Data S1

Single nucleotide variant (SNV)/insertion/deletion (Indel) analysis. SNVs and Indels were detected using DeepVariant v1.4.0 (1). The variant filtration and interpretation were performed using the Franklin platform v72 (franklin.genoox.com), developed by Genoox, which is based on the guidelines of the American College of Medical Genetics and Genomics (2). First, variants were filtered by inheritance pattern. Three modes were considered: *De novo* autosomal dominant, autosomal recessive and compound heterozygous. To identify candidate variants, two approaches were used. The first approach filtered for only variants that were pathogenic or likely pathogenic based on Franklin classification. The second approach filtered for variants whose effects were either missense, nonsense, frameshift, non-frameshift or splice site variants or variants with aggregated prediction >0.15 (calculated by Franklin) with a minor allele frequency <1% (from the gnomAD database, gnomad.broadinstitute.org). Candidate variants from either approach were confirmed for their presence visually using Integrative Genomics Viewer (IGV) v2.15.4 (3). Next, it was checked if the particular variants were rare (allele frequency <1%) in the in-house Thai control database at the Excellence Center for Genomics and Precision Medicine, King Chulalongkorn Memorial Hospital, Bangkok, Thailand.

Structural variant (SV) analysis. SVs (Indels, inversions and duplications) were detected using pbsv v2.8.0 (github.com/PacificBiosciences/pbsv). Signatures of structural variation were identified by *pbsv discover* . Structural variations were then jointly called and genotyped by pbsv call using structural variation signatures obtained from *pbsv discover* with the default settings. Variant calls for four samples (both twins and their parents) were output in a single VCF file. Only SVs with an alternate allele depth ≥10 were considered for further analysis. First, variants were filtered by the three inheritance pattern then annotated by ANNOVAR v2023Mar15 (4). Candidate variants were filtered by pathogenicity or effect. Variants were retained if their classification was pathogenic or likely pathogenic or their effect was either non-frameshift, frameshift, stop-loss, stop-gain, start-loss, start-gain or splice site variants with a minor allele frequency <1% (from the gnomAD database). Finally, as in the SNVs/Indels analysis, variants were screened for obvious false positives and checked against the in-house database.

Copy number variation (CNV) analysis. CNVs were detected using HiFiCNV v0.1.7 (github.com/PacificBiosciences/HiFiCNV). Alignments from BAM files were used to construct bins of average depth across each chromosome (bin size, 2 kb), assign genotypes and generate VCF files. The called variants were compared between twins and parents to identify variations that have the same value in both twins but different from both parents. All variants were then annotated and pathogenicity score calculated by ClassifyCNV v1.1.1 (5). Variants with IGV were checked for rarity using the in-house unaffected control subject database.

Methylation profile analysis. 5-Methylcytosine information was added to the BAM files in MM and ML tag format using jasmine v5.1.2 (github.com/pacificbiosciences/jasmine). Small variant calls from DeepVariant v1.4.0 (github.com/google/deepvariant) were haplotyped and phased using WhatsHap v1.7.0 (6). Haplotype assigned reads (Haplotagged.bam) were used as an input for pb-CpG-tools v1.1.0 (github.com/PacificBiosciences/pb-CpG-tools) to estimate the modification probability. Differentially methylated regions (DMRs) between twins and parents were detected using metilene v0.2-8 (7). A difference >10% and q-value <0.05 was considered significant. DMRs with <10 CpGs and regions on sex chromosomes were discarded. All DMRs were annotated by ChIPseeker v1.28.3 (8) using gencode.v44 (gencodegenes.org/human/release_44.html) for annotation. Binomial genomic region enrichment analysis was performed using GREAT (v.4.0.4) (9), which quantified the significance of enrichment in terms of binomial fold enrichment and P-value. Gene Ontology (geneontology.org/) biological process terms were used for prediction purposes. The enriched annotations with a false discovery rate <0.05 were considered significant.

References

1. Poplin R, Chang PC, Alexander D, Schwartz S, Colthurst T, Ku A, Newburger D, Dijamco J, Nguyen N, Afshar PT, *et al*: A universal SNP and small-indel variant caller using deep neural networks. Nat Biotechnol 36: 983-987, 2018.
2. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, *et al*: Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. Genet Med 17: 405-424, 2015.
3. Thorvaldsdóttir H, Robinson JT and Mesirov JP: Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. Brief Bioinform 14: 178-192, 2012.
4. Wang K, Li M and Hakonarson H: ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38: e164, 2010.
5. Gurbich TA and Ilinsky VV: ClassifyCNV: A tool for clinical annotation of copy-number variants. Sci Rep 10: 20375, 2020.
6. Martin M, Patterson M, Garg S, Fischer SO, Pisanti N, Klau GW, Schöenhuth A and Marschall T: WhatsHap: Fast and accurate read-based phasing. bioRxiv: Nov 14, 2016 (Epub ahead of print).
7. Jühling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF and Hoffmann S: Metilene: Fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. Genome Res 26: 256-262, 2016.
8. Yu G, Wang LG and He QY: ChIPseeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics 31: 2382-2383, 2015.
9. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM and Bejerano G: GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 28: 495-501, 2010.

Figure S1. Mean expression of *TYMS* mRNA in PBMCs. Reverse transcription-quantitative PCR was performed on PBMCs from control (n=3) and patient samples (n=2). PBMC, peripheral blood mononuclear cell; TYMS, thymidylate synthase.

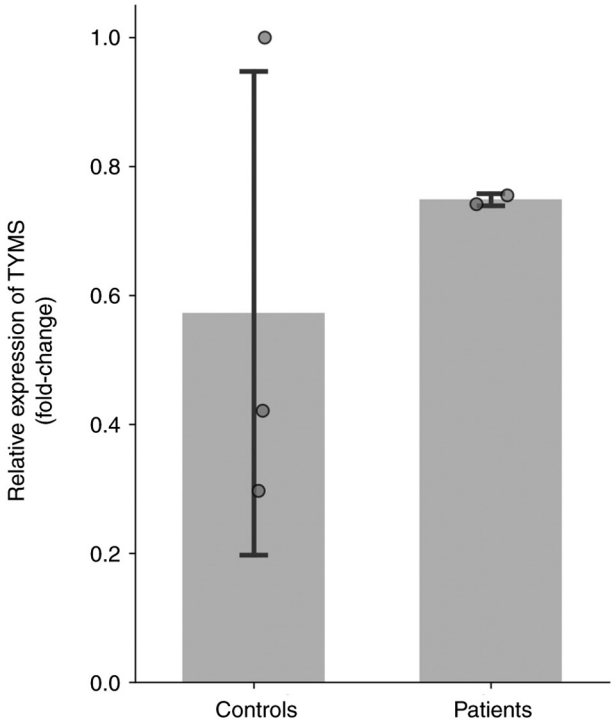

Figure S2. Original blots of thymidylate synthase detection.

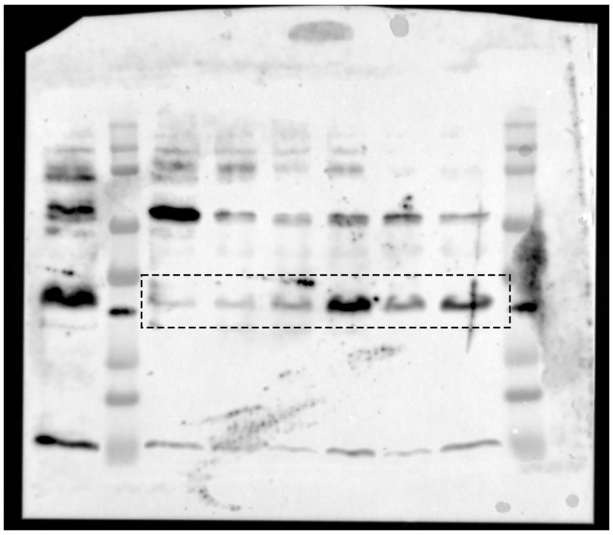

Figure S3. Original blots of β-actin detection.