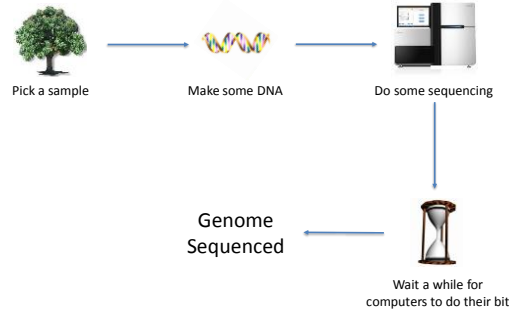


Sequencing (Complex) Genomes



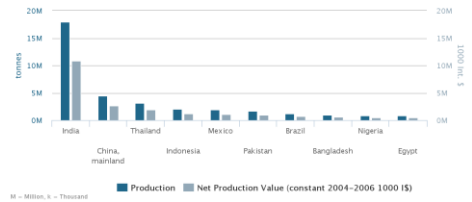
Mango

- *Mangifera indica* L.
- Member of the Anacardiaceae family
 - Cashew, pistachio, sumac
- Allotetraploid
 - 2n = 40
- Small genome size
 - 440 – 485 Mbp
- Australian mango genepool collection contains over 300 accessions



Mango as a commodity

- Global production > 40M tonnes annually (FAO 2013 – note includes mangosteens and guava)



Mango in Australia

Rank	Area	Production (tonnes)	Production (1000\$ Int)
1	India	18,002,000.00	\$10,786,204.33
2	China, mainland	4,450,000.00	\$2,666,293.15
3	Thailand	3,141,950.00	\$1,882,552.76
4	Indonesia	2,058,607.00	\$1,233,449.38
5	Mexico	1,901,871.00	\$1,139,538.34
6	Pakistan	1,658,562.00	\$993,755.62
7	Brazil	1,163,000.00	\$696,831.22
8	Bangladesh	950,000.00	\$569,208.65
9	Nigeria	850,000.00	\$509,291.95
10	Egypt	834,543.00	\$500,030.63
40	Congo	41,104.00	\$24,628.16
41	Australia	40,797.00	\$24,444.22
42	Chad	34,500.00	\$20,671.26

Mango in Queensland

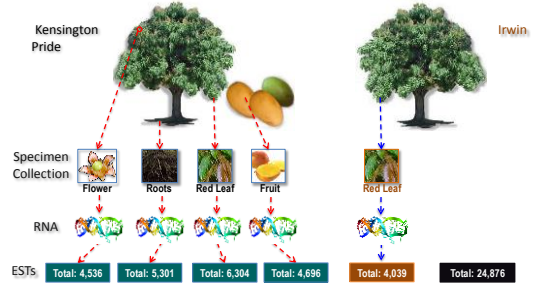
- Valued at >\$A60M annually
- Kensington Pride has over 60% of the market
- Out of the breeding program
 - R2E2
 - Selected in 1982
 - Released into production in 1991
 - Calypso
 - Released in 2007
 - Three new varieties released in 2014



Mango Genomics Initiative

- An initiative of the Queensland Government (Australia)
- Commenced in 2005 to:
 - Identify genes and markers for traits of interest
 - Fruit colour and flavour
 - Novel bioactives
 - Production traits – small tree, high productivity initiative
 - Toolkit for mango breeders
 - Improved breeding efficiency / targeted breeding
- Resources
 - A long history in mango breeding
 - Commercially released mango varieties
 - Global collection of 377 accessions
 - 3 hybrid populations (F1)

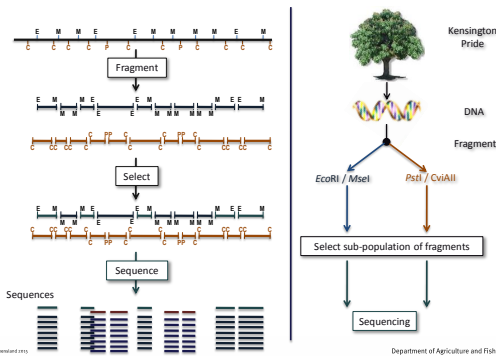
ESTs (Sanger)



Excellent sequence resource
EST-SSR Markers used for diversity analysis (Dillon et al 2013)

Needed to do more sequencing

Genome Complexity Reduction (454)

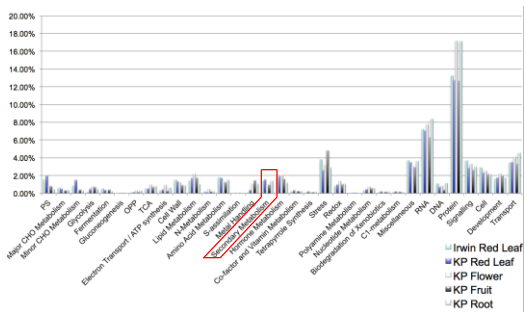


Sequence Data

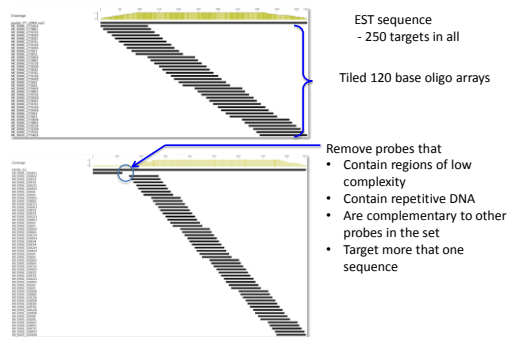
Library	Type	Annotated Sequences	No Annotation	Number of Sequences	Average Seq Length
KP Red Leaf	cDNA	5,290	1,014	6,304	473
KP fruit skin	cDNA	4,518	177	4,695	623
KP Flower	cDNA	4,215	285	4,500	550
KP Root	cDNA	5,242	60	5,302	704
Ir Red Leaf	cDNA	3,463	576	4,039	564
Total	cDNA	22,728	2,112	24,480	579
EM Library	Genomic	4,909	77,132	82,041	111
PC Library	Genomic	12,819	56,942	69,761	141
Total	Genomic	17,728	134,074	151,802	124
Total	All	40,456	136,186	176,282	

Needed to do more sequencing

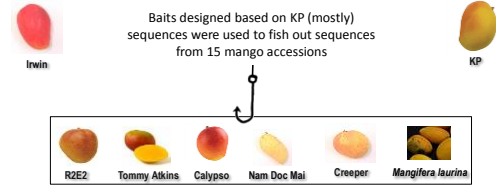
More Sequences



Target Sequencing

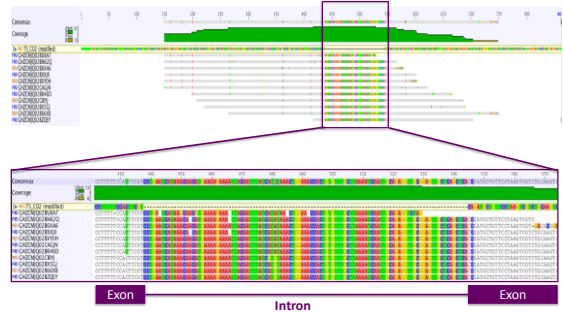


Going Fishing



© State of Queensland 2015 Department of Agriculture and Fisheries

Fishing for DNA Fragments



© State of Queensland 2015 Department of Agriculture and Fisheries

Targeted Sequencing



144 SNPs have been provided to Dr David Kuhn for contribution to a linkage map
60 SNPs being assess as potential markers for tree architecture traits

Opportunity to do more sequencing

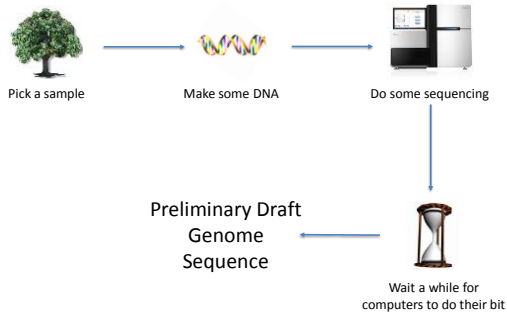
© State of Queensland 2015 Department of Agriculture and Fisheries

Sequencing the Kensington Pride Genome

- Joint initiative lead by Horticulture Innovation Australia Ltd with
 - Department of Agriculture and Fisheries (Qld)
 - International Crop Research Institute for the Semi-Arid Tropics (ICRISAT)
 - Beijing Genomics Institute (BGI)
- Strategy
 - Libraries insert sizes: 170, 200, 500, 800, 2000 and 5000 bp
 - 87 Gb of cleaned reads
 - SOAP de novo assembly
- Very preliminary results
 - N50: Contig - 16 kb; Scaffold - 129 kb
 - Genome size 407 Mbb
 - Depth: 210X
 - Highly heterozygous

© State of Queensland 2015 Department of Agriculture and Fisheries

Sequencing (Complex) Genomes



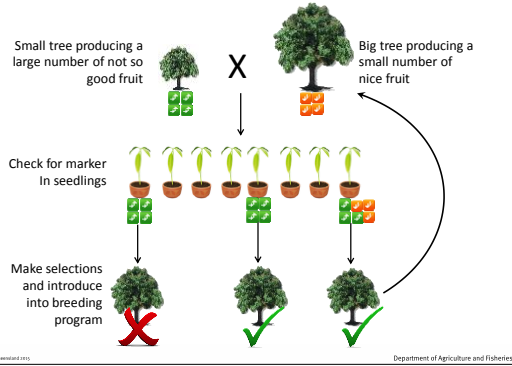
© State of Queensland 2015 Department of Agriculture and Fisheries

What's Next

- We need to do some more sequencing
 - Waiting on additional sequence data from 20 kb insert library
 - Annotation
- Skim / re-sequencing of the mapping populations
 - Markers
 - KP common parent
- Genetic map
- What are we trying to achieve...
 - A useful / usable KP mango genome sequence

© State of Queensland 2015 Department of Agriculture and Fisheries

The Overall Goal



Thanks to

Department of Agriculture and Fisheries
Natalie Dillon (Project lead)
Ian Bally
Roger Broadley
Louise Hucks
Matt Webb
Luke Devitt

ICRISAT
Rajeev Varshney
Prasad Bajaj
BGI
Yao Ming
Xuan Li
Xiaolei Gao

Queensland Alliance for Agriculture and Food Innovation (University of Queensland)
Ralf Dietgen
Tim Holton (currently with ILRI)

Subtropical Crops Research Institute
Ru Lin Zhan
Wu HongXia

Horticulture Innovation Australia Ltd
Alok Kumar

Thanks to Horticulture Innovation Australia and the Queensland Government for their ongoing support of this project