

The **Affymetrix GeneChip® Brassica Exon 1.0 ST Array** was initiated in the UK as part of a BBSRC-funded project on *Brassica* biofortification (Broadley, Hammond, King, Rawlings et al.). The array is now in production (Jan 2010) and will be made available as a catalogue item. In addition, the NASC Affymetrix Service (<http://affymetrix.arabidopsis.info/>) will offer this array as a service item, with costs comparable to existing advertised services.

Summary information for the Affymetrix GeneChip® Brassica Exon 1.0 ST Array:

- 5µM 49-7875 format array
- 2.4 million 25 base oligonucleotide probes
- representing 135,201 gene models
- 15 probes per gene, distributed over exons

The design process was carried out by Chris Love (Rothamsted) and involved collating gene model and transcript data available in late 2009.

- The starting point was the pre-existing JCVI Unigene set (Trick et al., (2009) BMC Plant Biology 9:50 <http://ukpmc.ac.uk/articlerender.cgi?accid=pmcA2685394>) which has been used to develop the Agilent 95k oligo array.
- Due to the JCVI Unigene set being composed of assemblies of ESTs from different *Brassica* species, it is not easy to provide a detailed breakdown of unique genes by species. In addition, the dataset has gone through a number of filtering steps to avoid redundancy and orientate, where possible.
- Additional datasets included those from gene predictions derived from publicly available *B. rapa* Chiifu KBr BAC sequences (BrGSP), from 'digital transcriptome' of *B. napus* Tapidor and Ningyou7 (Trick et al., 2009. Plant Biotechnol J 7:334-46), additional *B. oleracea* and other ESTs, and finally a small subset of Arabidopsis gene models that were not otherwise represented by *Brassica* orthologues.
- Discrimination of the 135,201 gene models was based on an e-value cut-off of 1^{-5} , with $\leq 98\%$ sequence identity.
- This dataset was further filtered using the Affymetrix probe selection pipeline and standard Affymetrix Arabidopsis control and reporter sequences were added (89 Affymetrix controls and reporter genes). Probesets were selected based on 15 probes per gene.
- The following number of genes are represented with probes that passed the Affymetrix probe selection process:

Source	Number of gene models
JCVI unigene set not represented by gene predictions	89,216
PASA gene predictions	14,254
SNAP gene predictions	13,306
JIC <i>Brassica napus</i> N/T solexa (Velvet) assemblies	9,300
Arabidopsis gene models	4,517
AAFC/WHRI <i>B. oleracea</i> assemblies (cap3 94%id)	2,215
ESTs not represented in JCVI unigene set extracted from Genbank on 14/05/09 (assembled with cap3 94% id)	142
Arabidopsis controls	176
Total set:	135,201

Further information will become available via <http://www.brassica.info/resource/transcriptomics.php>