Non-normalized cDNA
(Multiple individuals, 11 tissues)

Illumina paired-end reads

Trimming of low quality and too short reads with length < 15 bp

Reference assembly
(Input: 493.4 million reads; Reference: 105,182 catfish EST consensus sequences)

104,870 contigs averaging 773 bp
(336.0 million reads assembled)

De novo assembly of remaining reads
(157.4 million reads)

Total of 526,099 contigs used for putative gene and SNP identification

Gene identity and annotation
(BLASTX to known protein database)

SNP identification and filtration from three sets of assembly

Gene Ontology analysis using Blast2GO

Inter-specific and intra-specific SNPs characterization