Supplementary Materials: Germline (epi)genetics reveals high predisposition in females: a 5-year, nation-wide, prospective Wilms tumor cohort

Supplementary methods

Tumor classification

Histology and stage were assessed and centrally reviewed for all tumors according to the SIOP-staging after preoperative chemotherapy[1], and dictated the intensity of post-operative chemotherapy and sometimes radiotherapy for all patients.

Patient inclusion & Germline DNA analyses

Following diagnosis the patients consented to take part in the *Sequencing of Tumor and Germline DNA - Implications and National Guidelines* (STAGING) project. The current study focuses on patients with WT included in the STAGING study from July 1st 2016 until July 1st 2021. Inclusion procedures and germline sequencing protocols have been published elsewhere[2].

Briefly, leukocyte DNA was isolated from peripheral blood samples drawn alongside standard blood-sampling executed as part of treatment. When possible, parental blood samples were taken to establish whether detected pathogenic variants were inherited or occurred *de novo*.

Whole genome sequencing (WGS) was performed using the HiSeqX or NovaSeq platforms

Stoltze UK, *et al. J Med Genet* 2023;0:1–8. doi: 10.1136/jmg-2022-108982

(Illumina, San Diego, CA, USA) with paired-end sequencing of 150-bp reads and target 30X average coverage. Reads were mapped to the hg19 reference genome sequence (GRCh37.p13; RefSeq assembly accession GCF_000001405.25) using GATK version 3.8 or the DNAseq pipeline (Sentieon, San Jose, CA, USA). VarSeq software (version 2.2.3, Golden Helix, Bozeman, MT, USA) was used to annotate variants.

Rare variants (gnomAD frequency less than 0.1%) in a panel of 390 cancer related genes selected from the existing medical literature[3,4] were reviewed by a multidisciplinary team of clinical geneticists, pediatric oncologists and bioinformaticians and classified in accordance with current international standards[5]. Variants classified as "pathogenic" or "likely pathogenic" may be referred to collectively as "pathogenic" in this study.

In the whole genome detection of predicted loss-of-function (pLoF) variants, structural variants (SVs) were called for the full STAGING cohort based on aligned WGS data using Manta (1.4), CNVnator (0.3.3), CNV kit (0.9.6), Delly2 (0.8.1) and ExpansionHunter (2.5.6). Any SVs also detected in an in-house non-cancer cohort were removed, as were all non-exonic and/or non-deletion SVs. Similarly, using R (3.6.1), called single nucleotide variants (SNVs) were filtered by removing intronic and non-LoF SNVs and by application of the following quality control (QC) parameters; coverage >15X, VAF >0.3 and <0.70, strand bias <10, allele count =2, indel size <10. SNVs with >2 exact matches among non-WT cancer patients were removed.

The SV/SNVs remaining post-filtration were considered putative pLoF variants and subject for constraint gene analysis, which we have presented for childhood cancer predisposition

investigation previously[6]. Briefly, pLoF variants were filtered to those present in constrained genes only. Gene constraint was defined as any gene having a pLoF observed vs. expected upper bound fraction (LOEUF) score lower than 0.35. LOEUF scores were derived from canonical transcripts in Supplementary Dataset 11 in Karczewski et al[7]. The resulting variants underwent manual curation based on visual analysis of WGS data using Integrated Genome Viewer, comparison to The Genome Aggregation Database (gnomAD v2.1)[7] for population frequencies and ClinVar[8] for variant classification as well as scientific literature review.

Epigenetic germline and tumor analyses

Peripheral blood DNA was isolated from individuals with WT and age and sex-matched controls. An individual with molecularly confirmed BWS and IC1 hypermethylation was included as a positive control. When available, tumor DNA was isolated from patients. Bisulfite conversion was performed on 200 ng of DNA using an EZ-DNA Methylation-Gold kit (Zymo Research, Irvin, CA, USA), according to the manufacturer's instructions, and eluted in a volume of 10 μL. Approximately 20 ng of bisulfite converted DNA was amplified in triplicates using a Pyromark PCR kit (Qiagen, Hilden, Germany) and primers targeting a 208 bp region of IC1. Methylation quantification at five CpG sites at IC1 was carried out with 10 μL of PCR product on a Pyromark Q48 autoprep, and analyzed with Pyromark Q48 software. The PCR and pyrosequencing primers and parameters are detailed by Pignata et al.[9]. The mean of the triplicates was calculated, and the average methylation value of the five analyzed CpG sites was used. An individual was considered to have GOM at IC1 if the methylation level was higher

than the normal range, which was defined as the average methylation level of the controls plus/minus three standard deviations. A normal IC1 methylation range for kidney tissue established by Pignata et al. (34-66%)[9] was used to detect GOM in tumor tissue.

Methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA, ME030-C3, MRC Holland, Amsterdam, The Netherlands) was conducted according to the manufacturer's instructions. The mean of the four probes at IC1 was calculated and used in the analysis. The threshold for GOM at IC1 is set as > 0.65 in the clinical set-up.

A linear mixed model with technical replicate number as a random effect was used to assess the association between blood IC1 methylation levels by pyrosequencing and Wilms tumor or macrosomia. Other statistical tests used are specified in the text.

Statistical analyses

Statistical analyses were conducted using IBM SPSS Statistics (v.25) and R (v.3.6.1). The statistical tests used are specified.

Supplementary Table 1: Classification of both single nucleotide variants (SNV) and structural variants (SV) found in 390 genes across all patients

	SNV	SV
Characteristic	**N = 498**[1]	**N = 198**[1]
Variant classification		
benign	7 (1.4%)	0 (0%)
likely benign	80 (16%)	51 (26%)
variant of unknown significance	116 (23%)	36 (18%)
pathogenic (risk factor)	4 (0.8%)	0 (0%)
likely pathogenic	4 (0.8%)	1 (0.5%)
pathogenic	0 (0%)	0 (0%)
dismissed (poor quality)	287 (58%)	110 (56%)

[1] n (%)

Supplementary Table 2: Causative variants (both genetic and epigenetic) and predicted loss-of-function variants (pLoF) in constrained genes.

Causative genetic and epigenetic variants (n=4)											
Pt #	**Age at Dx** [in months]	**WT stage**	**Gene**	**Ontology**	**HGVS c.** [region affected for SV]	**HGVS p.** [size for SV]	**VAF** [alt/X]	**Inheritance**	**Family history**	**Extended phenotype**	**gnomAD alleles [freq. (count/number)]**
1	36-72	III	*REST*	Deletion	chr4:57,761,129-59,377,004	1.62Mb (spanning 7 genes[a])	49% [22/45]*	Paternal	Two WT cases[b]	Unremarkable	0.00 (0/21,478)
2	>72	IV	*FBXW7*	Nonsense	NM_033632.3:c.832C>T	NP_361014.1:p.Arg278Ter	58% [16/38]	Maternal	Unremarkable	Facial scoliosis etc.[c]	0.00 (0/249,772)
3	<36	I	*WT1*	Frameshift	NM_024426.4:c.332del	NP_077744.3:p.Pro111Argfs*47	33% [20/40]	*De novo*	Unremarkable	Fetal hydronephrosis[d]	0.00 (0/114,890)
4	36-72	I	UPD11	Chromosoma l	chr11:204,228-47,983,477	47.78Mb (spanning p15.5-11.2)	20-25%	*De novo*	Unremarkable	Classic BWS[e]	N/A

Additional pLoF variants in constrained genes (n=9)											
Pt #	**Age at Dx** [in months]	**WT stage**	**Gene**	**Ontology**	**HGVS c.** [region affected for SV]	**HGVS p.** [size for SV]	**VAF** [alt/X]	**Inheritance**	**Family history**	**Extended phenotype**	**gnomAD alleles [freq. (count/number)]**
1	36-72	III	*POLR2B*	Deletion	chr4:57,761,129-59,377,004	1.62Mb (spanning 7 genes[a])	49% [22/45]*	Paternal	Two WT cases[b]	Unremarkable	0.00 (0/21,478)
2	>72	IV	*ZCCHC8*	Frameshift	NM_017612.4:c.1074_1077del	NP_060082.2:p.Tyr359Ilefs*53	52% [32/61]	N/A	Unremarkable	Facial scoliosis etc.[c]	0.00 (0/270,388)
4	36-72	I	*SLIT2*	Nonsense	NM_004787.3:c.1849C>T	NP_004778.1:p.Arg617Ter	37% [10/27]	N/A	Unremarkable	Classic BWS[e]	0.00 (0/249,578)
10	36-72	I	*KCNA4*	Nonsense	NM_002233.3:c.1348C>T	NP_002224.1:p.Arg450Ter	38% [16/42]	N/A	Unremarkable	Unremarkable	0.00 (0/249,644)
12	36-72	III	*FRMD4A*	Frameshift	NM_001318337.1:c.678dup	NP_001305266.1:p.Leu227Thrfs*51	45% [15/33]	N/A	Unremarkable	Small stature	0.00 (0/282,512)
			SMC2	Frameshift	NM_006444.2:c.398_402del	NP_006435.2:p.Ser133Trpfs*3	49% [17/35]	N/A			0.00 (0/234,676)
14	<36	II	*OTUD4*	Nonsense	NM_001102653.1:c.2635C>T	NP_001096123.1:p.Arg879Ter	68% [28/41]	N/A	Unremarkable	Unremarkable	0.00 (0/282,758)
20	>72	III	*FRMD4A*	Nonsense	NM_001318337.1:c.142C>T	NP_001305266.1:p.Gln48Ter	49% [33/68]]	N/A	Agenesis of the kidney (father)	Wide fontanelle, sclera with blue hue	4.69e-5 (7/149,308)
			CTNND1	Nonsense	NM_001085458.1:c.2540C>A	NP_001078927.1:p.Ser847Ter	53% [25/47]	N/A			0.00 (0/242,410)

Pt #; patient number (females in bold), Dx; diagnosis, WT; Wilms Tumor, HGVS; Human Genome Variation Society, c.; coding DNA, SV; structural variant, p.;protein, VAF; variant allele frequency, X; coverage, UPD; uniparental disomy, BWS; Beckwith-Wiedemann Syndrome, pLoF; predicted loss-of-function.

[a]*IGFBP7, IGFBP7-AS1, LOC101928851, NOA1, POLR2B, REST, UBE2CP3*

[b]Paternal uncle [2nd-degree relative, WT (archived pathology report) at 4 years old, deceased at 6 years old, not tested] & Paternal grandmother's sister's son [4th-degree relative, WT (archived pathology report) at 4 years old, alive and well, carries an identical 1.62Mb [chr4:57,761,129-59,377,004 heterozygous deletion]. The proband's father, and the obligate carriers (paternal grandmother and her sister) were unaffected.

[c]Epichantus, facial scoliosis, septal heart defect and two congenital accessory skin tags on the cheek and behind the ear. Several teeth were later surgically removed, although this may be related to chemo treatments. All were noted on this study's phenotype checklist.

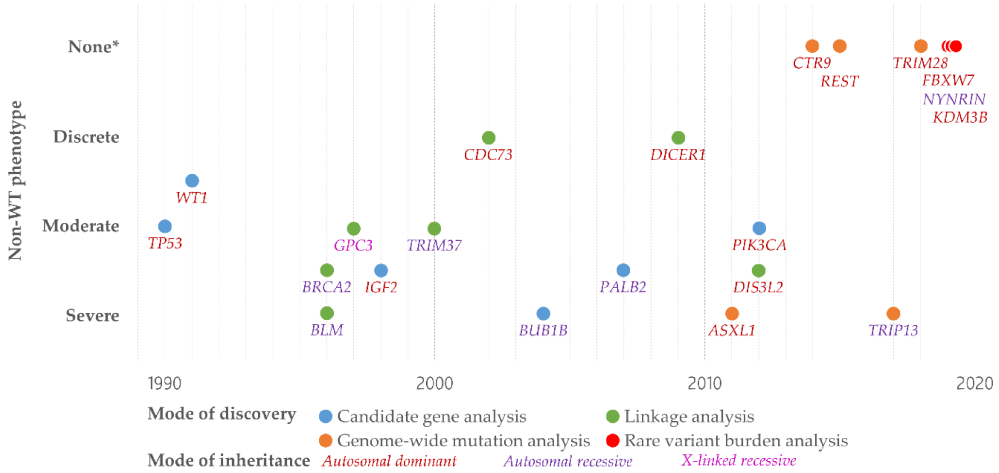
[d]Affected the left kidney where the patient later developed WT.

[e]Macrosomia, lateralized overgrowth of the left leg [+3 cm circumference, +1.5cm length], macroglossia & epicanthus

*VAF estimated across the span of the deletion and the bordering regions.

Supplementary Table 3: Variants of unknown significance (VUSs) of interest.

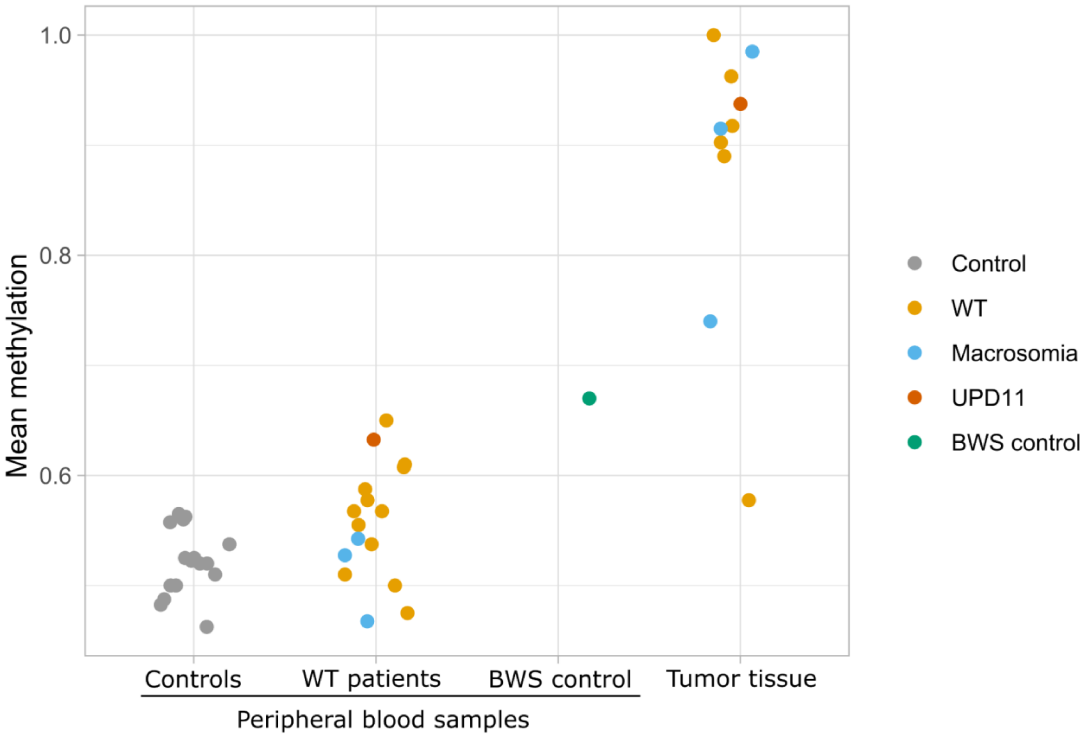
Pt #	Gene Names	Ontology	Protein change [nucleotide change]	VAF[alternate/total]	CADD score	gnomAD alleles [freq. (count/number)]	Clinvar
5	*NYNRIN*	Missense	p.Ala175Val [c.524C>T]	0.46[42/92]	23.3	0.00 (0/225,860)	Not reported
5	*CTR9*	Missense	p.Tyr385Cys [c.1154A>G]	0.40[14/35]	25.7	1.62e-5 (4/247,004)	Not reported
5	*ASXL1*	Missense	p.Ala1312Val [c.3935C>T]	0.55[26/47]	11.1	8.17e-4 (231/282,778)	LB
6	*DICER1*	Missense	p.Thr60Ile [c.179C>T]	0.57[17/30]	15.5	4.95e-5 (14/282,740)	VUS
6	*NSD1*	Initiator	p.Asp23Gly [c.68A>G]	0.56[20/36]	23	3.98e-6 (1/251,442)	VUS
9	*NYNRIN*	Missense	p.Gly353Arg [c.1057G>A]	0.46[13/28]	15.1	6.07e-5 (17/280,254)	Not reported
9	*BARD1*	Deletion	chr2:215,591,264-215,774,591	~0.51[24/49]	N/A	0.00 (0/21,694)	Not reported
12	*NBN*	Frameshift	p.Gln279Thrfs*6 [c.834dupA]	0.42[13/31]	37	0.00 (0/251,318)	Not reported
15	*NYNRIN*	Missense	p.Thr1172Met [c.3515C>T]	0.26[7/27]	20.4	8.03e-6 (2/248,982)	Not reported
15	*NYNRIN*	Missense	p.Glu420Met [c.1258_1259delinsAT]	0.41[14/34]	24.3	0.00 (0/277,536)	Not reported
19	*CTNNB1*	Missense	p.Glu155Asp [c.465A>T]	0.51[20/39]	19.2	0.00 (0/251,200)	Not reported
22	*PALB2*	Missense	p.Glu211Gly [c.632A>G]	0.58[23/40]	0	0.00 (0/251,406)	VUS
24	*REST*	Missense	p.Pro141Arg [c.422C>G]	0.43[19/44]	23.3	1.98e-4 (56/282,762)	VUS



Supplementary Figure 1: Illustrates the mode of discovery over time, trending toward NGS-based discoveries uncovering monogenic diseases with a narrow phenotype when compared to earlier “overt” syndromes linked to WT predisposition. *no phenotype other than WT risk reported.

Elaboration on Supplementary Figure 1: While Beckwith-Wiedemann syndrome was described as early as 1963, the identification of specific genes involved in the CPSs underlying WT predisposition was only possible beginning in the 1990s. Though not yet understood to include WT in the phenotypic spectrum at the time, *TP53* was the first of the 21 genes related to WT predisposition discovered in 1990. The year after *WT1* was discovered as the cause of WT1 disorder. Both genes were identified through candidate gene analysis. This approach, along with linkage analysis, ushered in an era of discovery that uncovered the molecular basis for many of the "overt" syndromes which had been recognized clinically for decades. Broadly, these syndromes tend to have moderate to severe non-WT phenotype.

The advent of next generation sequencing, genome-wide mutation analysis and, more recently, rare variant burden analysis, has led to discoveries of new "covert" syndromes, where the phenotype appears to be restricted to increased WT risk. This shift is illustrated in Supplementary Figure 1.



Supplementary Figure 2: Jitter plot showing IC1 methylation levels using MS-MLPA of all individuals included in the methylation analysis. The methylation levels are calculated as an average of the four sites analyzed. WT, Wilms tumor, BWS, Beckwith-Wiedemann syndrome, UPD11, uniparental disomy of chromosome 11.

Supplementary references

1. Pritchard-Jones, K. *et al.* Omission of doxorubicin from the treatment of stage II-III, intermediate-risk Wilms' tumour (SIOP WT 2001): an open-label, non-inferiority, randomised controlled trial. *Lancet Lond. Engl.* **386**, 1156–1164 (2015).
2. Byrjalsen, A. *et al.* Nationwide germline whole genome sequencing of 198 consecutive pediatric cancer patients reveals a high incidence of cancer prone syndromes. *PLoS Genet.* **16**, (2020).
3. Zhang, J. *et al.* Germline mutations in predisposition genes in pediatric cancer. *N. Engl. J. Med.* **373**, 2336–2346 (2015).
4. Rahman, N. Realizing the promise of cancer predisposition genes. *Nature* (2014) doi:10.1038/nature12981.
5. Richards, S. *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.* (2015) doi:10.1038/gim.2015.30.
6. Stoltze, U. K. *et al.* Population-based whole-genome sequencing with constrained gene analysis identifies predisposing germline variants in children with central nervous system tumors. 2022.03.01.22271578 Preprint at https://www.medrxiv.org/content/10.1101/2022.03.01.22271578v1 (2022).
7. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
8. Landrum, M. J. *et al.* ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **46**, D1062–D1067 (2018).
9. Pignata, L. *et al.* Both Epimutations and Chromosome Aberrations Affect Multiple Imprinted Loci in Aggressive Wilms Tumors. *Cancers* **12**, 3411 (2020).