Site-nonspecific mutation

- Chemical, physical, and insertional mutagen
  - ✓ How do those mutagens create mutations?
  - ✓ What are characters of the induced mutations
  - ✓ How to use the induced mutations to study gene's function
    - Reverse genetics
    - Forward genetics
  - ✓ How to create gene-index catalogue for induced mutants (locate mutation sites for all mutants)

Gene-indexed catalogue of physical and chemical mutagens induced mutants

- Whole genome sequencing
- Exome capture sequencing
  - Exome is all exons of protein coding genes in a genome
  - Exome capture sequencing is a technique for sequencing all exons of the protein coding genes in a genome

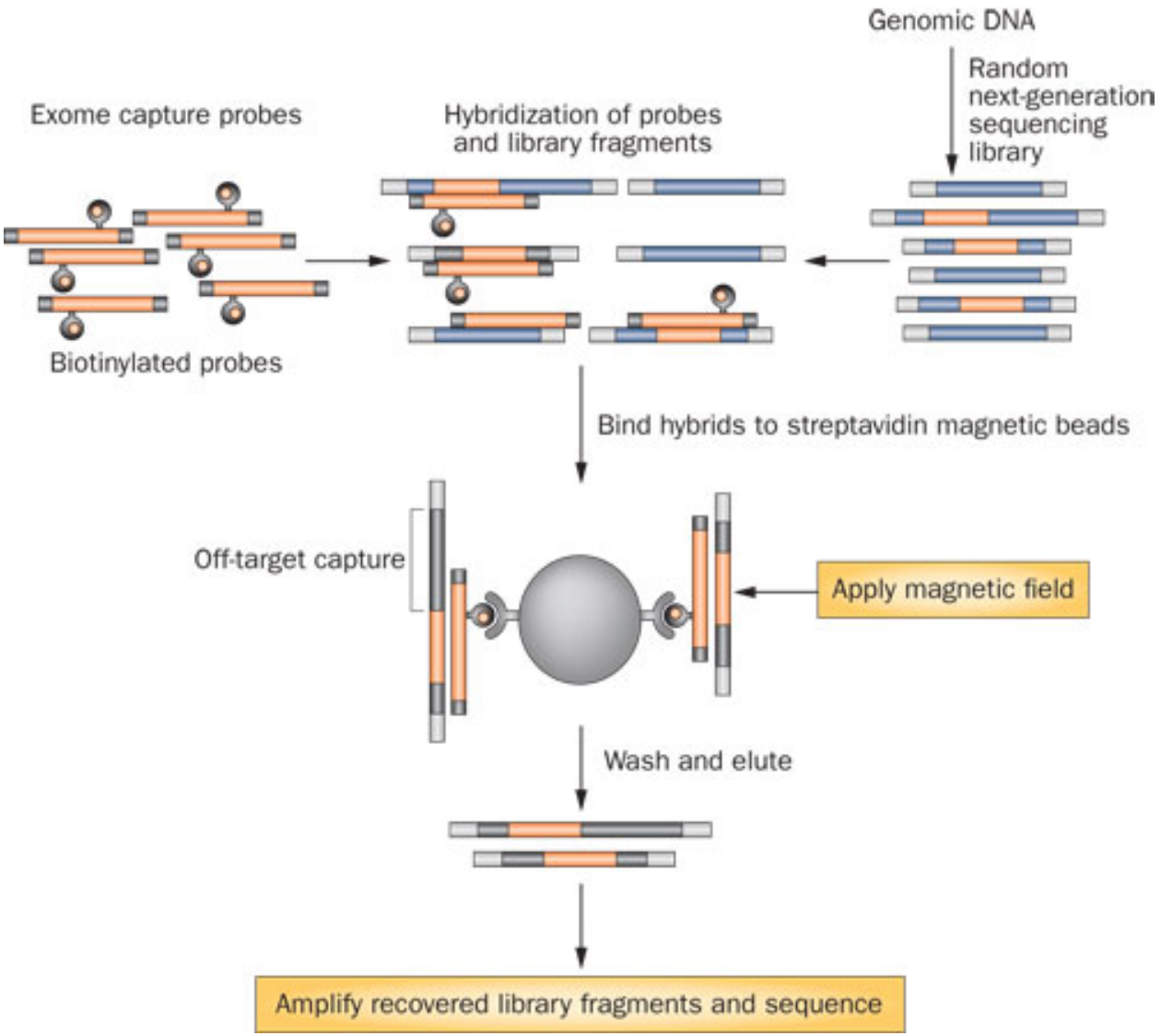
Whole-genome sequencing (WGS) or exome capture sequencing (ECS)

$2000/lane of Illumina HiSeq2000, 200 million reads of SE-100

Species	Analysis	Size of Target (Mb)	Reads for 20× Coverage (Million)	Capture Cost ($)	Total Cost ($)
Most plant species	**ECS**	**40**	**20**	**40**	**307**
Arabidopsis	WGS	120	12	0	180
Rice	WGS	380	38	0	507
Tomato	WGS	900	90	0	1,200
Maize	WGS	~2,300	230	0	3,067
wheat	WGS	~15,960	1,596	0	21,280

Henry et al., The Plant Cell, 2014

Exome capturing



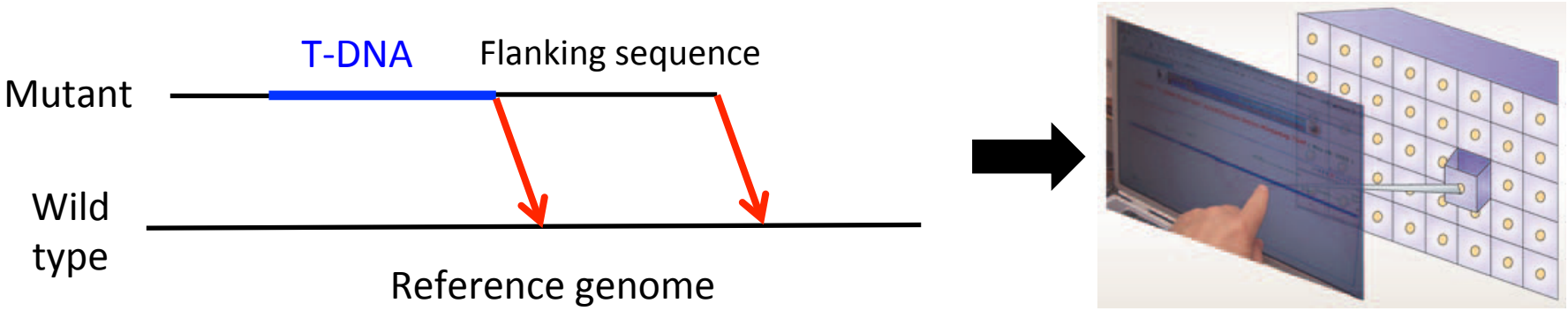
Exome capture sequencing of wheat EMS-induced mutants

- A wheat exome capture platform was developed and used to sequence the coding regions of 2,735 wheat mutants
- The targeted exon space in wheat was selected
  - An 84-Mb exome capture assay including overlapping probes covering 82,511 transcripts
- Over 10 million mutations were identified from the 2,735 mutants
- On average, 2,705-5,351 mutations per mutant, 35–40 mutations per kb

Krasileva et al., PNAS, 2017

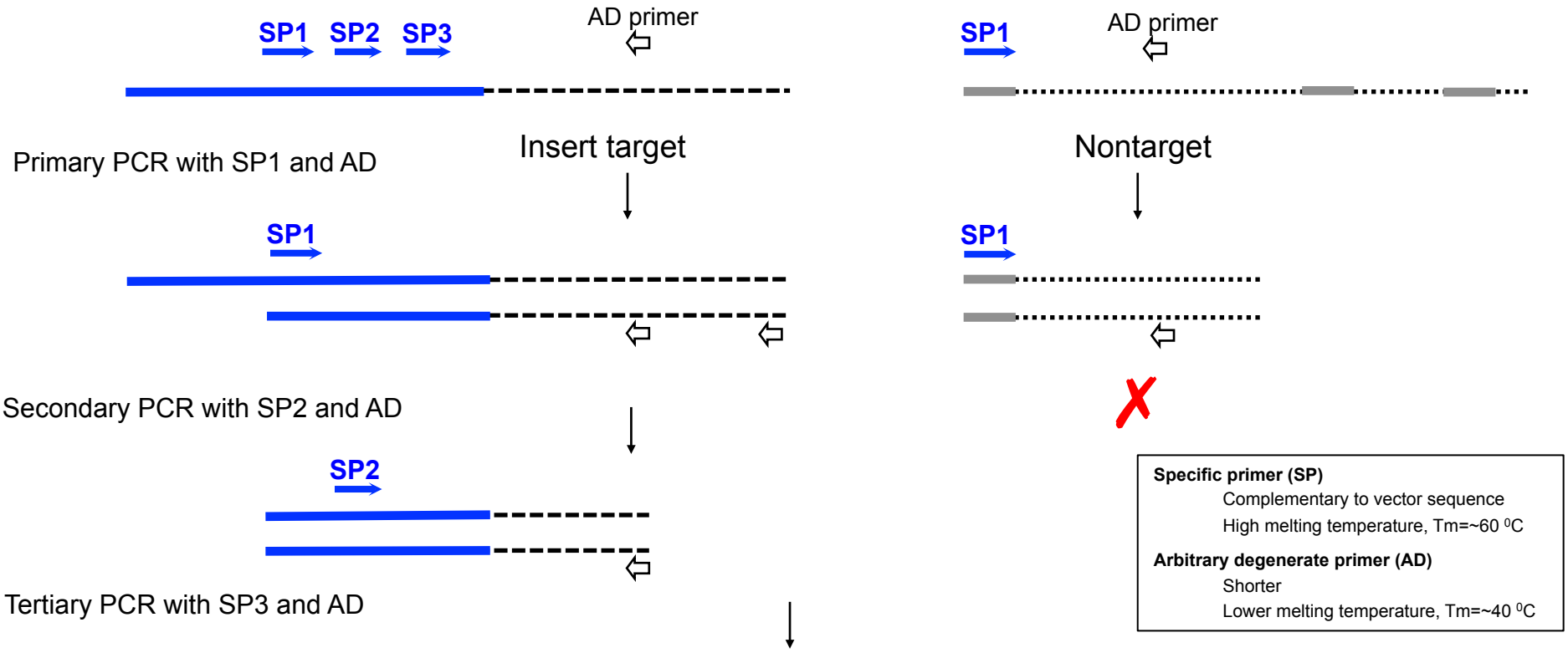
Gene-index catalogue for T-DNA insertional mutants

- Hundreds of thousands of T-DNA insertional mutants available in Arabidopsis
- A method to locate T-DNA inserts in all mutants is to get flanking sequences of the insertion site and then align the flanking sequences to reference genome



Thermal asymmetric interlaced PCR (TAIL-PCR) to obtain flanking sequences

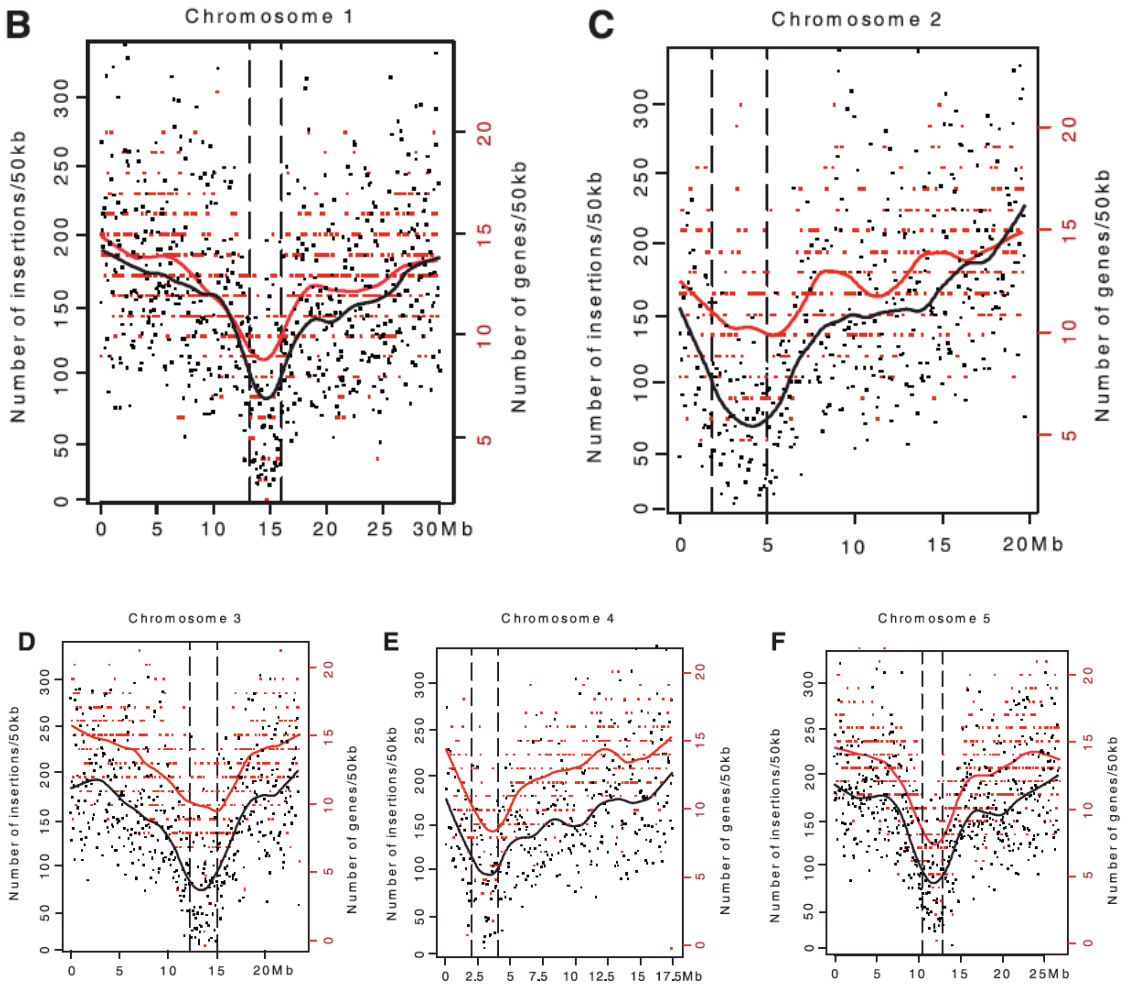
TAIL-PCR: Three PCR reactions are carried out sequentially to amplify target sequences using nested T-DNA-specific primers (SP1-3) on one side and an AD primer on the other side



Liu et al., Plant Journal 1995

Example: A gene-indexed catalogue of insertion mutants in Arabidopsis

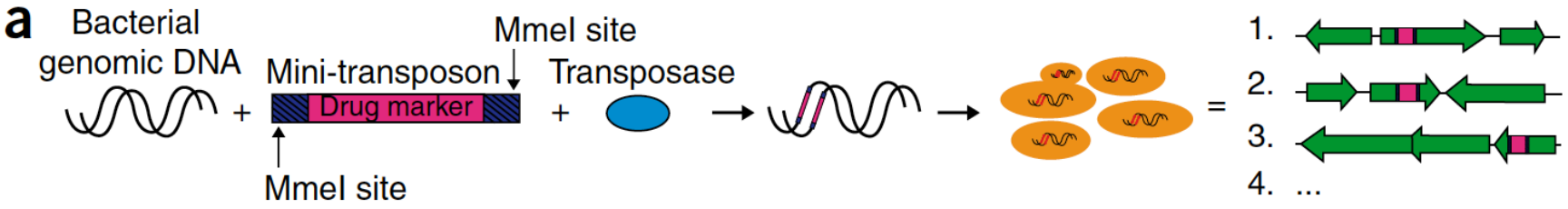
- Screened 127,706 mutants using TAIL-PCR
- 88,122 insertion sites identified in 21,799 genes
- Fewer T-DNA integration events were consistently observed in regions surrounding each of the five centromeres



Alonso et al., Science 2003

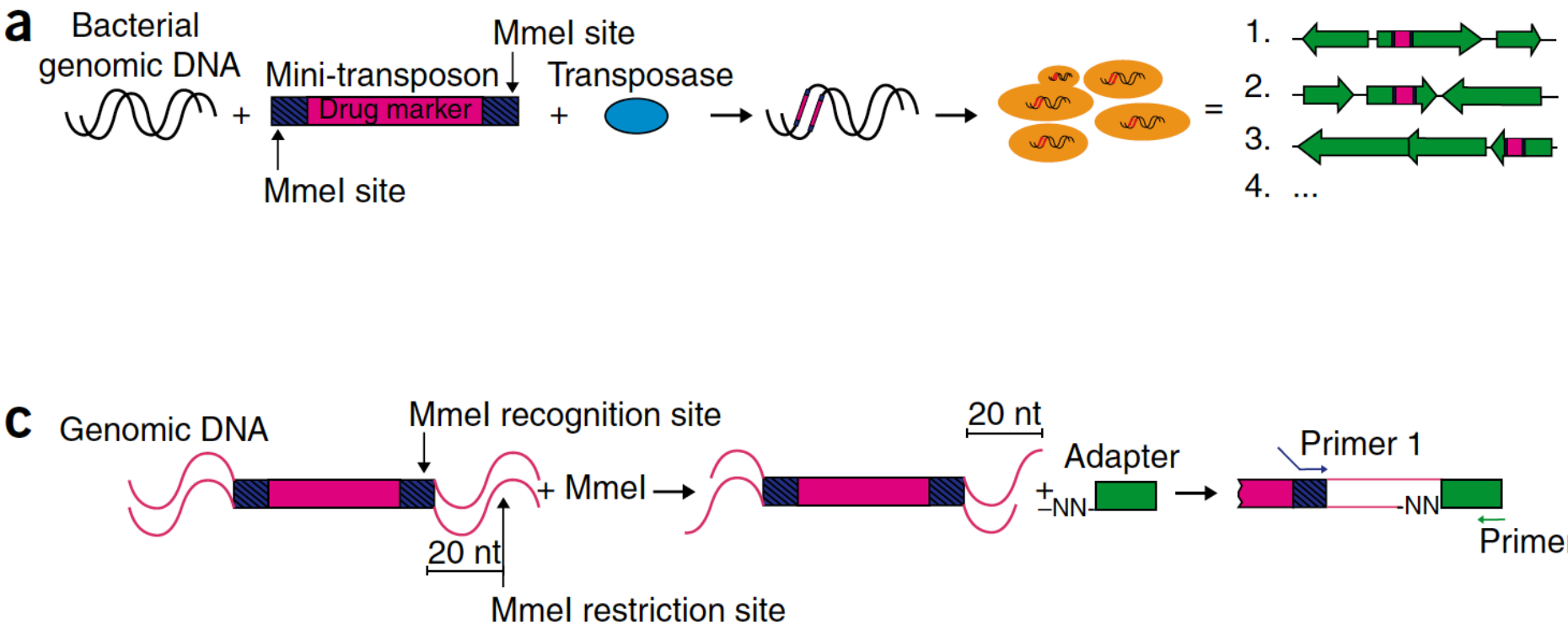
Create an insertional mutation library of *S. pneumoniae* using mini-transposon

- A gene disruption library is constructed by first transposing the mini-transposon into bacterial genomic DNA **in vitro** and then transforming a bacterial population with the transposed DNA
- The transpose inserts randomly in the genome, requiring only a TA dinucleotide at the insertion site



van Opijnen et al., Nature methods 2009

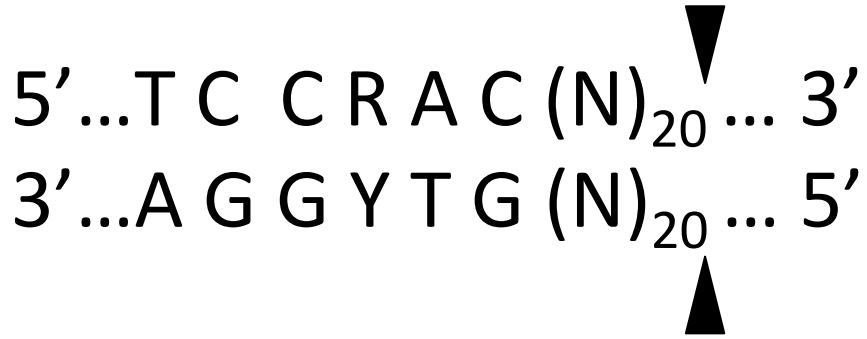
Locate insertional sites in an insertional mutation library of *S. pneumoniae*



van Opijnen et al., Nature methods 2009

Restriction enzymes recognize specific sequence

- Restriction enzymes is a protein that can cleave DNA or RNA into fragments
- Restriction enzymes recognize a specific sequence and cut the DNA within, near or remote from the specific site
- For example, *MmeI* recognizes a specific sequence of 6 bp



What need to know for final exam

- Can describe exome capture sequencing
- Can describe how to use TAIL-PCR to obtain flanking sequences of inserts and locate T-DNA insertional sites for a large number of insertional mutants

References

1. Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R. and Gadrinab, C., 2003. Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science*, *301*(5633), pp.653-657.
2. Alonso, J.M. and Ecker, J.R., 2006. Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. *Nature Reviews Genetics*, *7*(7), p.524.
3. Azpiroz-Leehan, R. and Feldmann, K.A., 1997. T-DNA insertion mutagenesis in Arabidopsis: going back and forth. *Trends in Genetics*, *13*(4), pp.152-156.
4. Feldmann, K.A. and Marks, M.D., 1987. Agrobacterium-mediated transformation of germinating seeds of Arabidopsis thaliana: a non-tissue culture approach. *Molecular and General Genetics MGG*, *208*(1-2), pp.1-9.
5. Forsthoefel, N.R., Wu, Y.W., Schulz, B., Bennett, M.J. and Feldmann, K.A., 1992. T-DNA insertion mutagenesis in Arabidopsis: prospects and perspectives. *Functional Plant Biology*, *19*(4), pp.353-366.
6. Henikoff, S. and Comai, L., 2003. Single-nucleotide mutations for plant functional genomics. *Annual Review of Plant Biology*, *54*(1), pp.375-401.

References

7. Henry, I.M., Nagalakshmi, U., Lieberman, M.C., Ngo, K.J., Krasileva, K.V., Vasquez-Gross, H., Akhunova, A., Akhunov, E., Dubcovsky, J., Tai, T.H. and Comai, L., 2014. Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing. *The Plant Cell*, *26*(4), pp.1382-1397.
8. Huang, M., Bai, Y., Sjostrom, S.L., Hallström, B.M., Liu, Z., Petranovic, D., Uhlén, M., Joensson, H.N., Andersson-Svahn, H. and Nielsen, J., 2015. Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proceedings of the National Academy of Sciences*, *112*(34), pp.E4689-E4696.
9. Jiao, Y., Burke, J.J., Chopra, R., Burow, G., Chen, J., Wang, B., Hayes, C., Emendack, Y., Ware, D. and Xin, Z., 2016. A sorghum mutant resource as an efficient platform for gene discovery in grasses. *The Plant Cell*, pp.tpc-00373.
10. Krasileva, K.V., Vasquez-Gross, H.A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., Simmonds, J., Ramirez-Gonzalez, R.H., Wang, X., Borrill, P. and Fosker, C., 2017. Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences*, *114*(6), pp.E913-E921.
11. Krysan, P.J., Young, J.C. and Sussman, M.R., 1999. T-DNA as an insertional mutagen in Arabidopsis. *The Plant Cell*, *11*(12), pp.2283-2290.

References

12. Krysan, P.J., Young, J.C., Tax, F. and Sussman, M.R., 1996. Identification of transferred DNA insertions within Arabidopsis genes involved in signal transduction and ion transport. *Proceedings of the National Academy of Sciences*, *93*(15), pp. 8145-8150.
13. Li, G., Jain, R., Chern, M., Pham, N.T., Martin, J.A., Wei, T., Schackwitz, W.S., Lipzen, A.M., Duong, P.Q., Jones, K.C. and Jiang, L., 2017. The sequences of 1504 mutants in the model rice variety Kitaake facilitate rapid functional genomic studies. *The Plant Cell*, *29*(6), pp.1218-1231.
14. McCallum, C.M., Comai, L., Greene, E.A. and Henikoff, S., 2000. Targeting induced locallesions in genomes (TILLING) for plant functional genomics. *Plant physiology*, *123*(2), pp.439-442.
15. Page, D.R. and Grossniklaus, U., 2002. The art and design of genetic screens: Arabidopsis thaliana. *Nature Reviews Genetics*, *3*(2), p.124.
16. Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., Kim, H., Cross, J., Morreel, K., Araujo, P. and Welsh, L., 2013. Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in Arabidopsis. *Science*, *341*(6150), pp.1103-1106.
17. van Opijnen, T., Bodi, K.L. and Camilli, A., 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature methods*, *6*(10), pp.767-772.