**Supplementary methods**

**NGS data processing**

Sequencing data were analyzed through aligning the clean reads to the reference human genome (hg38) using BWA18 (version 0.7.12-r1039) [13]. Single-nucleotide variants as well as small insertions and deletions (indels) were identified by MuTect19 (version 1.1.4) [14]. A somatic mutation was confirmed if it possessed a variant allele fraction ≥ 1% and at least five high-quality reads (Phred score ≥ 30, mapping quality ≥ 30, and absent paired-end reads bias). Mutations were annotated to the genes using ANNOVAR20 software [15]. Copy number variations were detected by CONTRA21 [16].

**Statistics**

The associations between disease subtype and clinical characteristics were analyzed using the chi-squared test. Differences between groups were assessed by the Student’s *t*-test or one-way analysis of variance. Statistical significance was based on two-tailed tests at *p* < 0.05. SPSS 23.0 software (IBM Corp., Armonk, NY, USA) was used for statistical analyses.