Supplementary information: Genome restrict digestion simulation

To screen the best restrict enzyme combination for genotyping by sequencing (GBS) library construction, a genome restrict digestion simulation was performed using the assembled scaffolds from the genome survey sequencing as a reference. Common used rare cutter (six bases restriction enzymes) including BamHI, EcoRI, HindIII, PstI, SalI, XbaI and XholI and frequent cutter (four bases restriction enzymes) including MseI and MspI were used for simulation and both the number of fragments and the length distribution were counted and compared. Among all restriction enzyme combinations in the two groups, i.e. rare-cutter-MseI and rare-cutter-MspI, EcoRI-MseI and EcoRI-MspI produced more fragments than other combinations in each group (Fig. S1, S2). Since the more smashed the genome been digested the more likely the fragments would be evenly distributed on the whole genome, EcoRI-MseI and EcoRI-MspI were selected as alternative restrict digestion combination for reduced representation library construction.

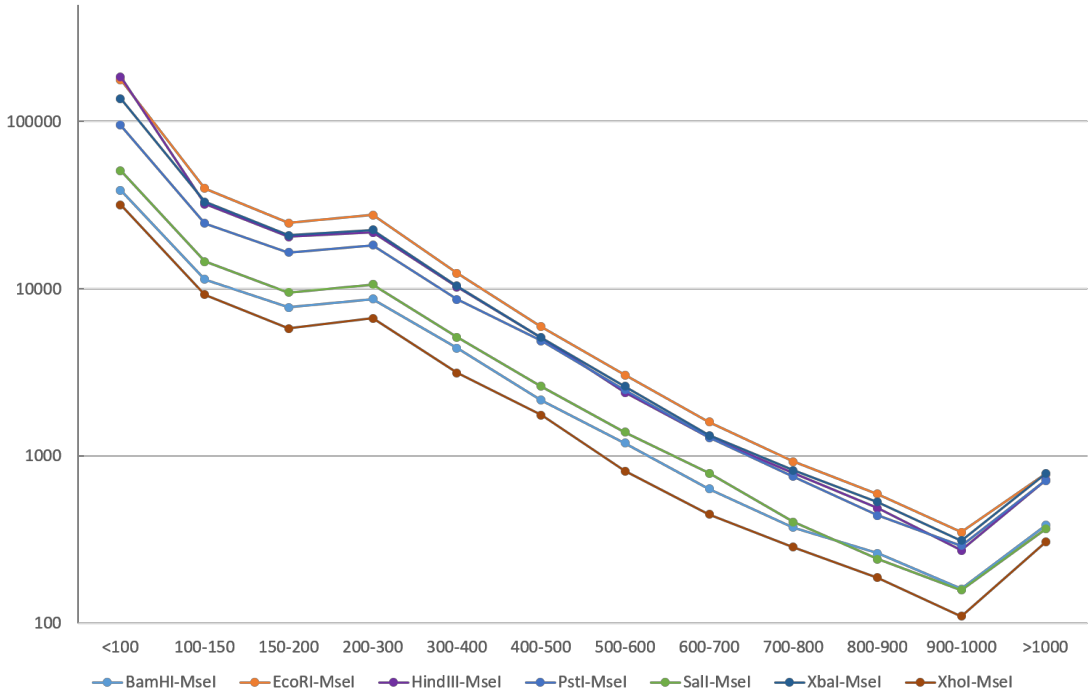

Fig. S1 Length distribution fragments digested by common restrict enzyme-MseI

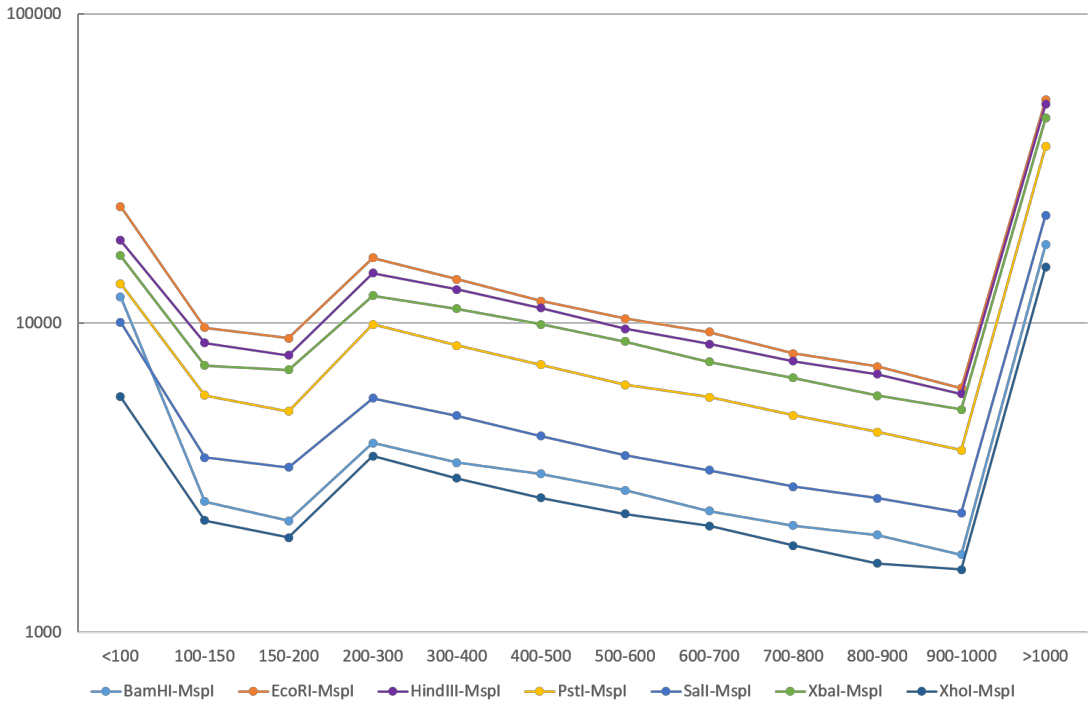

Fig. S2 Length distribution fragments digested by common restrict enzyme-MspI

The genome DNA of bay scallop was used for trial test for library construction so as to assess the capacity of the library. Library DNA were cloned into pMD-19 T vector and twenty clones randomly collected from each of the two libraries were Sanger sequenced. Multi-sequence alignment analysis revealed that two clones in the EcoRI-MseI library were the same, implying the possibility that the library contained large part of duplicate genome sequences which would reduce the effectiveness of sequencing. Thus, the EcoRI-MspI combination was used for GBS library construction.