Genome Select Oil Palm: Innovation to Commercialization
Dr. Harikrishna Kulaveerasingam

November 2016

What is required to take research from idea to commercial realization?

Palm oil is used in more than half of packaged supermarket products today

Despite this, Palm Oil is Faced with Many Challenges

Competing commodity
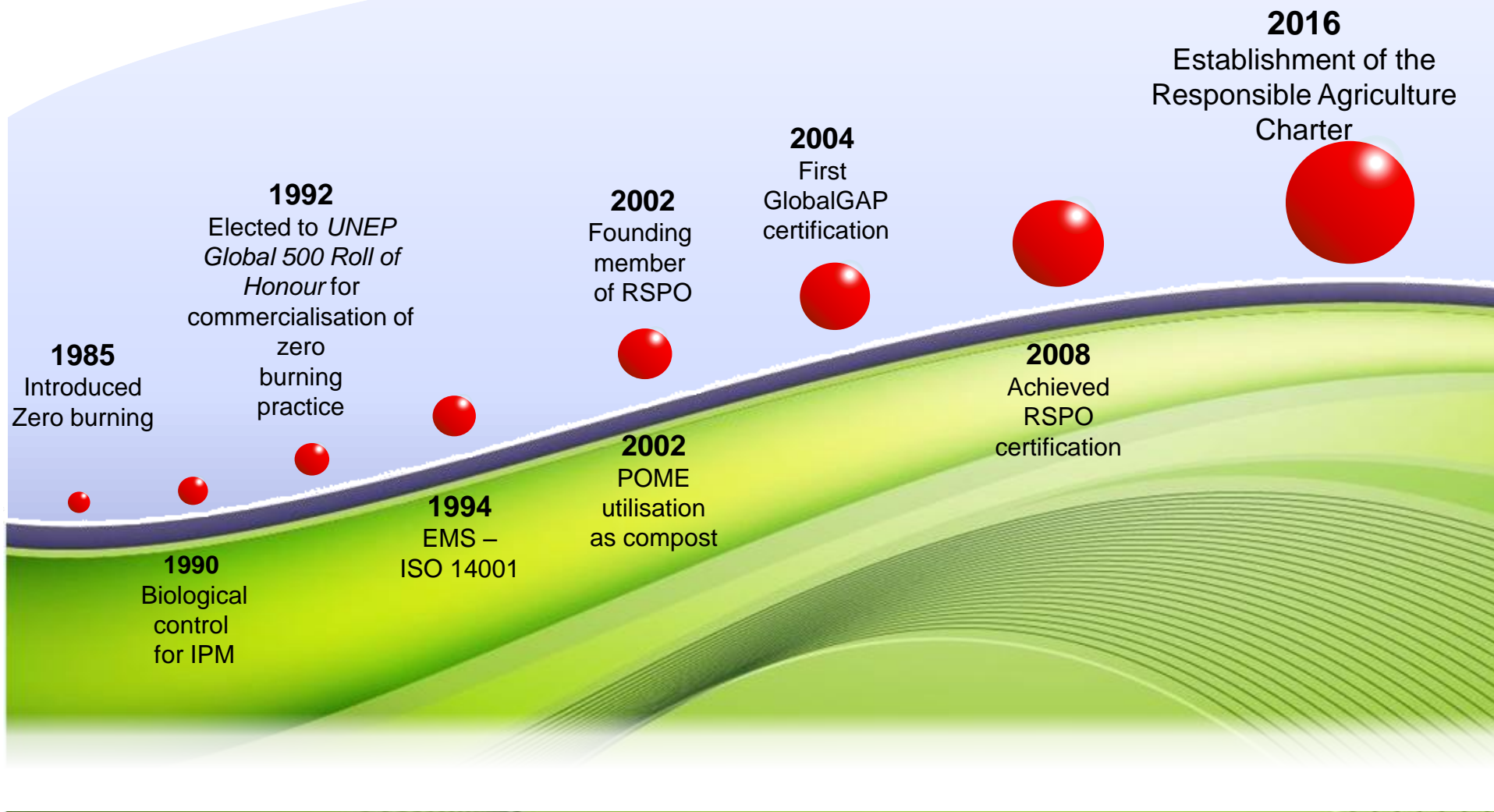
High labour requirement

Changing weather patterns

Lack of new arable land

Our Sustainability Journey

In addition, Sime Darby is committed to uncompromised sustainability

Sime Darby Plantation

1985
Introduced Zero burning

1990
Biological control for IPM

1992
Elected to *UNEP Global 500 Roll of Honour* for commercialisation of zero burning practice

1994
EMS – ISO 14001

2002
Founding member of RSPO

2002
POME utilisation as compost

2004
First GlobalGAP certification

2008
Achieved RSPO certification

2016
Establishment of the Responsible Agriculture Charter

Sustainability Practices

- Sime Darby Plantation pioneered many sustainable plantation initiatives which are considered best practices today:

Good Water Management

Excellent Plantation Management Practices

Ground Cover Crops for Soil and Moisture Conservation

Best Mill Practices

Balanced Age Profile Through New & Replanting Exercise
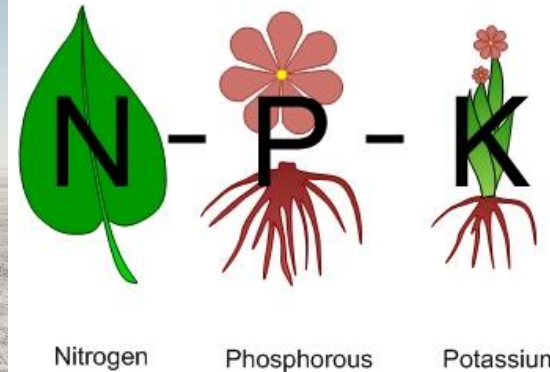

Integrated Pest Management

Zero Burning Replanting Technique

Palm oil mill effluent (POME) Treatment

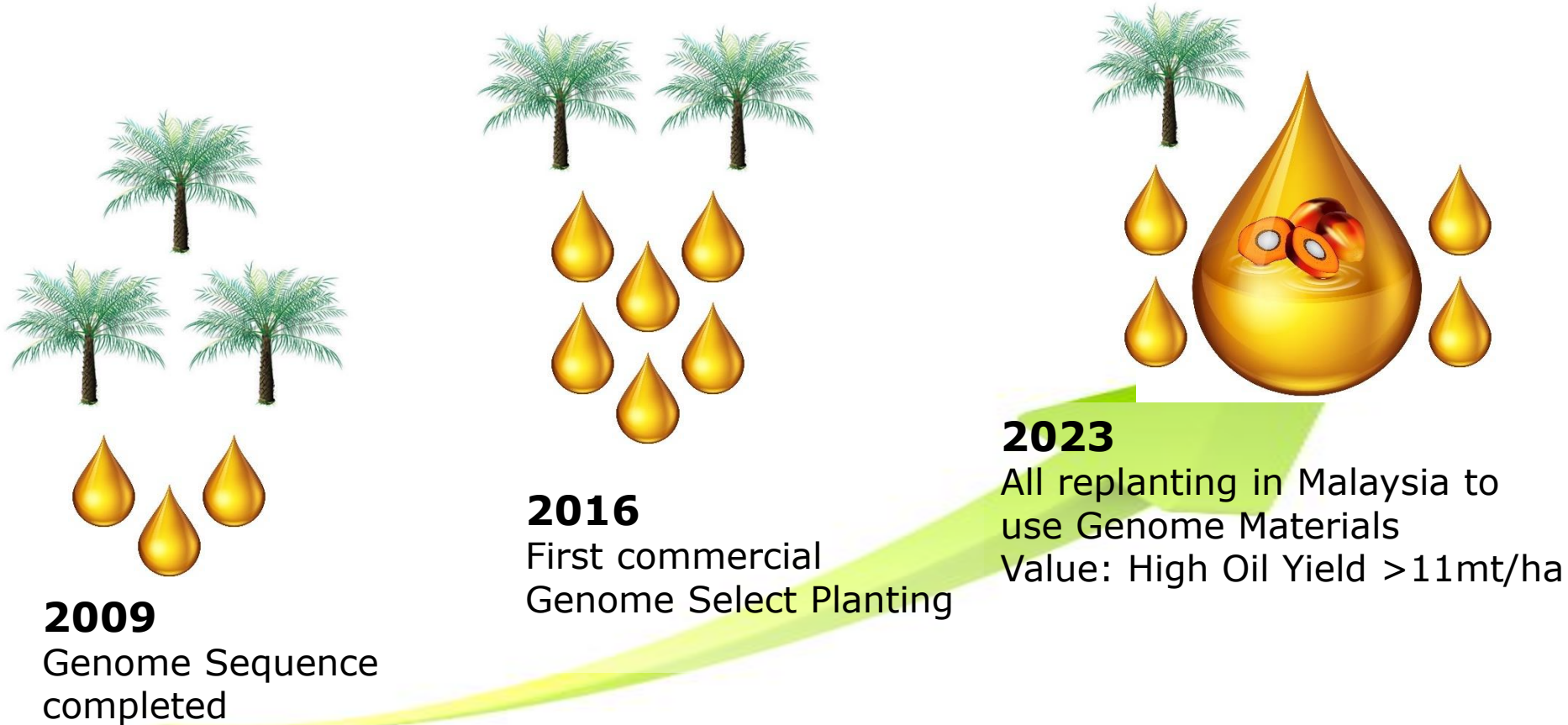# Climate impact, costs and resource scarcity
Selecting for yield across all environments is essential for the future

Nitrogen   Phosphorous   Potassium

Resource efficiency

Climate change problems; floods or drought, need to produce more from less Compounded by labour constraints

Oil Palm Genome: A Vision to Address the Bottleneck in Productivity Improvements

2009
Genome Sequence completed

2016
First commercial Genome Select Planting

2023
All replanting in Malaysia to use Genome Materials
Value: High Oil Yield >11mt/ha

Delivering a yield improvement equivalent to 50,000 ha

What Did We Do To Deliver The Genome Select Planting?

Best crosses Segregated and tracked to field

80,000 Seedlings sampled June 2015 to March 2016

80 million Genetic tests

New lab and IT systems Databases, analysis, Apps

Combined effort R&D-Operations Logistics, timing

FY2020/21 FFB difference apparent 48 months after planting

The Gap Between Idea and Commercialisation

Transforming the Team is Key

Basic Research

- Lots of ideas
- [More] flexible deadlines
- Discrete teams
- Focused research
- Specialists
- Small scale – low throughput
- Iterative

Commercialisation

- Singular strategy
- Backup plans
- "Drop-dead" dates and non-negotiable standards
- Multi-disciplinary teams
- Logistics and Operations Management
- Quality Control
- Multi-year and multi-faceted plan

1. Building the Team

Core pillars:

1. Multi-disciplinary teams
2. Scientific Excellence
3. Scientific Communication
4. Scientific Discussion and Challenge
5. Success Culture
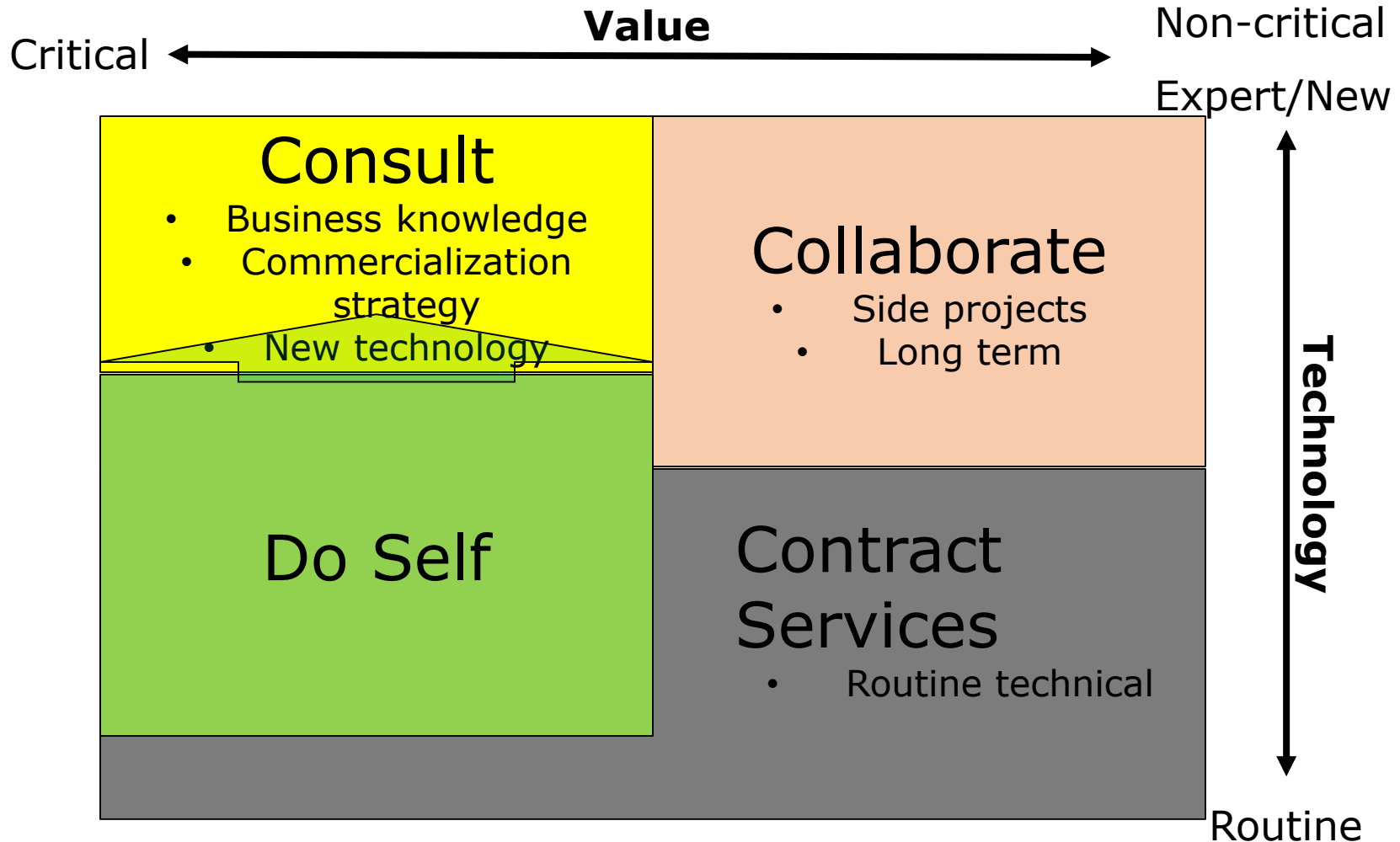6. Focus on Fundamental Understanding and Result Quality

The Team

2. Building Expertise

Leveraging on the Right Partnerships to have expertise in-house

Value

Critical ←——————————→ Non-critical

Expert/New

Consult
- Business knowledge
- Commercialization strategy
- New technology

Collaborate
- Side projects
- Long term

Do Self

Contract Services
- Routine technical

Technology

Routine

Do **not** outsource:
Experimental design
Result analysis

6-7 years of OP Genome research in Sime Darby

230

Genome sequences

200,000 **X** **>10,000**

Genome code
differences

tested palms

x 52 traits

1000 differences
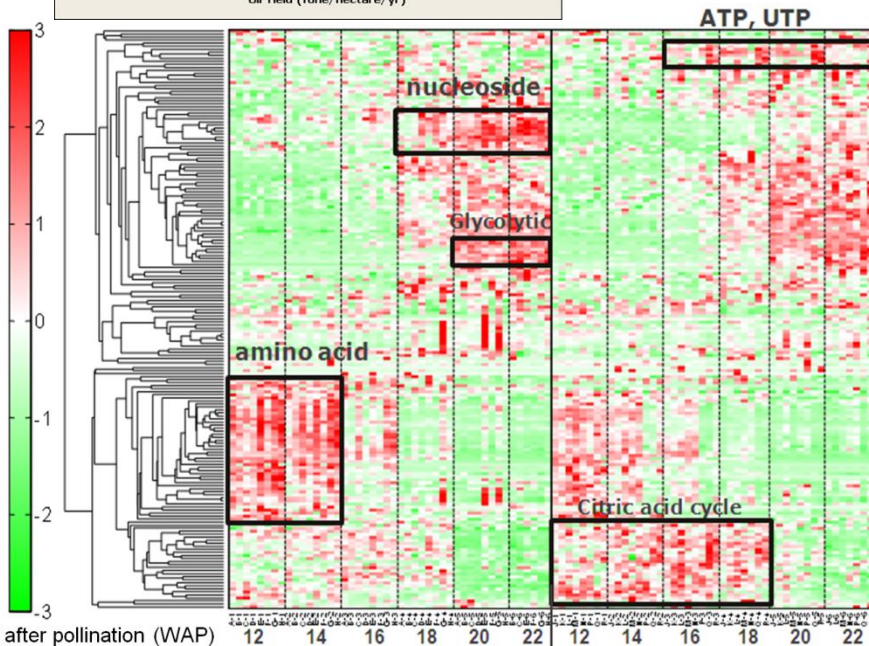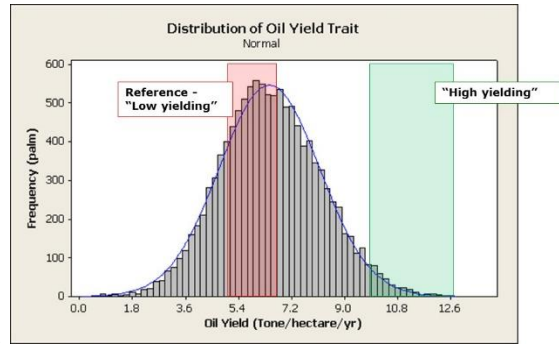for **Yield**

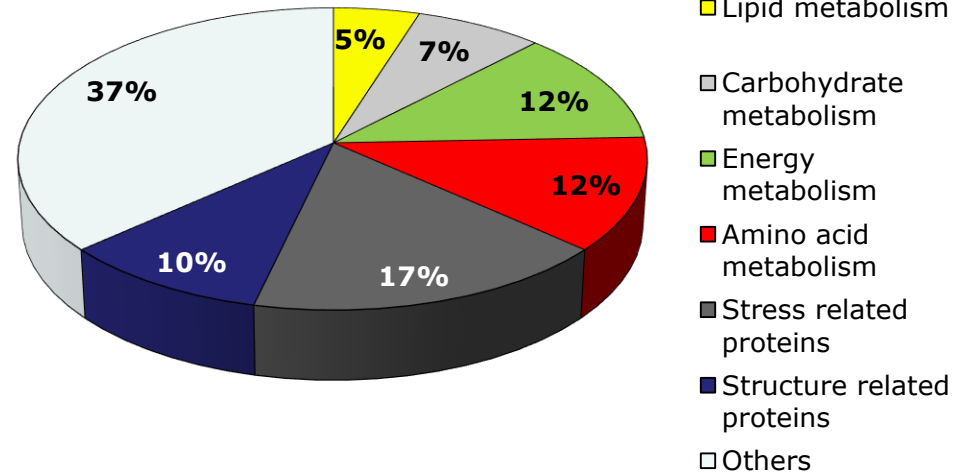# Genetic Test for Selection

Naturally Pollinated With No Change to the DNA
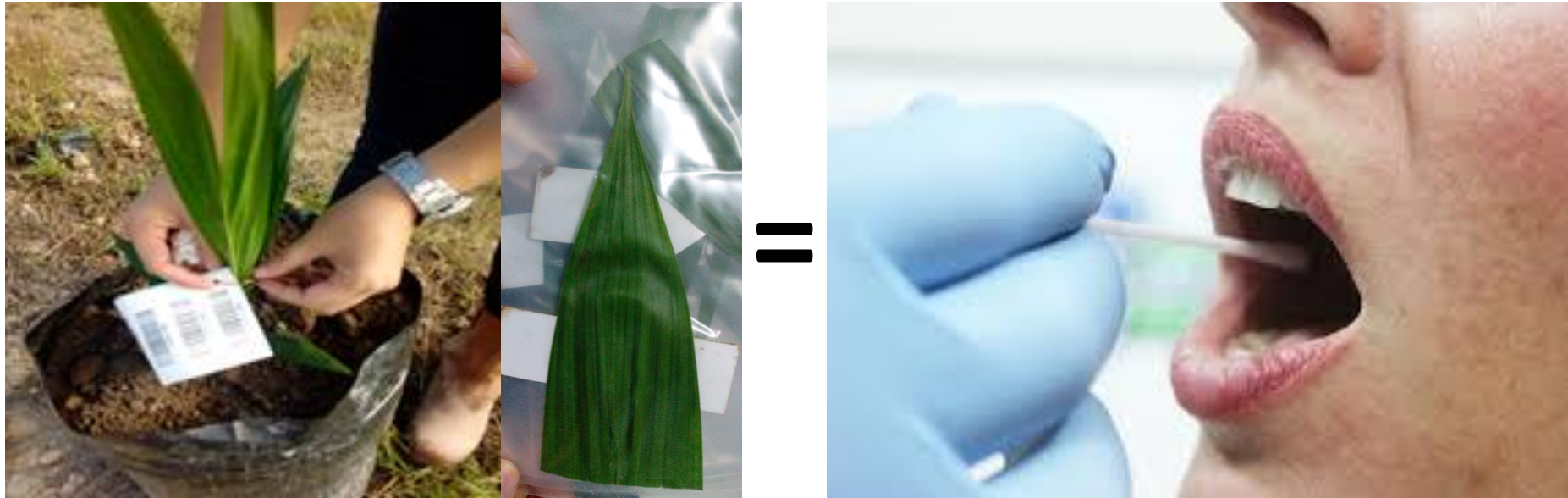
Select seedlings in the nursery

Increasing fundamental understanding

Omics analysis gave us insights into yield

Week 14 → Week 18 → Week 22

Mesocarp metabolite levels
  Teh *et al* PlosOne 2013
  Neoh BK et al JAFC 2013

Differentially expressed proteins

- ☐ Lipid metabolism
- ☐ Carbohydrate metabolism
- ☐ Energy metabolism
- ■ Amino acid metabolism
- ■ Stress related proteins
- ■ Structure related proteins
- ☐ Others

5% 7% 12% 12% 37% 10% 17%

Ooi EK *et al* Proteome Science 2015
Wong YC *et al* Microarrays 2014

+ >1000 differentially expressed genes

3. Building Infrastructure

From zero to 100 in 6 months

Best crosses Segregated and tracked to field

80,000 Seedlings sampled June 2015 to March 2016

80 million Genetic tests

New lab and IT systems Databases, analysis, Apps

Combined effort R&D-Operations Logistics, timing

FY2020/21 FFB difference apparent 48 months after planting

3. Building Infrastructure

1. **Sampling** (80,000 samples in 6 months)

2. **Labeling** (2 stages of nursery to field and beyond)

3. **Logistics** (collecting, storing, shipping, timing)

4. **Laboratory** (selection, procurement, expertise)

5. **Data storage and Processing** (80 million datapoints)

6. **Field tracking** (long-lasting, robust, simple)

Barcode Labeling and Mass Sampling: 80,000 seedlings sampled in 6 months

Pre-Nursery Labeling

1
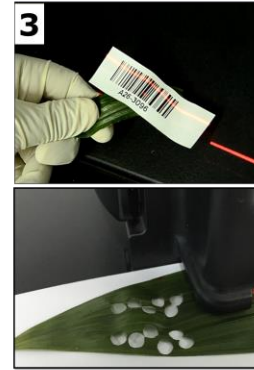
Throughput:
❖ 40,000 plants/month

3

PlantTrak LX
❖ Barcode scanner
❖ Leaf punching system
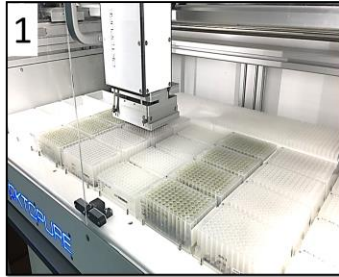
2

Main-Nursery Labeling

Labeling

Leaf Sampling

SNPLine High Throughput Genotyping:
80 million genotype data points processed in 10 months

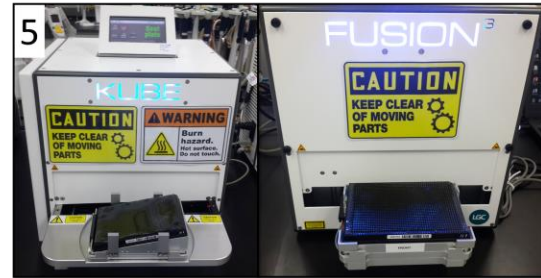

DNA extraction
❖ 30,000 samples/month

DNA dispensing
❖ Automated liquid handling
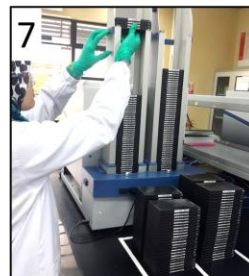
Plate drying
❖ Drying of plates in oven
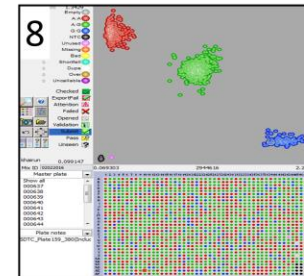
Assay dispensing
❖ 5,000,000 data points/month

Plate sealing
❖ Thermal and laser sealing of plates

DNA amplification
❖ Water-based thermal cycler

Plate reading
❖ Fluorescence based detection

Genotype scoring
❖ 5,000,000 data points/month

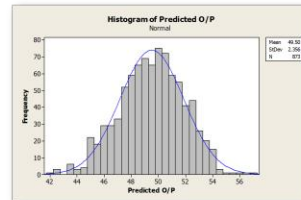# Genome Select and Breeding Database

GS Database
- ❖ >70,000 individual palms
- ❖ >80 millions genotypes records

Breeding Database
- ❖ >150,000 individual palms
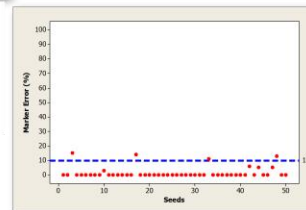- ❖ >7 million trait records

High Performance Server

GS Prediction
- ❖ Low repetition
- ❖ User friendly
- ❖ Oil yield prediction
- ❖ Combining ability prediction
- ❖ Elite palm selection

Legitimacy Test
- ❖ Seed purity assurance
- ❖ User friendly
- ❖ Auto report generation
- ❖ Accurate and fast

Data Retrieval
- ❖ Data co-location
- ❖ Data dissemination
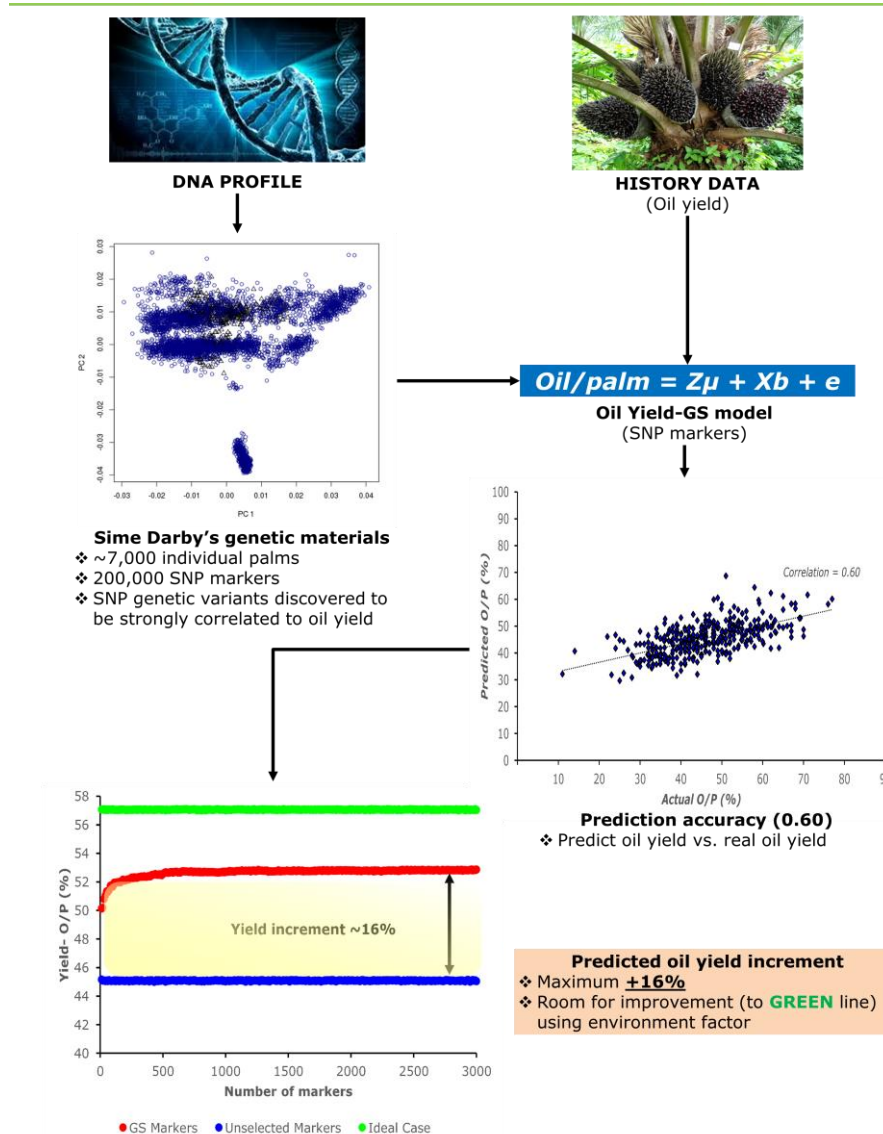- ❖ Auto error detection
- ❖ Auto report generation

Genome Select: Big Data Science

Marker identification and Genomic Selection requires billions of datapoints

DNA PROFILE

HISTORY DATA
(Oil yield)

$$Oil/palm = Z\mu + Xb + e$$

Oil Yield-GS model
(SNP markers)

Sime Darby's genetic materials
- ~7,000 individual palms
- 200,000 SNP markers
- SNP genetic variants discovered to be strongly correlated to oil yield

Correlation = 0.60

Prediction accuracy (0.60)
- Predict oil yield vs. real oil yield

Yield increment ~16%

Predicted oil yield increment
- Maximum **+16%**
- Room for improvement (to **GREEN** line) using environment factor

● GS Markers ● Unselected Markers ● Ideal Case

Segregated and Tracked Genome Select Nursery and Field Management

Genome select pre-nursery handling assisted by Sime Darby Seeds & Agricultural Services (SDSAS)

Main nursery handling assisted by Dusun Durian Estate

More than Scientists to Commercialise!
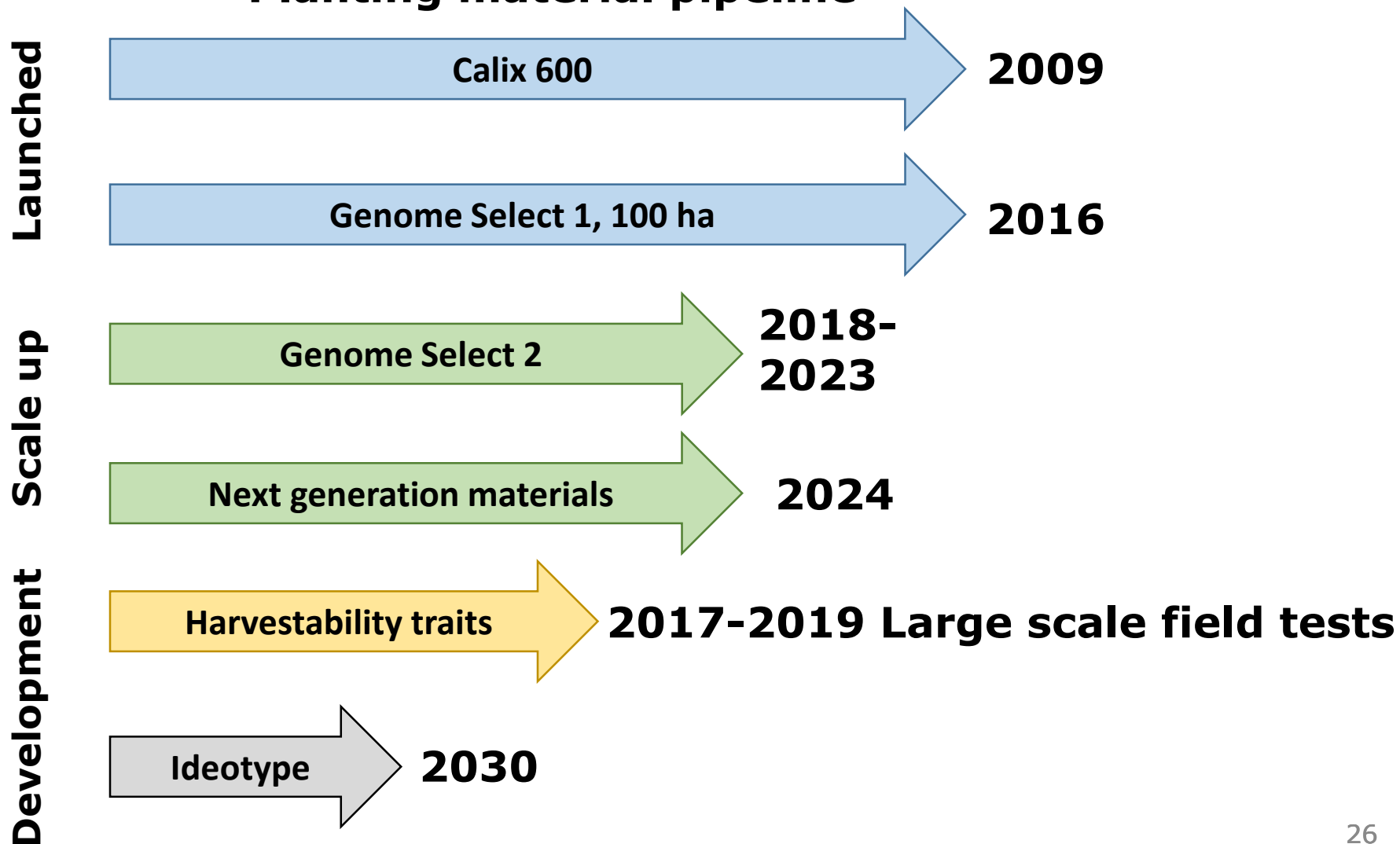
Communication is Key

Simple, direct ways to show value

Planting material pipeline

Launched
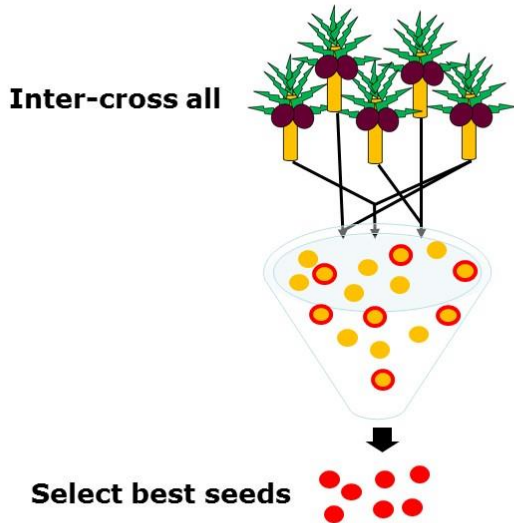
Calix 600 → **2009**

Genome Select 1, 100 ha → **2016**

Scale up

Genome Select 2 → **2018-2023**

Next generation materials → **2024**

Development

Harvestability traits → **2017-2019 Large scale field tests**

Ideotype → **2030**

What works best in the lab is unlikely to be the best method commercially

Seed Selection

Best commercial parents

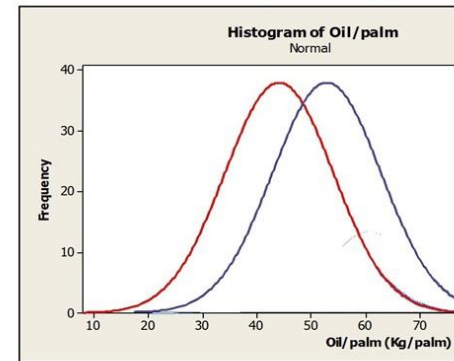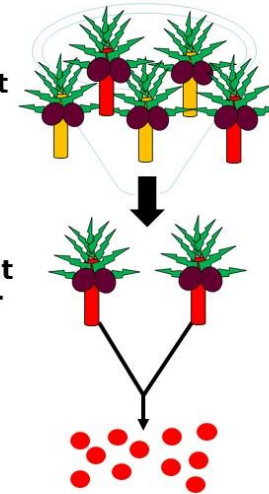Inter-cross all

Select best seeds
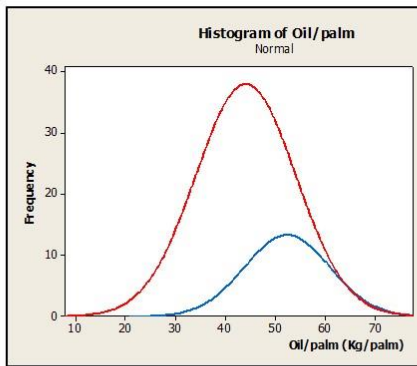
High yielding
Narrow CV
High cost

Parent Selection

Marker Selection of best Parents

Select best parent combinations for crossing

Low cost
High productivity
Higher CV
But, future improvement

Breeding, Molecular Biology, Tissue culture, IT, Operations, SAC

Thank you