**Glossary of terms**

***Aiptasia***: a genus of globally distributed, tropical sea anemones

***Symbiodinium*:** a genus of dinoflagellate algae, many species of which form symbiotic relationships with invertebrate animals including corals and sea anemones.

***Primer****:* a short strand of nucleotides, usually about 20 bases or so, which serves as a starting point for DNA synthesis.

***Oligo:*** see definition for ‘primer’.

***Allele:*** One of two or more alternative forms of a gene.

***Locus:*** The specific location or position of a DNA sequence on a chromosome.

***SNP:*** single nucleotide polymorphism; a variation in a single nucleotide between two DNA sequences.

***Diploid:*** Having two complete sets of chromosomes.

***Genotype:*** The genetic makeup of an individual at a given location in the genome (noun) or the act of determining the genetic makeup of an individual (verb).

***Homozygous:*** For a diploid organism, having the same alleles at a given locus.

***Heterozygous:*** For a diploid organism, having different alleles at a given locus.

***GBS (Genotyping-by-Sequencing):*** a method for discovering large numbers of SNPs across many samples that involves reduction of genome complexity using restriction enzymes and high-throughput sequencing.

***2bRAD:*** A type of GBS that uses a type IIb restriction enzyme.

***reference-guided alignment:*** A type of analysis that involves use of a ‘reference’ genome. Sequence reads are mapped or aligned to the reference genome.

***de-novo assembly:*** A type of analysis in which a genome is created directly from sequence reads by assembling them together.

***BLAST:*** Basic Local Alignment Search Tool; finds regions of similarity between biological sequences.

***Illumina:*** A company that develops instruments for high-throughput genetic sequencing.

***Read:*** the data from a high-throughput sequencing experiment. Reads can vary in length and are usually provided in fastq format.

***Contig:*** A long stretch of sequence created from overlapping reads

***Scaffold:*** A longer stretch of sequence created from joined contigs.

***barcode:*** A short stretch of nucleotides (~6 bp) unique to a sample that allows for bioinformatic separation of reads from different samples.

***Fastq:*** A format for DNA sequences that includes quality information.

***Fasta:*** A format for DNA sequences that includes only the sequence and a descriptive header.

***sam***: common file format for read alignments.

***vcf:*** common file format for variant calls.