Supplemental Table S10. Assembly of the "model genome" (human 10q11 segmental duplication) with *k*-mer validation and hybrid assembly methods. Note that only LH *k*-mer validation produces a correct assembly with two contigs.

| Method* | Both contigs | | | Left contig (401 kb) | | | Right contig (201 kb) | | |
|---|---|---|---|---|---|---|---|---|---|
| | % sequence lost | % duplicated | Number of contigs | % sequence lost | % duplicated | Number of contigs | % sequence lost | % duplicated | Number of contigs |
| M | 6.3 | 19.2 | 11 | 4.2 | 15.0 | 6 | 10.6 | 27.7 | 5 |
| L | 3.3 | 1.0 | 3 | 2.5 | 0 | 1 | 4.9 | 3.0 | 2 |
| LH | 3.7 | 0 | 2 | 2.4 | 0 | 1 | 6.3 | 0 | 1 |
| Hybrid | 4.5 | 17.5 | 12 | 3.5 | 8.4 | 5 | 6.5 | 35.7 | 7 |
| DCA | 1.9 | 7.2 | 5 | 1.5 | 7.8 | 3 | 2.7 | 6.0 | 2 |

* "M", standard MHAP; "L", MHAP with low-frequency *k*-mer masking; "LH", MHAP with low and high frequency *k*-mer masking; "Hybrid", standard hybrid assembly (Illumina reads used to correct the PacBio reads, followed by direct assembly); "DCA", doubly corrected assembly (Illumina reads used to correct the PacBio reads, followed by self correction of the long reads, then followed by the assembly).