Table S1. Sequencing metrics
The metrics of Pacbio SMRT sequencing reads and the paired-end Illumina sequencing reads.

Table S2. SNVs, indels and structural variations
The SNVs and indels between our assembly and the reference genome are shown in the SNV sheet. The repetitive regions with structural variants are masked for SNV and indel identification; the coordinates for the masked regions in our assembly and corresponding regions in the reference genome are found in SV_for_masking sheet. We identified two SNPs in the rDNA regions from the Illumina read alignment, and manually added them to the second-to-last repeat in our assembly: the SNV locations are shown in rDNA.polymorphism sheet. The structural variants detected by Assemblytics (Nattestad and Schatz 2016) are in SV.assemblytics sheet.

Table S3. Subtelomere Gene Annotation
The genes in re-annotated subtelomeres (and boundary non-subtelomeric genes) are shown in Annotation sheet. In addition to genes annotated using genes currently in the reference, we performed *de novo* ORF calling and annotated one more ORF, which is shown in denovo.ORF.calling sheet.

Table S4. Gene Comparison
The novel genes in our assembly are in Gene.New sheet. The genes in the reference that are removed in our assembly are in Gene.Remove sheet. The genes with structural variation (ORF genome length difference between our assembly and the reference > 50 nt) are in Gene.SV sheet. The genes with small variations (ORF genome length difference between our assembly and the reference less or equal to 50 nt) are in Gene.Variant sheet. There are four GPI-CWP genes with frameshifts in the ORF (see Table S4 for details). The length comparison of all single-exon genes in the reference genome that are also in our assembly is shown in Length.Comp sheet. The average Illumina coverage of those genes across the ORF region in our assembly and the reference is also shown in Length.Comp sheet. The average Illumina coverage of all the single-exon genes in our assembly is in Gene.Assembly.Illumina and that of all the single-exon genes in the reference genome is in Gene.Reference.Illumina.

Table S5. Putative GPI-anchored proteins
The putative GPI-anchored CWPs are in GPI.Adhesin sheet. The putative GPI-anchored proteins that are enzymes or structural genes are in GPI.enzyme_structural sheet. There are four GPI-CWPs with frameshifts in ORFs. Our corrections to predict the protein sequence are in Frame.Correction sheet. These corrections were only for the purpose of protein sequence comparison, and the genome sequence was not changed.

Table S6. GPI-CWP Adhesin-Like Gene Features
This table describes the following metrics: The number of beta-aggregation motifs (5 amino acids with > 10% aggregation potential predicted by Tango (Linding *et al.*, 2004; Fernandez-Escamilla *et al.*, 2004; Rousseau *et al.*, 2006)) in Tango > 10% column; the number of SFFIT, SHITT, TTITL motifs; the class of repeat region based on beta-aggregation potential; whether the repeat region contains a boundary beta-aggregation core.

Table S7. PCR primers for validation of SNVs and indels
The sequencing primers to the selected SNVs and indels in our assembly compared to the reference.

Table S8. PCR primers for Fosmid Analysis
The oligos to clone the subtelomeres into fosmids, and the oligos used for Sanger sequencing of the cloned subtelomeres.