



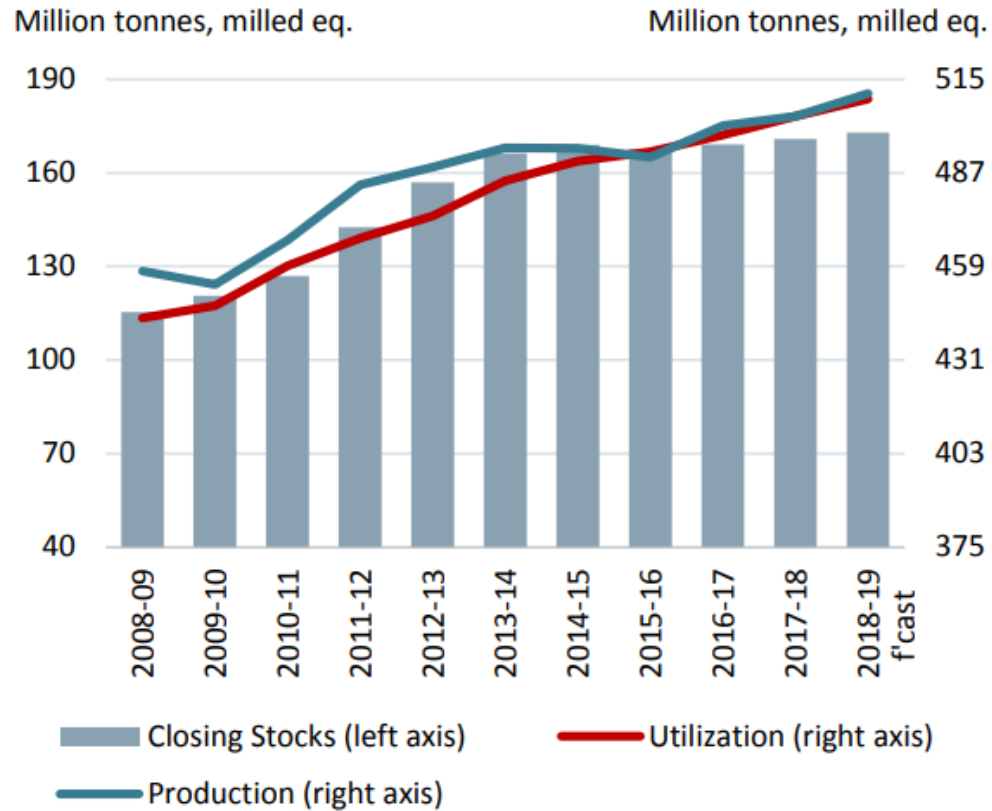
# Accelerating rice improvement in South Asia

Katherine Steele  
Bangor University





## Rice production, utilization and stocks



## Paddy production in Asia



# Most common targets for selection by rice breeders in Nepal, India and Pakistan

The two most common traits for improvement by marker assisted selection (MAS) are resistance to:

- **rice blast**

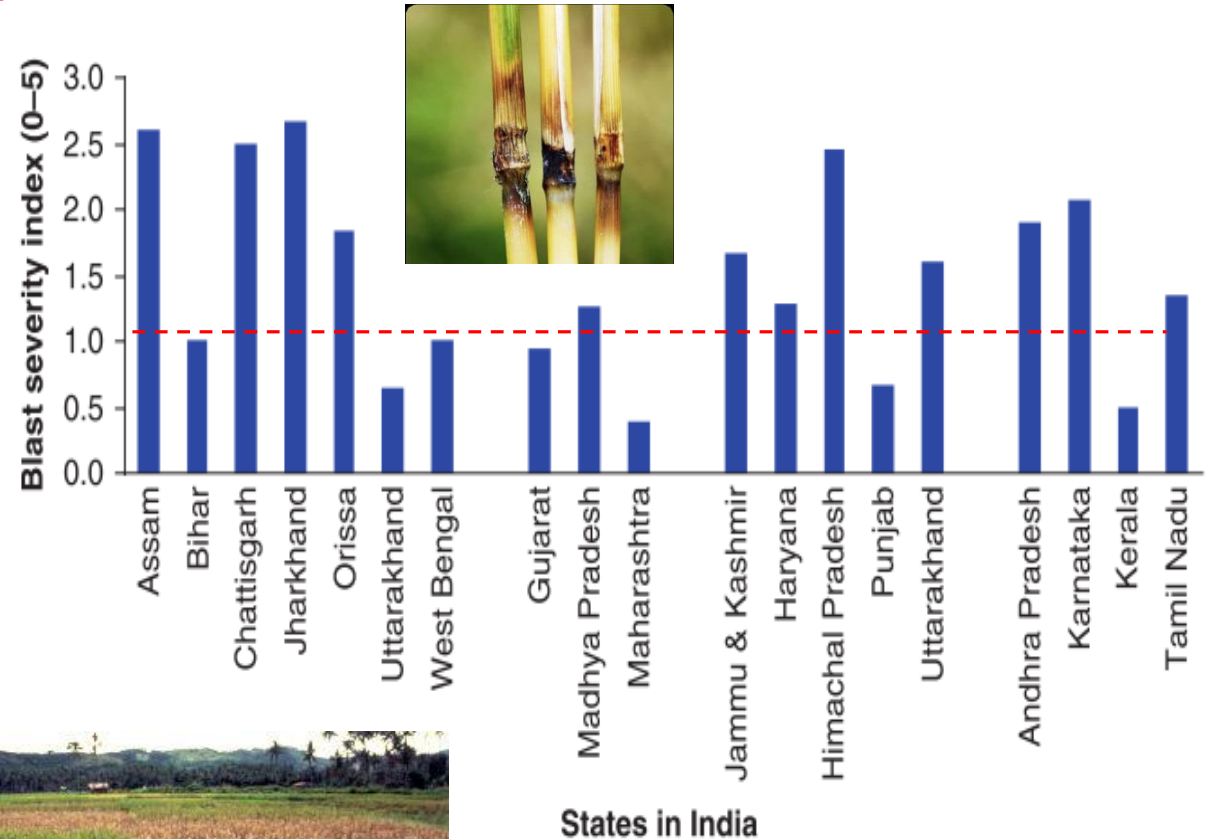
*Magnaporthe oryzae*

~100 race specific R loci known

- **bacterial leaf blight (BLB)**

*Xanthomonas oryzae* pv. *Oryzae*

~40 race specific R loci known



Rice BLB disease comparison with resistant lines  
Photo courtesy of IRRI Rice Knowledge Bank

# *Rice genomic variation*

- The complete rice genome is about 430 Mbp with
- ~39,045 protein coding sequences
- ~35% is repetitive sequences
- ~48,000 Simple Sequence Repeats (SSRs)

We estimate that any pairwise comparison between two rice varieties will differ by approximately

- **4,500** mapped SSRs
- **39,000** InDels
- **338,000** SNPs



Current microsatellites (SSRs)  
**18,000 available in rice**

KASP technology  
**2,055 available in rice**

Costly setup and labour intensive; high running costs

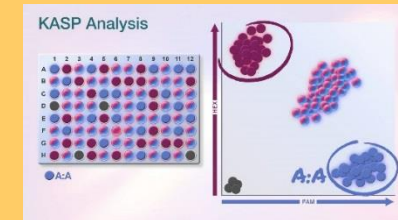
KASP from LGC Genomics is about three times cheaper than SSR genotyping (Semagn et al., 2014 Mol. Br. 33: 1-14)

Fewer possible SSR regions in the genome to convert into markers

Numerous SNPs and Indels exist in the genome; high conversion rate into markers

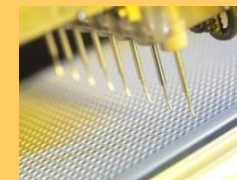
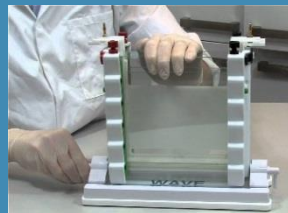
Safety hazards including use of toxic, carcinogenic and environmentally harmful acylamide gels or agarose gels with ethidium bromide staining

Safe to use. Unique chemistry and plate fluorescence to detect fluorescent signals



Low throughput, many repetitive steps, difficulty in automation

Automation gives extremely high throughput offering tailor made screens and a full service.



User sends leaf samples and LGC does the DNA extraction, KASP assays and provided genotypic scores in MS Excel

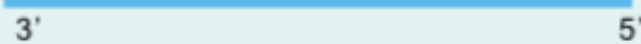
# KASP genotyping: assay components

## A) KASP Assay mix

Allele-specific forward primers:



Reverse primer:




## B) KASP Master mix




## C) DNA template (sample)



 Allele-1 tail FAM-labelled oligo sequence


 Allele-2 tail HEX-labelled oligo sequence

 Common reverse primer

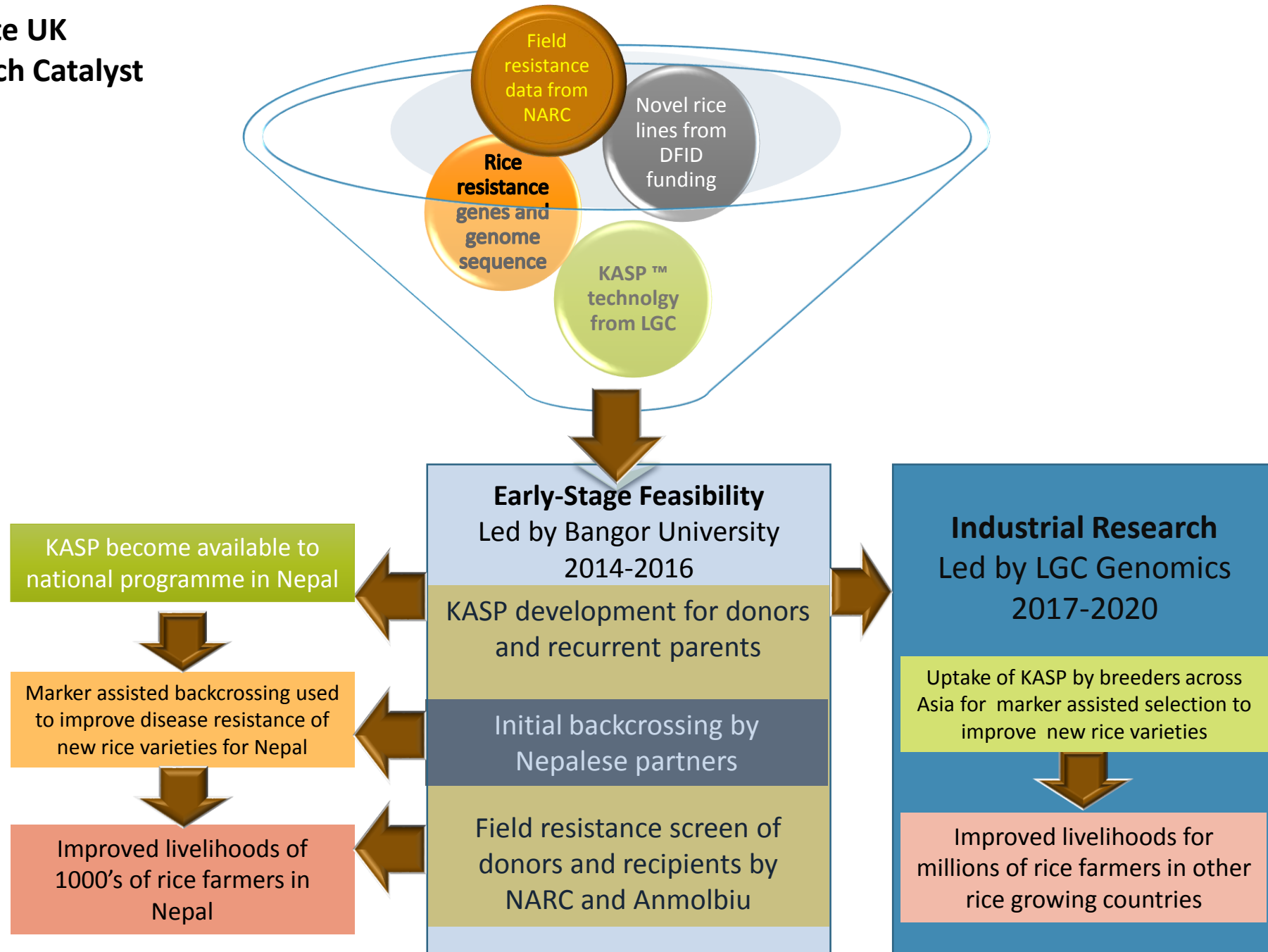
 FAM dye

 HEX dye

 Target SNP

 Quencher

# Innovate UK Agri-Tech Catalyst



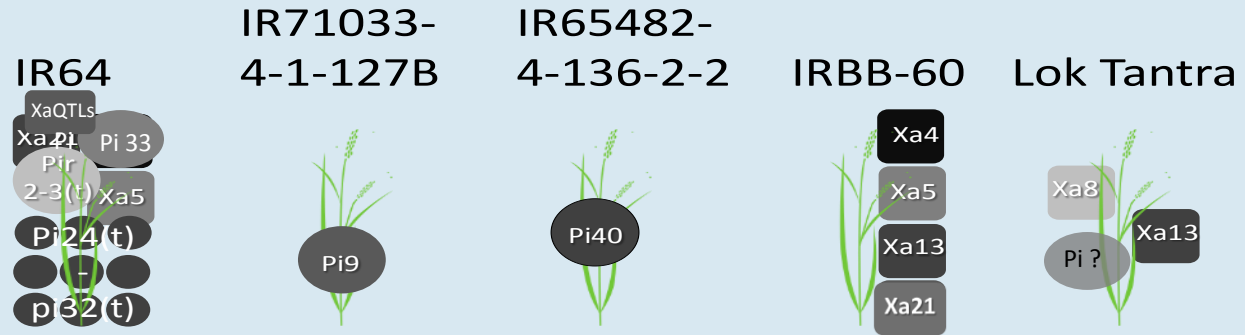


# *Aims of the Feasibility study*

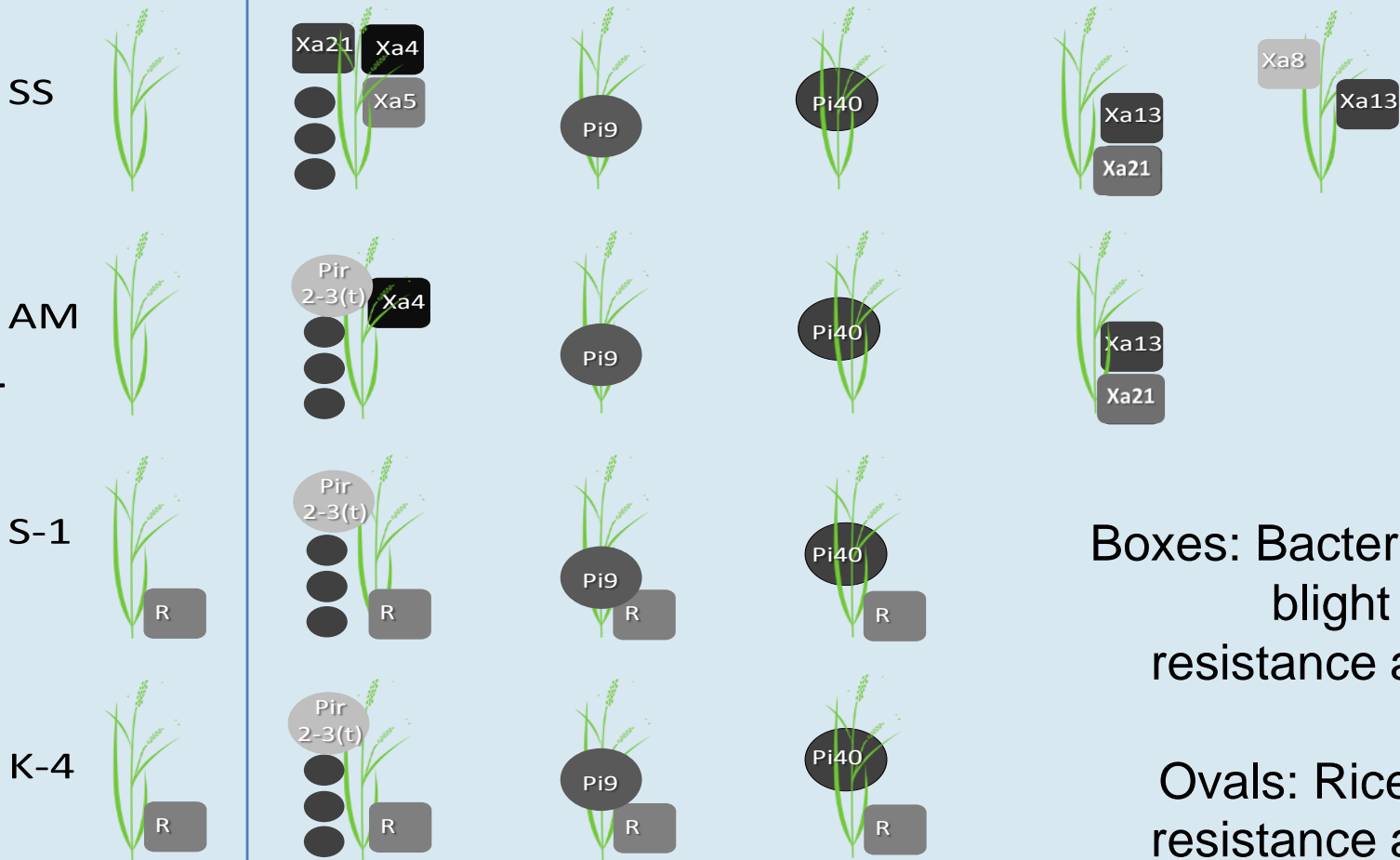
- Assess the potential of KASP technology for rice breeding
  - Identify variations in breeding lines (genome-wide)
  - Design KASP assays for identified variations
- Use KASP genotyping in Nepalese public sector breeding program



# Donor parents



## Recurrent parents



Boxes: Bacteria leaf blight (BLB) resistance alleles

Ovals: Rice blast resistance alleles

# Whole genome resequencing of rice lines

- Paired-end Illumina sequencing carried out on 9 indica rice lines
- Reads mapped against *Oryza sativa indica* reference genome
- Variants against *indica* reference sequence identified

| Rice line        | Mapping rate (%) | Genome coverage (%) | Gene region coverage (%) | Depth | Gene region depth |
|------------------|------------------|---------------------|--------------------------|-------|-------------------|
| IR64             | 92.2             | 89.2                | 94.2                     | 61.0  | 62.3              |
| IR71033          | 91.9             | 89.6                | 94.4                     | 55.2  | 55.2              |
| IR65482          | 90.8             | 88.9                | 94.0                     | 54.4  | 54.8              |
| Sunaulo Sugandha | 91.3             | 89.2                | 94.2                     | 56.8  | 58.3              |
| Anamol Masuli    | 90.9             | 88.9                | 94.1                     | 46.0  | 47.9              |
| Khumal-4         | 92.6             | 89.2                | 94.2                     | 55.2  | 55.6              |
| IRBB-60          | 92.8             | 89.4                | 94.4                     | 60.4  | 58.7              |
| Loktantra        | 93.8             | 88.7                | 93.9                     | 81.3  | 82.8              |
| Sugandha-1       | 92.4             | 89.3                | 94.3                     | 57.2  | 57.4              |

# Existing KASP assays

2,015 rice KASP assays available from LGC:

- Pariasca-Tanaka *et al.* (2015). Development of a SNP genotyping panel for detecting polymorphisms in *Oryza glaberrima/O. sativa* interspecific crosses. *Euphytica*, **201**, p. 67-78

1,491 assays align to the indica genome with 100% identity.

284-414 assays informative in pairwise crosses of the 9 lines used in this study.



# *KASP by Design*

- Customer submits a design sequence
  - Target variation indicated by square brackets
  - 50 bp flanking sequence either side of the target
  - Flanking sequence can incorporate ambiguous bases and/or InDels
    - Ambiguous bases indicated using IUPAC ambiguity codes
    - InDels indicated using Ns

CTTAGATCGACAGGTCTAAGAGCTGAAGAGCTAGCTGATTAAAGTCGAGC[S]  
AGCTGCTAGACGTCGCAGTCGACACAGCTAGCCTNNNACAAAGTCTCGTG

CTTAGATCGACAGGTCTAAGASCTGAAGAGCTAGCTGGCTAGCTGATTAA[ATAGACGTCGATCGT/-]  
AGCTGCTAGACGTCGCAGTCGACYCTGACGTCCTAGGACAAAGTCTCGTG

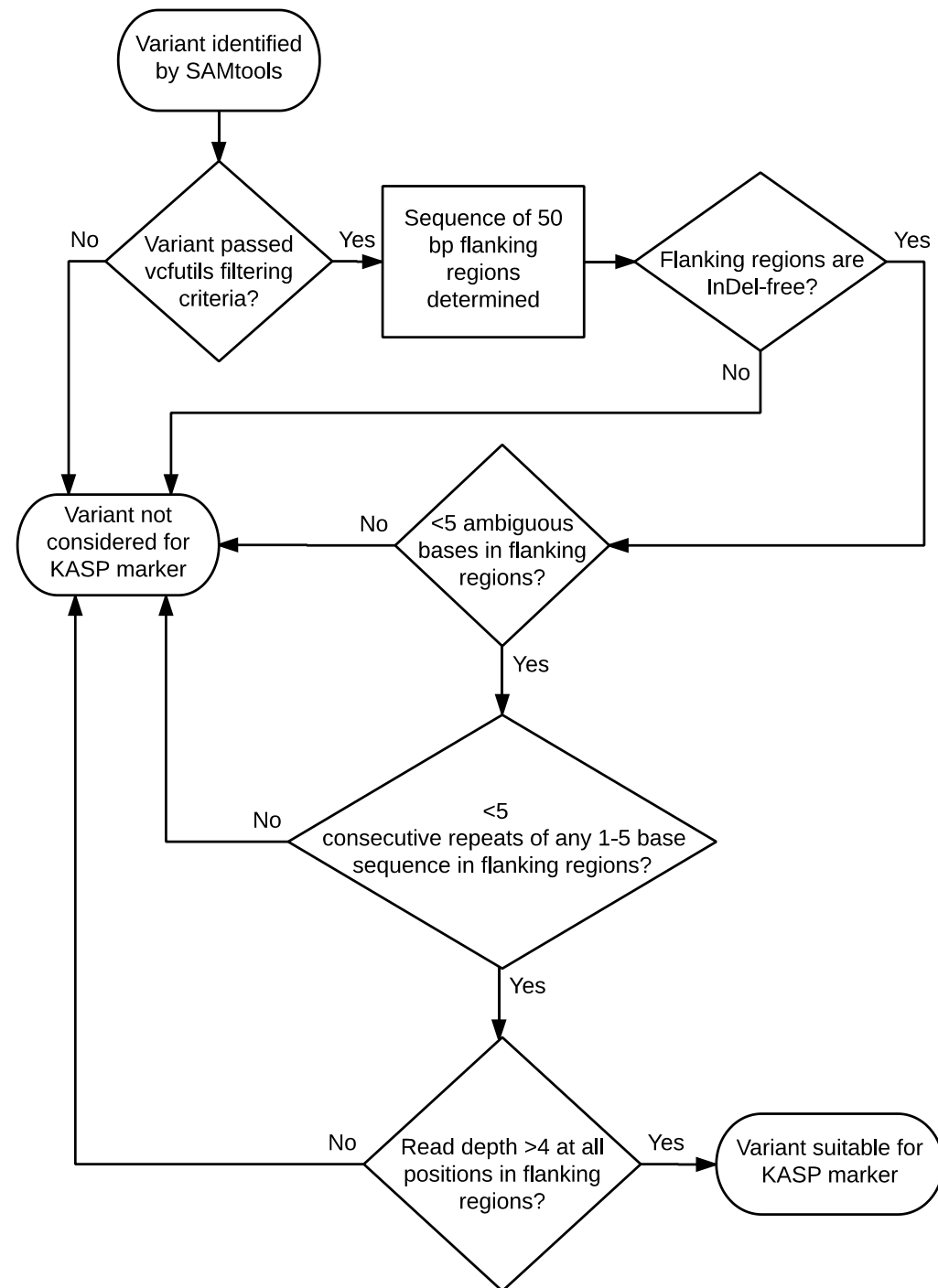
- LGC design assay using proprietary Kraken™ software system
  - 3 KASP primers specific to variation of interest
    - 2 allele-specific forward primers
    - 1 reverse primer

# Identifying potential KASP assays

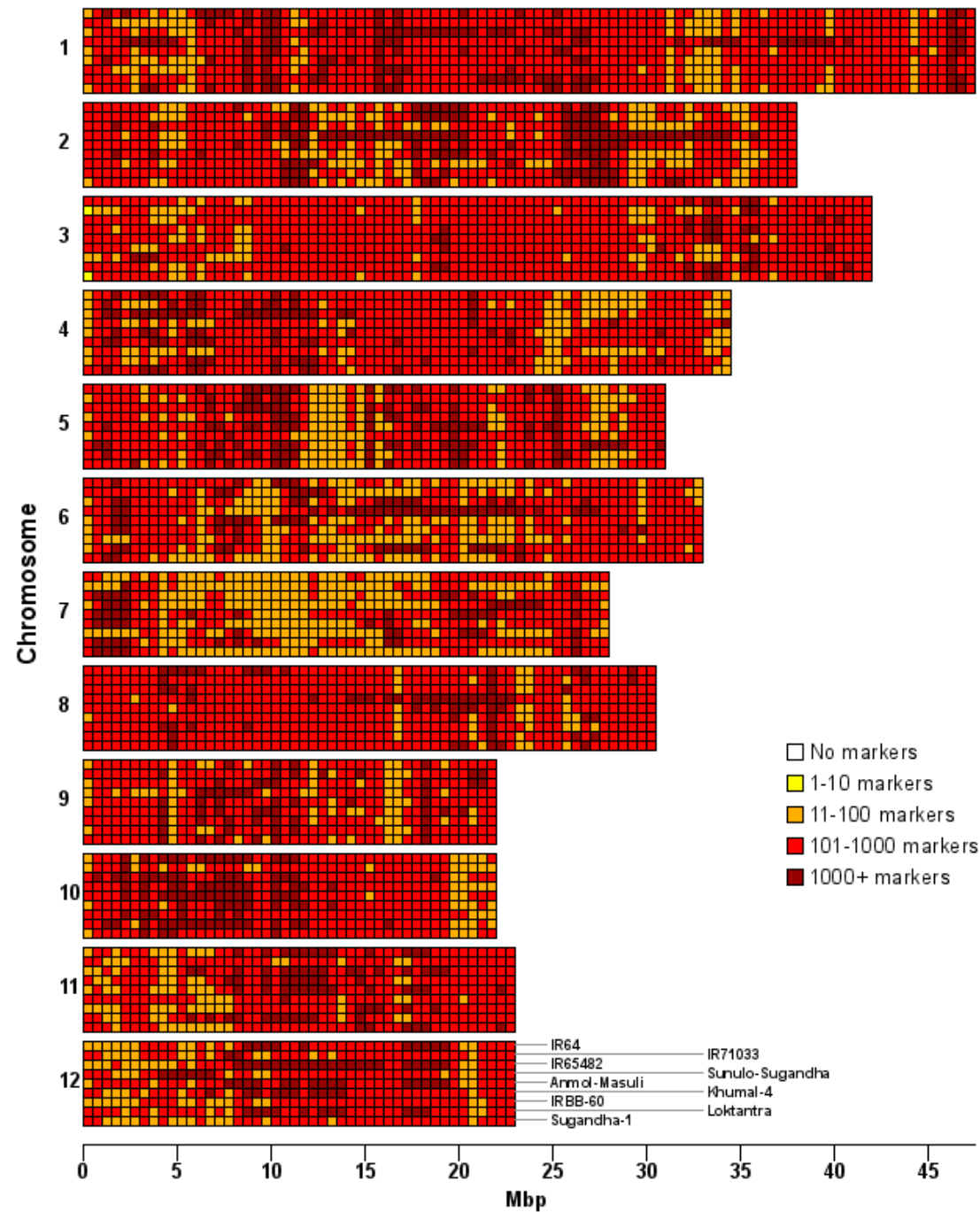
Not all variations are suitable for conversion to KASP assays.

Identified variations were filtered according to:

- evidence for variant
- InDels in flanking regions
- no. of ambiguous bases
- presence of repeat sequences
- read depth



|                  | IR64    | IR71033 | IR65482 | Sunaulo Sugandha | Anamol Masuli | Khumal-4 | IRBB-60 | Loktantra | Sugandha-1 | Indica |
|------------------|---------|---------|---------|------------------|---------------|----------|---------|-----------|------------|--------|
| IR64             |         | 361     | 413     | 456              | 377           | 511      | 382     | 492       | 441        | 480    |
| IR71033          | 245,367 |         | 419     | 453              | 470           | 434      | 345     | 453       | 386        | 377    |
| IR65482          | 355,518 | 322,602 |         | 503              | 488           | 473      | 430     | 442       | 469        | 490    |
| Sunaulo Sugandha | 444,337 | 418,294 | 511,006 |                  | 520           | 497      | 440     | 522       | 485        | 486    |
| Anamol Masuli    | 286,304 | 342,841 | 403,027 | 493,297          |               | 474      | 503     | 391       | 428        | 491    |
| Khumal-4         | 376,321 | 323,346 | 397,553 | 481,381          | 387,264       |          | 467     | 473       | 426        | 433    |
| IRBB-60          | 328,293 | 273,578 | 397,538 | 407,849          | 404,343       | 369,498  |         | 452       | 441        | 392    |
| Loktantra        | 362,689 | 346,699 | 385,651 | 460,348          | 332,649       | 391,608  | 378,459 |           | 407        | 496    |
| Sugandha-1       | 328,829 | 274,529 | 385,646 | 465,745          | 356,552       | 330,285  | 345,187 | 361,699   |            | 421    |
| Indica           | 388,347 | 309,369 | 447,904 | 459,229          | 433,769       | 369,572  | 316,757 | 434,001   | 337,913    |        |





# Outcomes of feasibility study


- Identified 1.3 million potential KASP assay designs:
  - 92,500 potential functional markers
  - 3.1 per 1 kb in crosses analysed
  - 370,000 per cross on average
- Validated 39 novel KASP assays by genotyping progeny from a range of crosses
- Utilisation of new and existing KASP to improve efficiency of public sector breeding in Nepal
- Published findings and software for KASP design generation:  
Steele *et al.* (2018). Accelerating public sector rice breeding with high-density KASP markers derived from whole genome sequencing of *indica* rice. *Molecular Breeding*, **38**, 38.



# Innovate UK Industrial Phase project

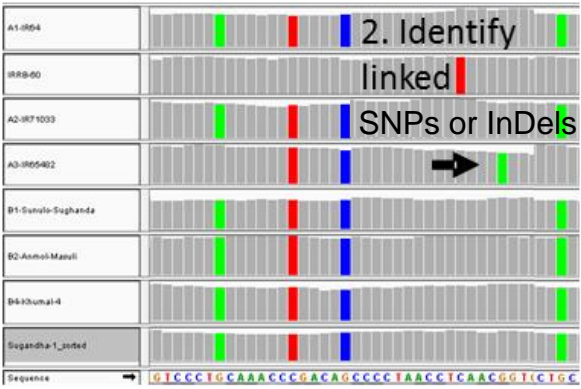
LGC, Bangor University, Agri-Epi Centre (UK); NIBGE (Pakistan), SKUAST (India), APL & NARC (Nepal)

Threats to rice production and food security



Bacterial leaf blight (BLB)    Blast    Bakanae    Brown spot    Abiotic stresses: drought and salinity

Applying KASP technology




|                    |   |   |  |  |  |
|--------------------|---|---|--|--|--|
| A1-0864            |   |   |  |  |  |
| IR8-00             |   |   |  |  |  |
| A2-9871003         |   |   |  |  |  |
| A3-0865402         |   |   |  |  |  |
| B1-Sunolo-Sughanda |   |   |  |  |  |
| B2-Aesho-Manali    |   |   |  |  |  |
| D4-Khumai-4        |   |   |  |  |  |
| Sugandha-1_sorted  |   |   |  |  |  |
| Sequence           | → | G T C C C T G C A A A C C C G A C A G C C C C T A A C C T C A A C G G T C T C G C |  |  |  |

1. Identification target genes in genomes of donor parents

2. Identify linked SNPs or InDels


3. Make crosses



4. Screen populations with KASP assays and select resistant and improved progeny

➔ **FOOD SECURITY**

Reduced yield penalties in locally adapted varieties – and reduced pesticide use



More efficient selection of disease resistant and better adapted varieties is likely to lead to economic benefits in developing country agriculture



# Nepal



**NARC - Biotechnology Division and National Rice Research Programme**



Far Western

Mid Western

Western

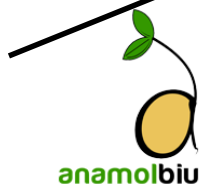
Central

Eastern

NARC Biotechnology Division  
Khumaltar  
**Kathmandu**  
BLB and blast screening

APL  
Surunga  
**Jhapa District**  
BLB Nursery  
Blast screening

Bhairawa  
**Rupandehi**  
BLB Nursery



**Anamolbiu Pvt Ltd (APL)**

Bharatpur  
**Chitwan District**

NARC - NRRP  
Janakpur  
**Hardinath**  
BLB and Blast screening

(1) Kathmandu  
(2) Bhaktapur  
(3) Lalitpur

# Aims

- Build on work carried out in feasibility study to make new KASP assays available for all rice breeders
- Generate KASP designs that can be utilised in a wide variety of rice lines
- Identify suitable SNPs and InDels to SSR markers currently used in rice breeding programs
- Validate 4,000 assays by genotyping lines from partners breeding programs
- Create database of new assays enabling customers to order them direct from LGC Genomics



## *Designing assays for the wider market*

A: CACCTCGGTGAATCAGGGCAGGAGTCATATTAGGCATACACT

B: CACCTCGGTGTATCAGGGCACAGGCATATTAGGCATACACT



CACCTCGGTGWATCAGGGCA [ S ] AGKCATATTAGGCATACACT

# Designing assays for the wider market

A: CACCTCGGTGAATCAGGGCAGAGTCATATTAGGCATACACT

B: CACCTCGGTGTATCAGGGCACAGGCATATTAGGCATACACT



CACCTCGGTGWATCAGGGCA [ S ] AGKCATATTAGGCATACACT

---

A: CACCTCGGTGAATCAGGGCAGAGTCATATTAGGCATACACT ✓

B: CACCTCGGTGTATCAGGGCACAGGCATATTAGGCATACACT ✓

C: CACCTCGGTGTATCAGGGCACAGACATATTAGGCATACACT ✗

D: CACCTTGGTGTATCAGGGCAGAGGCATATTAGGCATACACT ✗

# Designing assays for the wider market

A: CACCTCGGTGAATCAGGGCAGAGTCATATTAGGCATACACT

B: CACCTCGGTGTATCAGGGCACAGGCATATTAGGCATACACT



CACCTCGGTGWATCAGGGCA [ S ] AGKCATATTAGGCATACACT

---

A: CACCTCGGTGAATCAGGGCAGAGTCATATTAGGCATACACT ✓

B: CACCTCGGTGTATCAGGGCACAGGCATATTAGGCATACACT ✓

C: CACCTCGGTGTATCAGGGCACAGACATATTAGGCATACACT ✗

D: CACCTTGGTGTATCAGGGCAGAGGCATATTAGGCATACACT ✗



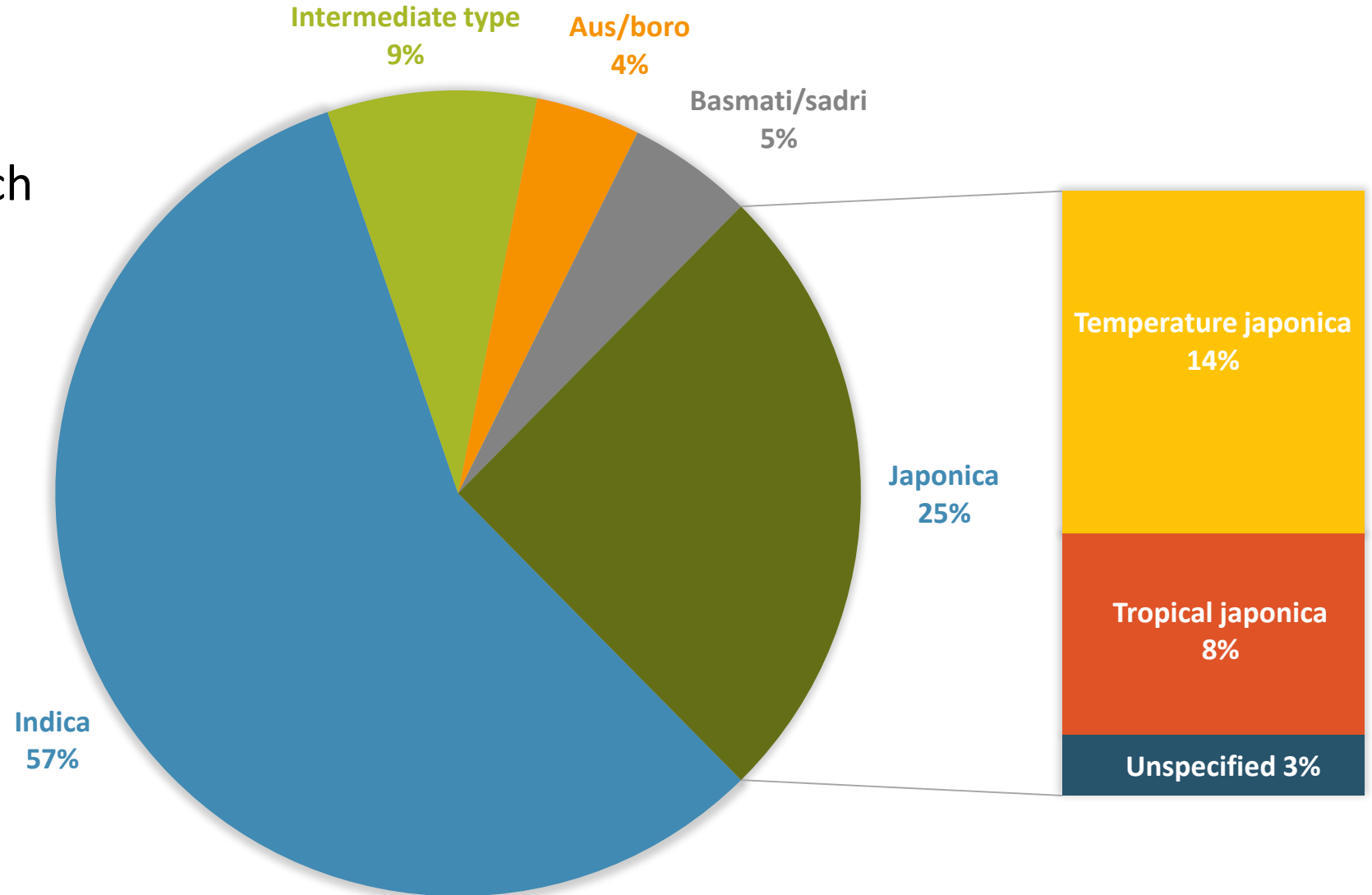
CACCTYGGTGWATCAGGGCA [ S ] AGDCATATTAGGCATACACT



# Utilising 3,000 Rice Genomes Project data

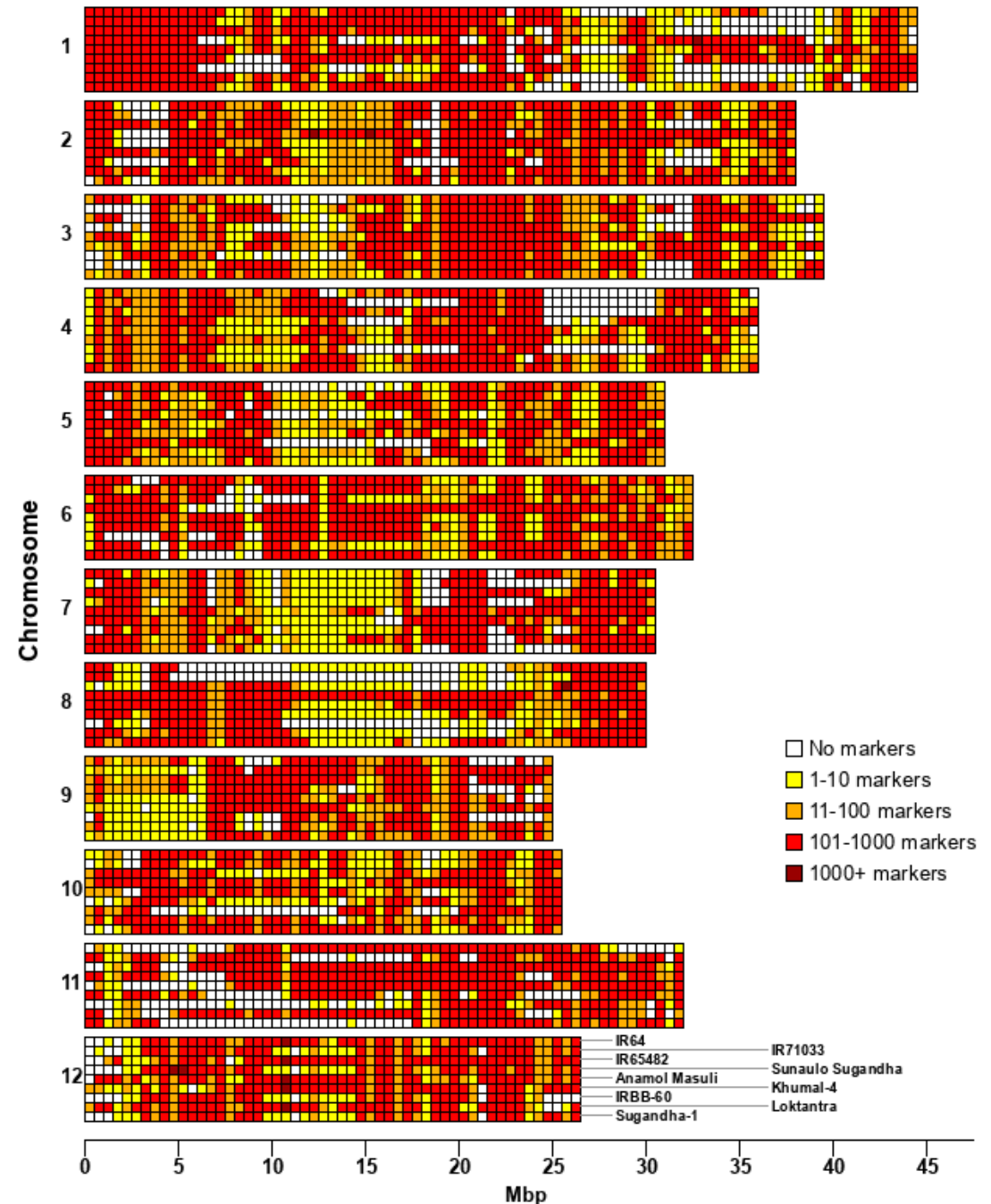
Wang et al.  
*Nature* 557.7703 (2018): 43.

- 119 lines selected:
  - at least one from each country of origin
  - mixture of subtypes
- Sequences aligned against *indica* reference genome
- Variants called
- KASP designs generated



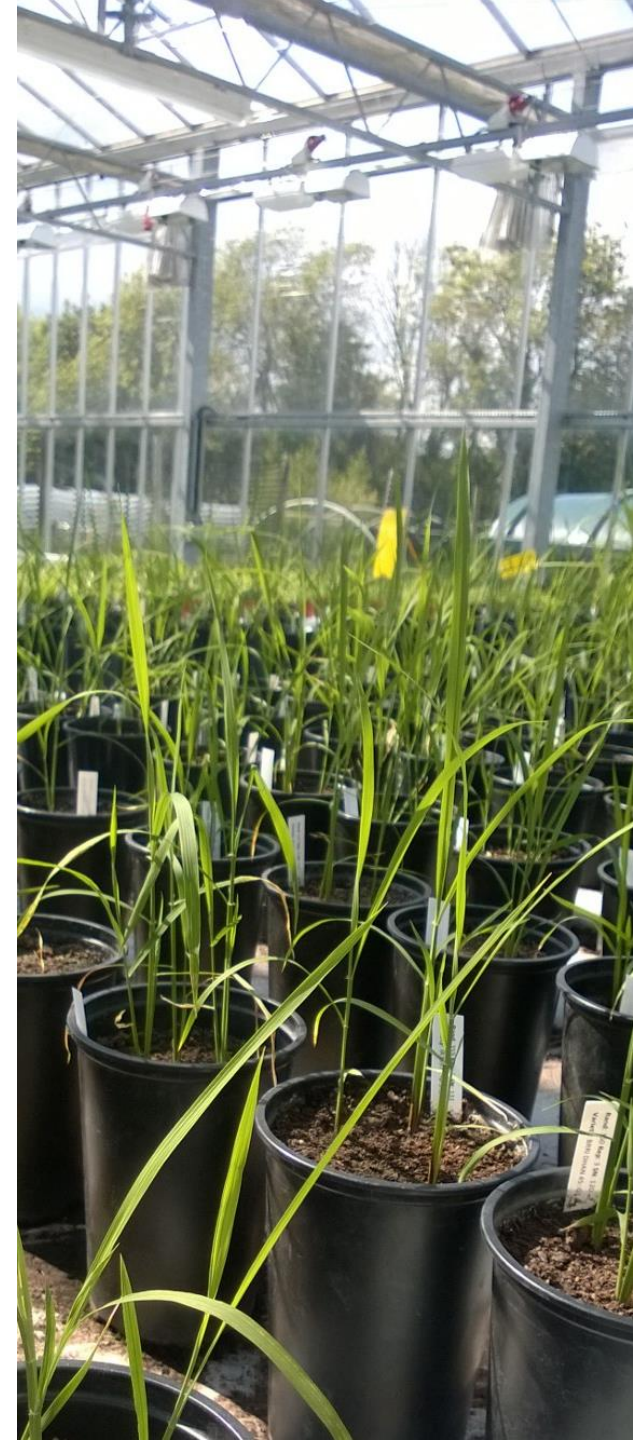
# Resulting KASP assay designs

- Designs based on 134 rice lines:
  - 119 from 3,000 Genomes Project
  - 13 additional resequenced lines including 9 from the feasibility study
  - *Indica* & *japonica* reference genome sequences
- 1,208,551 assay designs identified:
  - Target alleles found in  $\geq 90\%$  of lines tested
  - Align uniquely against both *indica* and *japonica* reference genomes



# *SSR marker position identification*

- Primer sequences of markers obtained:
  - 19,306 SSR marker primer pairs downloaded from Gramene database
  - 140 additional marker primer pairs (mostly SSRs) obtained from literature search of relevant rice breeding publications
  - 18,053 unique primer pairs identified
- Sequences BLASTed against *indica* and *japonica* references:
  - Unique alignment positions identified for 17,377 primer pairs
    - 30 *indica* only
    - 2,203 *japonica* only



# Making the conversion more convenient:

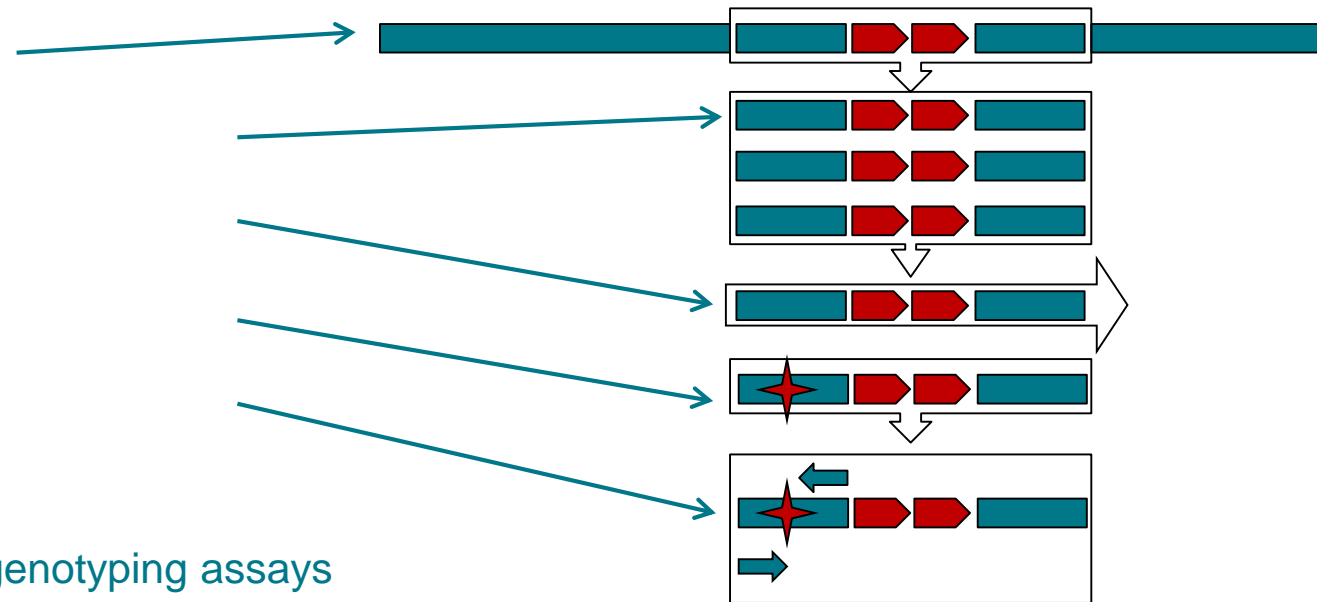


- How do you convert trait specific SSR's to SNP markers:

- DNA Extraction
- Long range PCR (3KB total)
- Next Generation Sequencing
- Data analysis & SNP scoring
- Assay design & genotyping

- What to do?

- Convert SSRs to KASP SNP genotyping assays



# *Future application in Basmati authentication?*

- The UK's Rice Association Code of Practice was updated in 2017 to include 26 new varieties that have been recently bred in India and Pakistan
- By screening them with thousands of KASP we can develop a suitable set of KASP assays for a screening test to use in authentication



# Acknowledgements

- Innovate UK Agri-Tech Catalyst and DFID for funding
- Dr Dominique Fauvin (LGC Genomics)
- Dr Darshna Vyas (LGC Genomics)
- Prof. John Witcombe (Bangor University)
- Dr Mark Quinton-Tulloch (Bangor University)
- Dr M. Arif (NIBGE, Pakistan)
- Resham Amgai (NARC, Nepal)
- Rajeev Dhakal, Anju Pandey and Mahesh Sharma (APL, Nepal)
- Freddie Reed (Agri-Epi Centre)

