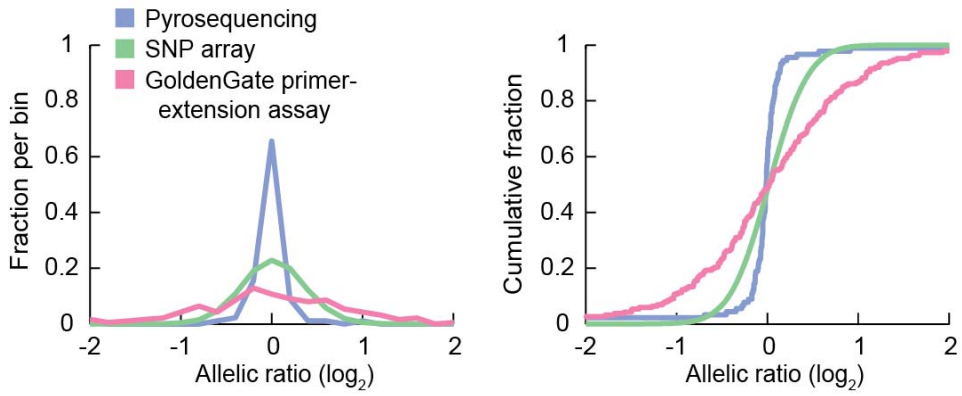
Supplementary Discussion and Figure

Another approach for detecting effects of regulatory SNPs is provided by studies of expression quantitative trait loci (eQTL)[1]. In eQTL studies, correlation between genotype of a polymorphic locus and expression of a gene is calculated for each locus:gene pair. In principle, eQTL studies involving unrelated individuals should preferentially identify polymorphic targets as cis-regulated because the SNPs in functional target sites and other linked SNPs should be associated with expression of the targets. However, when we analyzed the results of a large-scale eQTL study that used over 400 human liver samples[2], polymorphic miR-122 targets were not enriched among the genes identified as cis-regulated any more than were polymorphic miR-1 targets (data not shown). We attribute the greater sensitivity of AI-Seq to the internal reference provided by the non-target allele, which normalizes for environmental differences, trans-acting genetic differences and other sources of sample variability, thereby more effectively isolating the influence of the site on expression. Also important for the success of our approach in detecting the relatively subtle effect of miRNAs was the precision achieved by high-throughput sequencing. Previous studies using heterozygous SNPs to detect allelic expression imbalances rely on allele-specific hybridization or primer extension[3-6], which when compared at the gDNA level, were substantially noisier than our sequencing-based method (**Supplementary Fig. 1**).

Evaluation of previous allelic-imbalance measurement methods. For evaluating SNP arrays, we downloaded from the HapMap[7] website the raw signal-intensity data generated by hybridizing gDNA of a HapMap individual (NA19193) on the Affymetrix GeneChip 250K Nsp array. For evaluating the GoldenGate primer-extension assay, we downloaded from the Gene Expression Omnibus the raw signal-intensity data (GSM199494, GSM200074) generated from the Illumina GoldenGate assay with gDNA of a HapMap individual (NA10836)[5]. For both cases, SNPs annotated as heterozygous in the individual were identified from the HapMap genotype database, and the allele-specific

probe intensity values for the SNPs were used to calculate $\log_2$ ratios of one random allele to the other.

1. Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. Mapping complex disease traits with global gene expression. *Nat Rev Genet* **10**, 184-94 (2009).
2. Schadt, E.E. et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* **6**, e107 (2008).
3. Lo, H.S. et al. Allelic variation in gene expression is common in the human genome. *Genome Res* **13**, 1855-62 (2003).
4. Gimelbrant, A., Hutchinson, J.N., Thompson, B.R. & Chess, A. Widespread monoallelic expression on human autosomes. *Science* **318**, 1136-40 (2007).
5. Tan, A.C. et al. Allele-specific expression in the germline of patients with familial pancreatic cancer: an unbiased approach to cancer gene discovery. *Cancer Biol Ther* **7**, 135-44 (2008).
6. Serre, D. et al. Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic cis-acting mechanisms regulating gene expression. *PLoS Genet* **4**, e1000006 (2008).
7. Frazer, K.A. et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851-61 (2007).



Supplementary Figure 1 Noise of various allelic-imbalance measurement methods. Allelic-ratio measurements are compared for gDNA, which has an allelic ratio of 1:1, corresponding to 0.0 on a log2 scale. Standard (left, 0.2-unit bin) and cumulative (right) distribution of allelic ratios, measured using pyrosequencing (n = 90), the Affymetrix SNP array (n = 73,944), and the Illumina GoldenGate assay (n = 187).