RESEARCH HIGHLIGHTS

GENETIC SCREENS

RNA-seq into the toolkit

Forward genetic screens in model organisms are powerful approaches for identifying mutations underlying various phenotypes and disease states. Methods to identify the causal mutation generally involve genetic mapping to a megabase-scale chromosomal region, followed by focal sequence analyses, but this can be laborious for organisms with complex genomes. Two new studies have devised pipelines based on high-throughput RNA sequencing (RNA-seq) for facilitating various aspects of forward genetic screens.

Working initially with known, recessive, single-gene mutations in zebrafish, Miller *et al.* and Hill *et al.* developed pipelines that begin with bulk segregant analysis (BSA) for the first stage of mapping. BSA involves intercrossing heterozygous carriers followed by polymorphism analysis of affected versus unaffected pooled progeny to identify the signature of

The functional relevance of both mutations was validated



homozygosity in the region surrounding the causal mutation. This analysis can be achieved by sequencing-based single-nucleotide polymorphism (SNP) genotyping, which has the advantage of also providing data for the subsequent fine-mapping stage. Both teams used RNA-seq as a more economical means of genome-scale coverage for BSA compared with previous uses of genome-wide DNA sequencing.

Although most SNPs occur in intergenic regions, both groups showed that the transcriptomes contained enough SNPs (mostly in untranslated regions) to locate the chromosomal region harbouring the causal mutation, although this required bioinformatic corrections to compensate for noise in RNA-seq data. The resolution of the mapping was dependent on the number of individuals in each pool and the sequencing depth, but 10 million sequence reads of 60 individuals per pool resulted in a ~6 Mb-sized locus, which is amenable to fine-mapping.

RNA-seg data provide opportunities and challenges for pinpointing candidate causal mutations in the identified region. Unlike DNA sequencing, RNA-seq only analyses expressed transcripts, hence mutations in non-coding regions will be missed. Also, some relevant transcripts might be temporally regulated, but both teams extracted RNA from zebrafish embryos as soon as the differential phenotypes became visible to maximize the chances of detecting them. Advantages of RNA-seq include the detection of transcript levels and rearrangements that might report the effects of causal mutations.

Both groups used bioinformatic analyses across the large regions from the initial mapping step to prioritize the SNPs that are most likely to be deleterious. Hill et al. showed that the known nonsense mutations in nkx2.5 and tbx1 could be uncovered. Similarly, in their strains, Miller et al. found the known nonsense mutations in hoxb1b, nhsl1b and egr2b, and inferred that the egr2b mutation results in nonsense-mediated RNA decay based on the effect on transcript levels. They also found the splice-site mutation in vangl2 by detecting the resultant aberrant transcripts.

Crucially, Hill *et al.* also used their pipeline to characterize two mutant strains for which the causal mutation was not known. They found nonsense mutations in *ctr9* and *cds2* by identifying downregulation from RNA-seq data and confirming the mutation by RNA-seq or DNA sequencing. The functional relevance of both mutations was validated through phenotypic rescue when wild-type alleles were expressed.

It will be interesting to see the extent to which these pipelines can identify novel mutations of different molecular types underlying diverse traits, and to assess fully the merits of sequencing RNA, DNA or both for screens in model organisms with different genome complexities.

Darren J. Burgess

ORIGINAL RESEARCH PAPERS Miller, A. C. et al. RNA-seq based mapping and candidate identification of mutations from forward genetic screens. *Genome Res.* 8 Jan 2013 (doi:10.1101/ gr.147322.112) | Hill, J. T. et al. MMAPPR: Mutation Mapping Analysis Pipeline for Pooled RNA-seq. *Genome Res.* 8 Jan 2013 (doi:10.1101/gr.146936.112)