Supporting Information

Exome sequencing

DNA was extracted from cells and cSCC samples following the standard procedures. The exome sequences were enriched from 1.0 μg genomic DNA by Agilent liquid capture system (Agilent SureSelect Human All Exon V5) according to the protocol of manufacturer. Qualified genomic DNA had been randomly fragmented to the average size of 180-280 bp using Covaris S220 sonicator. The gDNA fragments had been end repaired and phosphorylated, followed by the A-tailing and ligation at the 3'ends with the paired-end adaptors (Illumina) with a single "T" base overhang and purification by AMPure from Agencourt. Next, the size distribution and concentration of libraries had been determined by Agilent 2100 Bioanalyzer and qualified by real-time PCR, respectively. DNA libraries were sequenced on Illumina Hiseq 4000 system for paired-end 150 bp reads.

SAMtools mpileup and BCFtools were used to do variant calling and identify Single Nucleotide Polymorphism (SNP) and Insertion-Deletion (INDEL) mutation. CoNIFER[1] was used to discover disruptive genic copy number variants (CNVs) in human genetic study of diseases. ANNOtate VARiation (ANNOVAR)[2] was performed to do annotation for Variant Call Format (VCF). Variants obtained were filtered by the minor allele frequency (MAF)>1% in the 1000 Genomes databases (1000 Genomes Project Consortium). Single nucleotide variants (SNVs) occurring in exons and canonical splice sites (splicing junction 10 bp) were further analyzed by Polymorphism Phenotyping v2 (PolyPhen-2)[3], Sorting Intolerant From Tolerant

(SIFT)[4], MutationTaster[5] and Combined Annotation Dependent Depletion (CADD)[6] for functional prediction.

References

1. Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP. *et al.* Copy number variation detection and genotyping from exome sequence data. *Genome Res.* 2012; 22: 1525-32

2. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010; 38: e164

3. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* 2013; Chapter 7: Unit7.20

4. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003; 31: 3812-4

5. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods.* 2010; 7: 575-6

6. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.* 2014; 46: 310-5