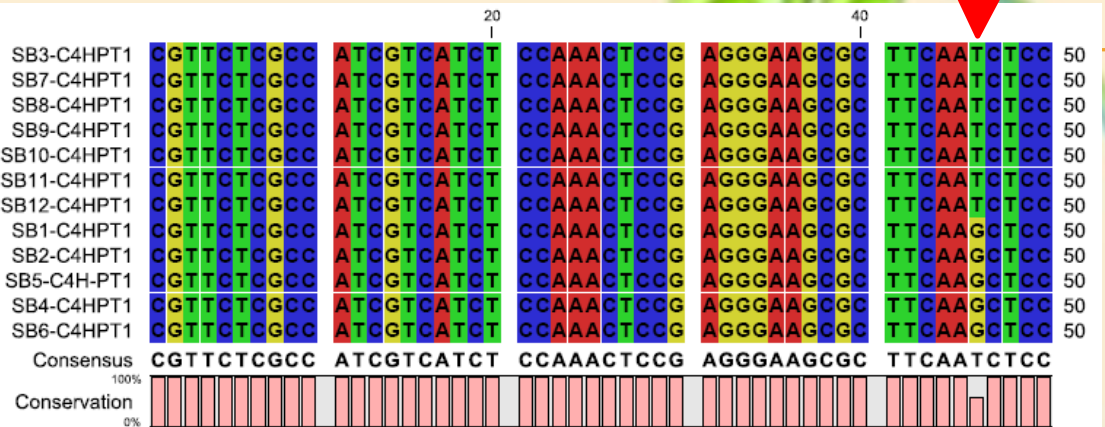


Figure The exonic and intronic regions found in partial *C4H* and *CAD* DNA sequence from *A. superbulk*.

DNA Sequence Variations



[illegible]

	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTAGT	ACACGCACGT	TATGTTCAACC	418
SB5-CAD	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTAGT	ACACGCACGT	TATGTTCAACC	418
SB7-CAD	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTAGT	ACACGCACGT	TATGTTCAACC	418
SB8-CAD	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTAGT	ACACGCACGT	TATGTTCAACC	417
SB12-CAD	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTATT	ACACGCACGT	TATGTTCAACC	417
SB3-CAD	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTATT	ACACGCACGT	TATGTTCAACC	417
SB4-CAD	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTATT	ACACGCACGT	TATGTTCAACC	417
SB9-CAD	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTATT	ACACGCACGT	TATGTTCAACC	418
SB6-CAD	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTAGT	ACACGCACGT	TATGTTCAACC	419
Consensus	TTCCTTCAAA	ACCCCACTTG	AGTGAAAAGA	AGTCATTAGT	ACACGCACGT	TATGTTCAACC	
Conservation	100%	100%	100%	100%	100%	100%	
	0%	0%	0%	0%	0%	0%	

Synonymous & Nonsynonymous Mutations

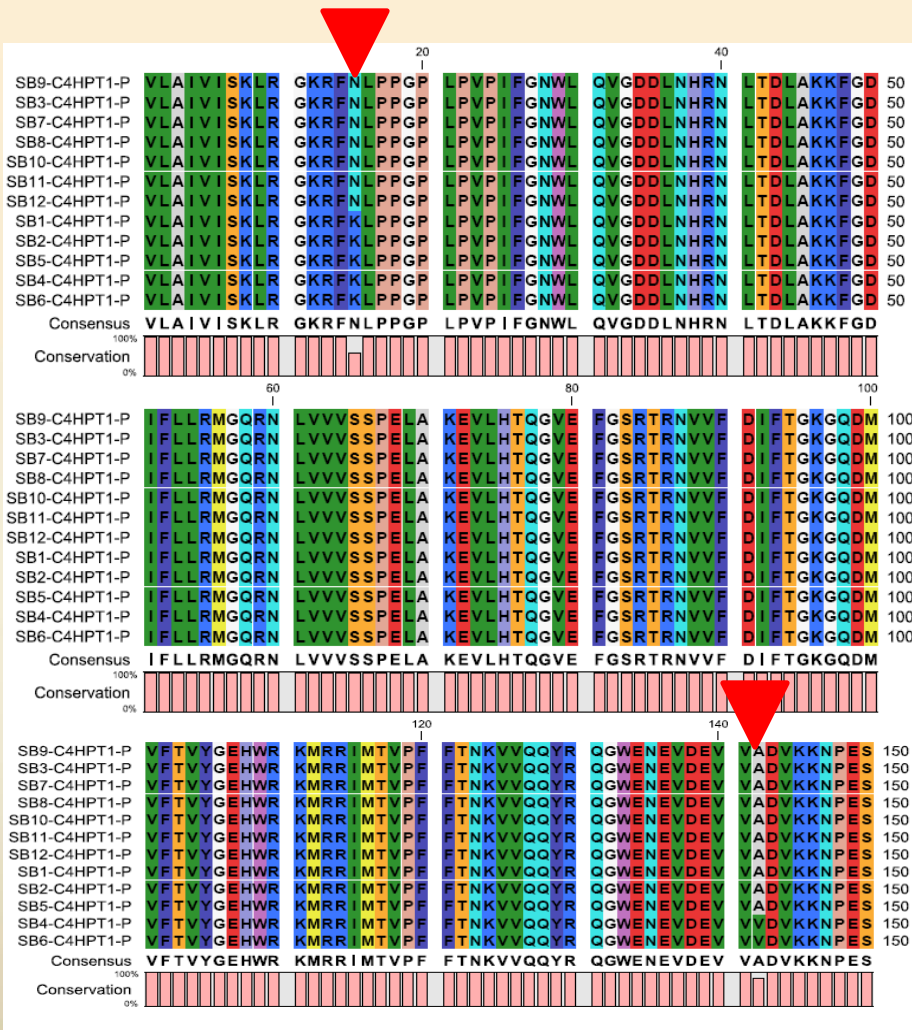


Table 3: Nonsynonymous mutations in partial DNA sequences of *C4H*.

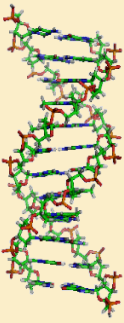
Amino acid position	Nucleotide changes	Amino acid alteration
15	T → G	N → K
142	C → T	A → V

Table 4: Number of synonymous and nonsynonymous mutations detected in the open reading frame of *C4H*.

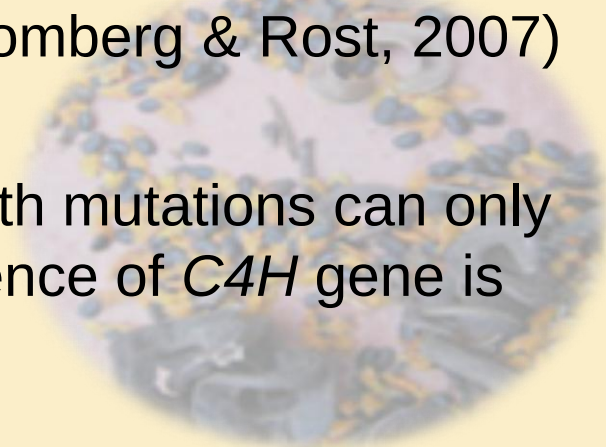
Total number of nonsynonymous mutations	2
Total number of synonymous mutations	-
Total number of amino acid	152
Proportion of nonsynonymous, dN	0.013
Proportion of synonymous, dS	-

Figure: Multiple alignment of *C4H* amino acid sequences.

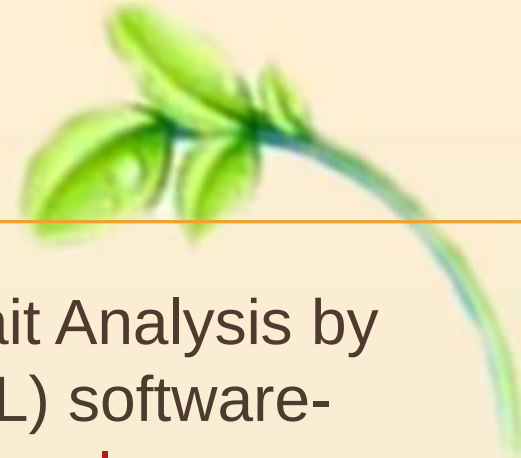
Synonymous & Nonsynonymous Mutations



- Synonymous mutations: caused by degenerative of genetic code but will not change the amino acid composition of the encoded protein
- Nonsynonymous mutations:
 - changes the encoded amino acid
 - may show functional or biochemical effects, therefore possibly lead to the changes in phenotypic characteristic of the trees (Bromberg & Rost, 2007)
- Decision on the popularity of both mutations can only be made when full length sequence of *C4H* gene is studied



ASSOCIATION ANALYSIS



- ✧ General linear model (GLM) provided by Trait Analysis by Association, Evolution and Linkage (TASSEL) software-available at <http://www.maizegenetics.net/tassel>

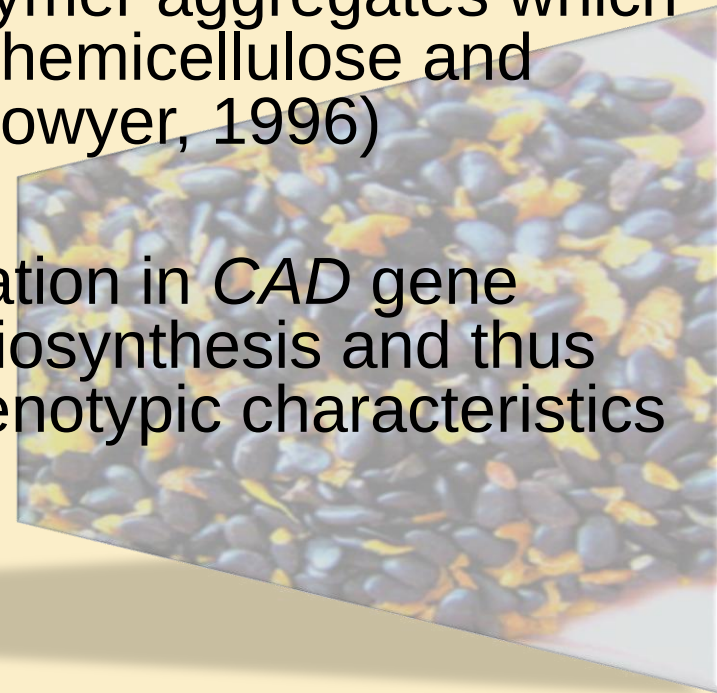
Table 5 Significant association between marker in CAD gene and wood properties (SG=specific gravity; WD=wood density; CWT=cell wall thickness).

Trait	SNP	Marker Effect		
		F	P	r ²
SG	M5	5.0755	0.0479	0.3367
	M6	5.5964	0.0396	0.3588
	M8	5.0755	0.0479	0.3367
WD	M6	6.7770	0.0263	0.4039
CWT	M5	9.2438	0.0125	0.4804
	M8	9.2438	0.0125	0.4804

Significant association: $p < 0.05$

ASSOCIATION ANALYSIS

- ✗ Findings (CAD + specific gravity & wood density) were similar with the previous studies done by Gonzalez-Martinez *et al.* (2007) and Yu *et al.* (2006), respectively
- ✗ Thickening of cell wall is affected by the arrangement of biopolymer aggregates which comprise of cellulose, hemicellulose and lignin(Haygreen and Bowyer, 1996)
- ✗ Single nucleotide mutation in *CAD* gene might alter the lignin biosynthesis and thus lead to changes in phenotypic characteristics



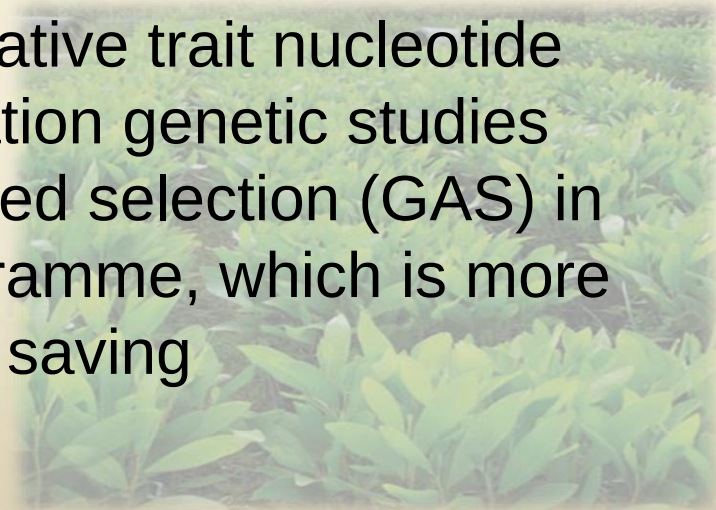
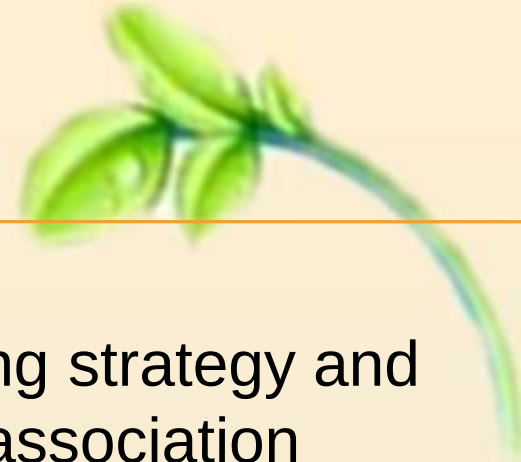
Conclusions

- 2 SNPs were found in the exon of *C4H*, of which all the SNPs caused nonsynonymous mutations
- 5 SNPs were identified in the *CAD* exons along with one deletion mutation. In addition, 2 SNPs were also identified in the *CAD* introns
- SNPs in *CAD* gene do associate with the specific gravity, wood density and cell wall thickness ($p < 0.05$)
- However, further study is needed to validate this finding



Conclusions

- Combination of SNP genotyping strategy and candidate genes approach in association genetic studies is feasible, especially in forest tree species with unknown and large genome
- Identification of quantitative trait nucleotide (QTN) through association genetic studies will lead to gene-assisted selection (GAS) in the tree breeding programme, which is more cost effective and time saving



Recommendations

- Discovery of the full length *C4H* and *CAD* genes which including the promoter, intron, exon and 5'/3'-untranslated regions is advisable
- Population structure parameter should be included in future study by using an independent set of random markers (Zhu *et al.*, 2008)
- Sample size (200-300) and the number of populations (natural population/ full-sib family mapping population) should be increased to obtain more powerful data for association mapping



Acknowledgements



Sincere appreciation and deepest gratitude to:

@ DR. HO WEI SENG

@ Assoc Prof. DR. ISMAIL JUSOH

@ DR. PANG SHEK LING (SFC)

@ MISS KAMALIAWATI

@ ALL LABMATES



References

- Bromberg, Y. & Rost, B. (2007). SNAP: Predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Research*, 35(11), 3823–3835.
- Abreu, H. S., Latorraca, J. V. F., Pereira, R. P. W., Monteiro, M. B. O. Abreu, F. A & Amparado, K. F. (2009). A supramolecular proposal of lignin structure and its relation with the wood properties. *Annals Of The Brazilian Academy Of Sciences*, 81 (1), 137-142.
- Baucher, M., Halpin, C., Petit-Conil, M. & Boerjan, W. (2003). Lignin: Genetic engineering and impact on pulping. *Critical Reviews in Biochemistry and Molecular Biology*, 38, 305-350.
- Fusari, C. M., Lia, V. V., Hopp, H. E., Heinz, R.A. & Paniego, N. B. (2008). Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach. *BMC Plant Biology*, 8(7).
- Gonzalez-Martinez, S. C., Wheeler, N. C., Ersoz, E., Nelson, C. D. & Neale, D. B. (2007). Association Genetics in *Pinus taeda* L. I. Wood Property Traits. *Genetics*, 175, 399-409
- Haygreen, J. G. and Bowyer, J. L. 1996. Forest Products and Wood Science, 3rd edition. IOWA State University Press, Ames. pp. 41-50.
- Ismail, J. & Farawahida, A. Z. (2007). Physical and Mechanical Properties of *Acacia mangium* x *Acacia auriculiformis* hybrid (*Acacia* hybrid) and *Acacia mangium* superbull Planted in Sarawak. *Paper presented in the Conference of Forestry and Forest Products Research (CFFPR 2007), November 27-29.*
- Spangenberg, G., Kalla, R., Lidgett, A., Sawbridge, T., Ong, E.K., & John, U. (2001). Transgenesis and Genomics in Molecular Breeding of Forage Plants. Online 10th Proceedings of the Australian Agronomy Conference. Science and Technology: Delivering Results for Agriculture, January 29. Hobart. Retrieved April 27, 2009, from <http://www.regional.org.au/au/asa/2001/plenary/6/spangenberg.html>
- Yu, Q., Li, B., Nelson, C. D., McKeand, S. E., Batista, V. B. & Mullin, T. J. (2006). Association of the *cad-n1* allele with increased stem growth and wood density in full-sib families of loblolly pine. *Tree Genetics and Genomes*, 2, 98-108.
- Gonzalez-Martinez, S. C., Krutovsky, K. V. & Neale, D. B. (2006). Forest-tree population genomics and adaptive evolution. *New Phytologist*, 170, 227-238.
- Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E. & Buckler, E. S. (2009). Association mapping: Critical consideration shift from genotyping to experimental design. *The Plant Cell*, 21, 2194-2202.

**THANK
YOU!**

