Child Neurology Open
Volume 7: 1-10

DOI: 10.1177/2329048X20939003
journals.sagepub.com/home/cno
SAGE

Expanded Phenotypic Definition Identifies Hundreds of Potential Causative Genes for Leukodystrophies and Leukoencephalopathies

Veronica M. Urbik, MS[1], Marilyn Schmiedel, BS[2], Haille Soderholm, BA[3], and Joshua L. Bonkowsky, MD, PhD[3,4,5]

Abstract

Background: The genes responsible for genetic white matter disorders (GWMD; leukodystrophies and leukoencephalopathies) are incompletely known. Our goal was to revise the list of genes considered to cause GWMD. We considered a GWMD to consist of any genetic disease causing T2 signal white matter changes in magnetic resonance images. **Methods and Results:** Using a systematic review of PubMed, Google, published literature reviews, and commercial gene panels, we identified 399 unique genes meeting the GWMD definition. Of this, 87 (22%) genes were hypomyelinating. Only 3 genes had contrast enhancement on magnetic resonance imaging (MRI): *ABCD1*, *GFAP*, and *UNC13D*. **Conclusions:** A significantly greater number of genes than previously recognized, 399, are associated with white matter signal changes on T2 MRI. This expansion of GWMD genes can be useful in analysis and interpretation of next-generation sequencing results for GWMD diagnosis, and for understanding shared pathophysiological mechanisms of GWMDs.

Keywords

leukodystrophy, genes, leukoencephalopathy, classification, diagnosis

Received April 05, 2020. Received revised May 22, 2020. Accepted for publication June 08, 2020.

Leukodystrophies are genetic disorders that affect development or maintenance of the white matter of the central nervous system (CNS).[1-3] Leukodystrophies have an incidence of almost 1 in 7500 live births, with significant morbidities and death in a third by age 8.[4] A confounding feature to understanding leukodystrophies is their apparent genetic and mechanistic heterogeneity.[5] Further, even with advanced next-generation sequencing (NGS) approaches, diagnosis rates remain below 70%,[6] suggesting that a quarter of disease-causing genes may not even be known.

A variety of approaches to define and categorize leukodystrophies have been pursued. An international committee of experts classified 30 disorders as leukodystrophies.[7] They defined leukodystrophies as genetic, with T2 signal abnormality on magnetic resonance imaging (MRI), and including glial or myelin sheath abnormalities in the CNS. Further, they termed "genetic leukoencephalopathies" to describe disorders that are heritable and result in white matter abnormalities but that did not necessarily meet their strict criteria as a leukodystrophy. Also, more recent classification schemes have been proposed for leukodystrophies, for example, recognizing the complex pathology of different cell types[5] or emphasizing the sorting of leukodystrophies into different types based on

[1] University of Utah School of Medicine, Salt Lake City, UT, USA
[2] NXP Semiconductors Austin Oak Hill, Austin, TX, USA
[3] Division of Pediatric Neurology, Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT, USA
[4] Brain and Spine Center, Primary Children's Hospital, Salt Lake City, UT, USA
[5] Primary Children's Center for Personalized Medicine, Salt Lake City, UT, USA

Corresponding Author:
Josh Bonkowsky, MD, PhD, Department of Pediatrics, University of Utah School of Medicine, 295 Chipeta Way/Williams Building, Salt Lake City, UT 84108, USA.
Email: joshua.bonkowsky@hsc.utah.edu

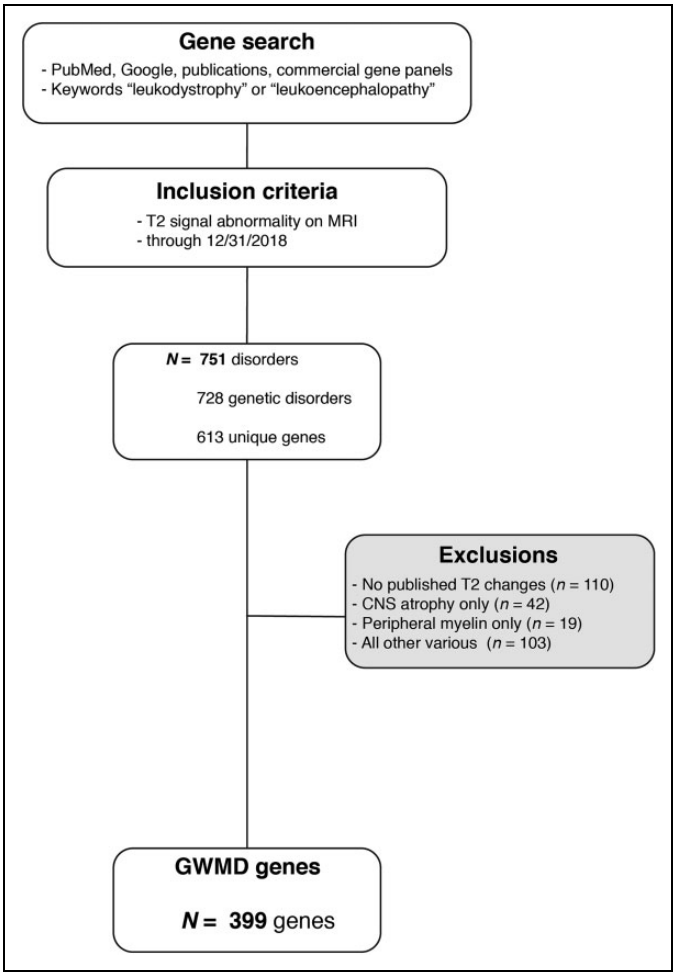

Figure 1. Schematic diagram of gene identification.

disease pathology such as hypomyelination or vasculature involvement.[8]

Our objective was to identify and include all genes that have been reported to cause T2 white matter abnormalities. Our hypothesis was that a more complete list of genes associated with leukodystrophies and leukoencephalopathies, which we will term "genetic white matter disorders (GWMD)," would be of utility for improving diagnostic yield in genetic testing, and would reveal unexpected shared mechanistic pathways. We chose not to exclude any apparent genetic cause, even if not historically considered as a leukodystrophy or leukoencephalopathy. A secondary aim was to determine whether there were any common genetic or mechanistic pathways identified by grouping similar disorders.

Methods

We conducted a systematic search using keywords "leukodystrophy" or "leukoencephalopathy," including of PubMed, Google, published literature reviews, and commercial gene panels (Figure 1). We included for consideration any publication reporting white matter signal changes on MRI in human patients. The timeline for publication was January 1, 1990, through December 31, 2018. Inclusion required a published report of abnormal T2 white matter signal abnormality on brain MRI. Exclusion criteria included any white matter change secondary to nongenetic cause, including traumatic, infectious, or autoimmune etiologies. We excluded any genomic-level structural chromosomal changes (deletion, duplication); we also excluded gray matter pathology without white matter involvement, brain iron disorders, and isolated atrophy, thinning, reduced volume, or absence of structures (eg, absence of the corpus callosum). Following review and manual curation, genes were characterized and grouped. We categorized genes as being hypomyelinating if they were specifically stated as such in published literature. We used the same criteria to identify genes reported to cause contrast enhancement. Each gene was linked with its Ensembl stable gene ID from Ensembl 92.[9]

Seven hundred fifty-one disorders of white matter were identified, including from publications[7,10-17]; lists from gene panel testing from GeneDx (https://www.genedx.com/test-catalog/available-tests/leukodystrophy-xpanded-panel/), the United Kingdom National Health Service (https://ukgtn.nhs.uk/find-a-test/search-by-disorder-gene/leukodystrophy-hypomyelinating-and-mitochondrial-leukoencephalopathy-96-gene-panel-899/), the scientific crowdsourcing resource Genomics England PanelApp version 1.60 (https://panelapp.genomicsengland.co.uk/panels/42/), Invitae (https://www.invitae.com/en/physician/tests/06155/), and the University of Chicago (https://dnatesting.uchicago.edu/sites/default/files/media/documents/Rett Angelman Information Sheet 4-27-17.pdf). The 751 disorders were limited to 728 genetic diseases, and then to 613 unique genes. Each gene was then reviewed in Online Mendelian Inheritance of Man, and if necessary searches were performed in PubMed to determine whether published examples of T2 MRI white matter changes were reported. Disorders involving nongenetic causes (eg, HIV, cytomegalovirus, dietary B12 deficiency) and portions of chromosomes (eg, 18q Deletion Syndrome, etc) were excluded. Disorders affecting only peripheral myelin were excluded.

Ensembl gene IDs were used to analyze the data on 2 platforms. To categorize the genes by biologic process and metabolic process, we used the Gene Ontology (GO) PANTHER classification system (PANTHER14.1).[18-20] To conduct pathway analysis, we used Reactome, a biological pathway and process analysis database and visualization tool.[21] Seventy-six leukodystrophy genes could not be mapped to a gene or process in Reactome.

Results

Using a comprehensive review of PubMed, Google, published literature reviews, and commercial gene panels, we identified 399 unique genes with white matter MRI pathology on T2 sequences (Figure 1; Table 1). Of this, 87 (22%) genes were hypomyelinating. Only 3 genes had contrast enhancement on MRI (*ABCD1*, *GFAP*, and *UNC13D*) (Table 2).

Gene Ontology term evaluation showed that the most frequent categories of GWMD genes (Figure 2A) were metabolic processes (n = 161), cellular processes (n = 120), localization (n = 49), biological regulation (n = 34), and response to stimulus (n = 14; Supplemental Table 1). Interestingly, although the overall number of genes was fewer, the distribution and type of GO biological processes was very similar to the canonical leukodystrophy genes (Figure 2B; Supplemental Table 2).

Table 1. List of All Identified Genetic White Matter Disorders (GWMD) Genes.

Gene name	Ensembl ID
AARS	ENSG00000090861
AARS2	ENSG00000124608
ABAT	ENSG00000183044
ABCA1	ENSG00000165029
ABCD1	ENSG00000101986
ACDB5	ENSG00000107897
ACER3	ENSG00000078124
ACOX1	ENSG00000161533
ACP33	ENSG00000090487
ACP5	ENSG00000102575
ACSF3	ENSG00000176715
ADAR	ENSG00000160710
ADGRG1	ENSG00000205336
ADSL	ENSG00000239900
AGA	ENSG00000038002
AHDC1	ENSG00000126705
AIMP1	ENSG00000164022
AIMP2	ENSG00000106305
ALDH3A2	ENSG00000072210
ALDH5A1	ENSG00000112294
ALDH6A1	ENSG00000119711
ALDH7A1	ENSG00000164904
ALG12	ENSG00000182858
ALG13	ENSG00000101901
ALG2	ENSG00000119523
ALG6	ENSG00000088035
ALG9	ENSG00000086848
AMACR	ENSG00000242110
AMPD2	ENSG00000116337
AP4B1	ENSG00000134262
AP5Z1	ENSG00000242802
APOPT1	ENSG00000256053
APP	ENSG00000142192
ARHGAP31	ENSG00000031081
ARHGEF10	ENSG00000104728
ARNT2	ENSG00000172379
ARSA	ENSG00000100299
ASL	ENSG00000126522
ASNS	ENSG00000070669
ASPA	ENSG00000108381
ASS1	ENSG00000130707
ASXL1	ENSG00000171456
ATN1	ENSG00000111676
ATP7B	ENSG00000123191
ATPAF2	ENSG00000171953
ATRX	ENSG00000085224
AUH	ENSG00000148090
B3GALNT2	ENSG00000162885
BCAP31	ENSG00000185825
BCKDHA	ENSG00000248098
BCKDHB	ENSG00000083123
BCS1L	ENSG00000074582
BOLA3	ENSG00000163170
BRAT1	ENSG00000106009
BTD	ENSG00000169814
CARS2	ENSG00000134905
CDKL5	ENSG00000008086
CLCN2	ENSG00000114859
CLN8	ENSG00000182372
CLP1	ENSG00000172409
CLPP	ENSG00000125656
CNTNAP1	ENSG00000108797
COA7	ENSG00000162377
COG7	ENSG00000168434
COL4A1	ENSG00000187498
COQ2	ENSG00000173085
COQ8A	ENSG00000163050
COQ9	ENSG00000088682
COX10	ENSG00000006695
COX14	ENSG00000178449
COX15	ENSG00000014919
COX6B1	ENSG00000126267
COX7B	ENSG00000131174
COX8A	ENSG00000176340
CSF1R	ENSG00000182578
CTC1	ENSG00000178971
CTDP1	ENSG00000282752
CTSA	ENSG00000064601
CTSD	ENSG00000117984
CTSF	ENSG00000174080
CYP27A1	ENSG00000135929
CYP2U1	ENSG00000155016
CYP7B1	ENSG00000172817
D2HGDH	ENSG00000180902
DAG1	ENSG00000173402
DARS	ENSG00000115866
DARS2	ENSG00000117593
DBT	ENSG00000137992
DCAF17	ENSG00000115827
DCX	ENSG00000077279
DDC	ENSG00000132437
DDHD2	ENSG00000085788
DEAF1	ENSG00000177030
DGUOK	ENSG00000114956
DHFR	ENSG00000228716
DLD	ENSG00000091140
DMPK	ENSG00000104936
DNM1L	ENSG00000087470
DOCK6	ENSG00000130158
DOLK	ENSG00000175283
DPAGT1	ENSG00000172269
DPM1	ENSG00000000419
DPYD	ENSG00000188641
EARS2	ENSG00000103356
EHMT1	ENSG00000181090
EIF2B1	ENSG00000111361
EIF2B2	ENSG00000119718
EIF2B3	ENSG00000070785
EIF2B4	ENSG00000115211
EIF2B5	ENSG00000145191
ELOVL4	ENSG00000118402
EPG5	ENSG00000152223
EPRS	ENSG00000136628
ERCC2	ENSG00000104884

(continued)

Table 1. (continued)

Gene name	Ensembl ID
ERCC3	ENSG00000163161
ERCC6	ENSG00000225830
ERCC8	ENSG00000049167
ETFDH	ENSG00000171503
ETHE1	ENSG00000105755
FA2H	ENSG00000103089
FAM126A	ENSG00000122591
FARS2	ENSG00000145982
FASTKD2	ENSG00000118246
FBXL4	ENSG00000112234
FH	ENSG00000091483
FIG4	ENSG00000112367
FKRP	ENSG00000181027
FKTN	ENSG00000106692
FMR1	ENSG00000102081
FOLR1	ENSG00000110195
FOXC1	ENSG00000054598
FOXRED1	ENSG00000110074
FUCA1	ENSG00000179163
GAA	ENSG00000171298
GALC	ENSG00000054983
GALT	ENSG00000213930
GAN	ENSG00000261609
GBA	ENSG00000177628
GBE1	ENSG00000114480
GCDH	ENSG00000105607
GFAP	ENSG00000131095
GFM1	ENSG00000168827
GJA1	ENSG00000152661
GJB1	ENSG00000169562
GJC2	ENSG00000198835
GLA	ENSG00000102393
GLB1	ENSG00000170266
GLRX5	ENSG00000182512
GLUL	ENSG00000135821
GLYCTK	ENSG00000168237
GM2A	ENSG00000196743
GNAO1	ENSG00000087258
GNS	ENSG00000135677
GPHN	ENSG00000171723
HEPACAM	ENSG00000165478
HEXA	ENSG00000213614
HHH/ SLC25A15	ENSG00000102743
HIBCH	ENSG00000198130
HIKESHI	ENSG00000149196
HLCS	ENSG00000159267
HMBS	ENSG00000256269
HMGCL	ENSG00000117305
HSD17B10	ENSG00000072506
HSD17B4	ENSG00000133835
HSPD1	ENSG00000144381
HTRA1	ENSG00000166033
IBA57	ENSG00000181873
IDS	ENSG00000010404
IDUA	ENSG00000127415
IFIH1	ENSG00000115267
ISCA1	ENSG00000135070
ISCA2	ENSG00000165898
ITPA	ENSG00000125877
IVD	ENSG00000128928
JAM3	ENSG00000166086
KCNT1	ENSG00000107147
L2HGDH	ENSG00000087299
LAMA1	ENSG00000101680
LAMA2	ENSG00000196569
LAMB1	ENSG00000091136
LARGE1	ENSG00000133424
LETM1	ENSG00000168924
LIAS	ENSG00000121897
LIPT1	ENSG00000144182
LMNB1	ENSG00000113368
LRPPRC	ENSG00000138095
LYRM7	ENSG00000186687
MAG	ENSG00000105695
MAN2B1	ENSG00000104774
MANBA	ENSG00000109323
MARS2	ENSG00000247626
MAT1A	ENSG00000151224
MCCC1	ENSG00000078070
MCOLN1	ENSG00000090674
MECP2	ENSG00000169057
MEF2C	ENSG00000081189
MFSD8	ENSG00000164073
MGP	ENSG00000111341
MLC1	ENSG00000100427
MLYCD	ENSG00000103150
MMACHC	ENSG00000132763
MMADHC	ENSG00000168288
MOCS1	ENSG00000124615
MOCS2	ENSG00000164172
MOGS	ENSG00000115275
MPLKIP	ENSG00000168303
MPV17	ENSG00000115204
MRPS16	ENSG00000182180
MRPS22	ENSG00000175110
MTATP6	ENSG00000198899
MTFMT	ENSG00000103707
MTHFR	ENSG00000177000
MTHFS	ENSG00000136371
MTND1	ENSG00000198888
MTND5	ENSG00000198786
MTND6	ENSG00000198695
MTTC	ENSG00000210140
MTTF	ENSG00000210049
MTTH	ENSG00000210176
MTTK	ENSG00000210156
MTTL1	ENSG00000209082
MTTQ	ENSG00000210107
MTTS1	ENSG00000210151
MTTS2	ENSG00000210184
NADK2	ENSG00000152620
NAGLU	ENSG00000108784
NAGS	ENSG00000161653
NAXE	ENSG00000163382

(continued)

Table 1. (continued)

(continued)

Table 1. (continued)

Gene name	Ensembl ID
NDUFA10	ENSG00000130414
NDUFA12	ENSG00000184752
NDUFA2	ENSG00000131495
NDUFA9	ENSG00000139180
NDUFAF1	ENSG00000137806
NDUFAF2	ENSG00000164182
NDUFAF3	ENSG00000178057
NDUFAF4	ENSG00000123545
NDUFAF5	ENSG00000101247
NDUFAF6	ENSG00000156170
NDUFB3	ENSG00000119013
NDUFB9	ENSG00000147684
NDUFS1	ENSG00000023228
NDUFS2	ENSG00000158864
NDUFS3	ENSG00000213619
NDUFS4	ENSG00000164258
NDUFS6	ENSG00000145494
NDUFS7	ENSG00000115286
NDUFS8	ENSG00000110717
NDUFV1	ENSG00000167792
NDUFV2	ENSG00000178127
NFU1	ENSG00000169599
NGLY1	ENSG00000151092
NKX6-2	ENSG00000148826
NOTCH1	ENSG00000148400
NOTCH3	ENSG00000074181
NPC1	ENSG00000141458
NPC2	ENSG00000119655
NUBPL	ENSG00000151413
OAT	ENSG00000065154
OCLN	ENSG00000197822
OCRL	ENSG00000122126
OPA1	ENSG00000198836
OPA3	ENSG00000125741
OSGEP	ENSG00000092094
OSTM1	ENSG00000081087
OTC	ENSG00000036473
PAFAH1B1	ENSG00000007168
PAH	ENSG00000171759
PC	ENSG00000173599
PCCA	ENSG00000175198
PCCB	ENSG00000114054
PDHA1	ENSG00000131828
PDHX	ENSG00000110435
PEX1	ENSG00000127980
PEX10	ENSG00000157911
PEX12	ENSG00000108733
PEX13	ENSG00000162928
PEX14	ENSG00000142655
PEX16	ENSG00000121680
PEX19	ENSG00000162735
PEX26	ENSG00000215193
PEX5	ENSG00000139197
PEX6	ENSG00000124587
PGAP1	ENSG00000197121
PGN	ENSG00000197912
PHGDH	ENSG00000092621

(continued)

Table 1. (continued)

Gene name	Ensembl ID
PHYH	ENSG00000107537
PIGA	ENSG00000165195
PLA2G6	ENSG00000184381
PLEKHG2	ENSG00000090924
PLP1	ENSG00000123560
PMM2	ENSG00000140650
PMP22	ENSG00000109099
POLG1	ENSG00000140521
POLG2	ENSG00000256525
POLR1A	ENSG00000068654
POLR1C	ENSG00000171453
POLR3A	ENSG00000148606
POLR3B	ENSG00000013503
POMGNT1	ENSG00000085998
POMK	ENSG00000185900
POMT1	ENSG00000130714
POMT2	ENSG00000009830
PPP1R15B	ENSG00000158615
PPT1	ENSG00000131238
PRF1	ENSG00000180644
PRKDC	ENSG00000253729
PRODH	ENSG00000100033
PRUNE1	ENSG00000143363
PSAP	ENSG00000197746
PSAT1	ENSG00000135069
PSEN1	ENSG00000080815
PURA	ENSG00000185129
PYCR2	ENSG00000143811
QARS	ENSG00000172053
RAB11B	ENSG00000185236
RARS	ENSG00000113643
RARS2	ENSG00000146282
RMND1	ENSG00000155906
RNASEH2A	ENSG00000104889
RNASEH2B	ENSG00000136104
RNASEH2C	ENSG00000172922
RNASET2	ENSG00000026297
RNF216	ENSG00000011275
RPIA	ENSG00000153574
RPS6KC1	ENSG00000136643
RRM2B	ENSG00000048392
RXYLT1	ENSG00000118600
SAMHD1	ENSG00000101347
SCO2	ENSG00000130489
SCP2	ENSG00000116171
SDHA	ENSG00000073578
SDHAF1	ENSG00000205138
SDHB	ENSG00000117118
SDHD	ENSG00000204370
SEPSECS	ENSG00000109618
SGSH	ENSG00000181523
SHPK	ENSG00000197417
SLC13A5	ENSG00000141485
SLC16A2	ENSG00000147100
SLC17A5	ENSG00000119899
SLC1A4	ENSG00000115902
SLC25A1	ENSG00000100075

(continued)

Table 1. (continued)

Gene name	Ensembl ID
SLC25A12	ENSG00000115840
SLC25A22	ENSG00000177542
SLC33A1	ENSG00000169359
SLC35A2	ENSG00000102100
SLC46A1	ENSG00000076351
SNIP1	ENSG00000163877
SNORD118	ENSG00000200463
SOD1	ENSG00000142168
SOX10	ENSG00000100146
SP110	ENSG00000135899
SPATA5	ENSG00000145375
SPG11	ENSG00000104133
SPG20	ENSG00000133104
SPTAN1	ENSG00000197694
SRD5A3	ENSG00000128039
STAMBP	ENSG00000124356
STN1	ENSG00000107960
STXBP1	ENSG00000136854
STXBP2	ENSG00000076944
SUCLA2	ENSG00000136143
SUMF1	ENSG00000144455
SUOX	ENSG00000139531
SURF1	ENSG00000148290
TACO1	ENSG00000136463
TAF2	ENSG00000064313
TBX1	ENSG00000184058
TCF4	ENSG00000196628
TCIRG1	ENSG00000110719
TM4SF20	ENSG00000168955
TMEM106B	ENSG00000106460
TMEM165	ENSG00000134851
TMEM70	ENSG00000175606
TMTC3	ENSG00000139324
TRAPPC9	ENSG00000167632
TREM2	ENSG00000095970
TREX1	ENSG00000213689
TRMT5	ENSG00000126814
TSC1	ENSG00000165699
TSEN54	ENSG00000182173
TUBB4A	ENSG00000104833
TUFM	ENSG00000178952
TWNK	ENSG00000107815
TYMP	ENSG00000025708
TYROBP	ENSG00000011600
UBE2A	ENSG00000077721
UBE3A	ENSG00000114062
UFM1	ENSG00000120686
UGT1A1	ENSG00000241635
UNC13D	ENSG00000092929
UPB1	ENSG00000100024
VARS2	ENSG00000137411
VPS11	ENSG00000160695
WT1	ENSG00000184937
WWOX	ENSG00000186153
ZFYVE26	ENSG00000072121
ZNF335	ENSG00000198026
ZNF9	ENSG00000169714

Table 2. List of Genes With Hypomyelination, List of Genes With Contrast Enhancement.

Hypomyelinating	
AARS	PRKDC
AIMP2	PRUNE1
ALG2	PURA
B3GALNT2	PYCR2
BCAP31	QARS
CLCN2	RARS
CNTNAP1	RMND1
DARS	RRM2B
DDC	SGSH
DPYD	SLC16A2
EPRS	SLC17A5
ERCC2	SLC1A4
ERCC3	SLC25A1
ERCC6	SLC25A12
ERCC8	SLC33A1
FAM126A	SNIP1
FOLR1	SOX10
FUCA1	SPATA5
GJA1	SPG11
GJC2	SPTAN1
GLB1	STAMBP
GLUL	STXBP1
HIKESHI	TMEM106B
HSPD1	TMTC3
MMADHC	TSC1
MPLKIP	TUBB4A
MTHFS	UFM1
NKX6-2	VSP11
NPC1	WT1
NPC2	ZNF335
OSTM1	
PAH	
PLP1	
POLR1C	
POLR3A	
POLR3B	
POMK	
Contrast enhancement	
ABCD1	
GFAP	
UNC13D	

A subgroup analysis of the single largest GO term of GWMD genes, "metabolic process," showed that the most frequent GO terms in this group were organic substance metabolic process (n = 119), cellular metabolic process (n = 63), primary metabolic process (n = 20), oxidation reduction process (n = 19), and catabolic process (n = 19; Figure 2C; Supplemental Table 3).

We used a biological pathway analysis tool, Reactome,[21] to identify whether GWMD genes were more represented in certain processes or shared common biological features (Table 3).

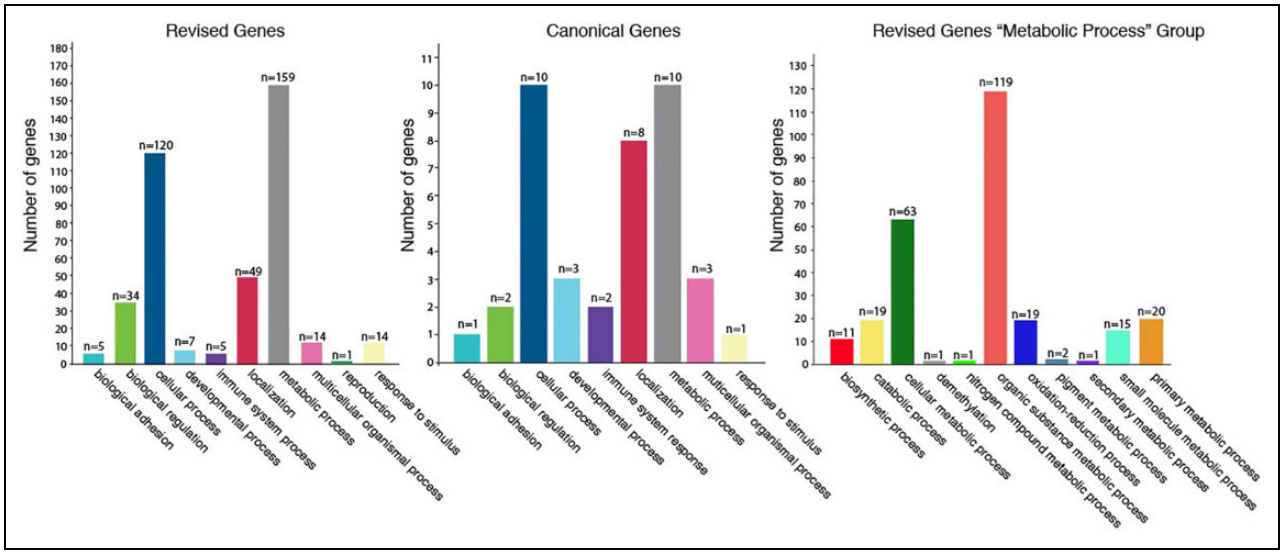

Figure 2. A, Revised genetic white matter disorders (GWMD) genes organized by Gene Ontology (GO) term biological process. B, Thirty canonical leukodystrophy genes organized by GO term biological process. C, Revised GWMD genes in the category "Metabolism" displayed by subtypes of metabolic processes.

Table 3. Reactome Pathway Listing of the 25 Most Overrepresented Biological Pathways, Grouped by Biological Mechanisms, and From Most to Fewest Number of Genes.[a]

	Genes		Reactions	
Pathway name	n/total	P	FDR	n/total
Metabolism				
Metabolism	177/5569	2.18e-9	2.33e-7	250/2213
Metabolism of amino acids and derivatives	41/931	1.54e-5	0.001	41/283
Diseases of metabolism	23/303	3.04e-7	2.31e-5	33/114
Metabolism of water-soluble vitamins and cofactors	20/377	2.56e-4	0.009	27/140
Defects in vitamin and cofactor metabolism	8/70	1.67e-4	0.008	9/22
Defects in biotin metabolism	6/34	1.11e-4	0.006	6/6
Biotin transport and metabolism	6/48	6.85e-4	0.023	9/13
Multiple carboxylase deficiency	5/32	7.18e-4	0.023	4/4
Mitochondrial				
Citric acid cycle and respiratory electron transport	45/404	1.11e-16	2.79e-14	30/65
Respiratory electron transport, ATP synthesis, heat production	38/273	1.11e-16	2.79e-14	20/29
Respiratory electron transport	37/215	1.11e-16	2.79e-14	17/19
Complex I biogenesis	25/144	3.66e-15	6.89e-13	13/13
Protein				
Protein localization	29/244	2.87e-13	4.30e-11	45/53
tRNA aminoacylation	15/232	1.99e-4	0.008	19/42
Recycling of eIF2:GDP	5/36	0.001	0.036	2/2
Peroxisomal				
Peroxisomal protein import	17/114	9.31e-10	1.16e-7	23/26
Class I peroxisomal protein import	9/40	3.08e-7	2.31e-5	6/6
Glycosylation				
Diseases of glycosylation	22/234	1.48e-8	1.39e-6	24/77
Diseases associated with glycosylation precursor biosynthesis	7/65	5.99e-4	0.021	8/16
Diseases associated with *N*-glycosylation of proteins	7/49	1.11e-4	0.006	8/23
Defective POMT1	3/5	1.90e-4	0.008	1/1
Defective POMT2	3/5	1.90e-4	0.008	1/1
Other				
Branched chain amino acid catabolism	10/106	1.26e-4	0.007	11/28
Mucopolysaccharidoses	6/37	1.75e-4	0.008	12/22
Loss of MECP2 binding to DNA	2/2	8.98e-4	0.028	1/1

Abbreviation: FDR, false discovery rate.
[a]Many genes are counted in more than one category (eg, metabolism, diseases of metabolism).

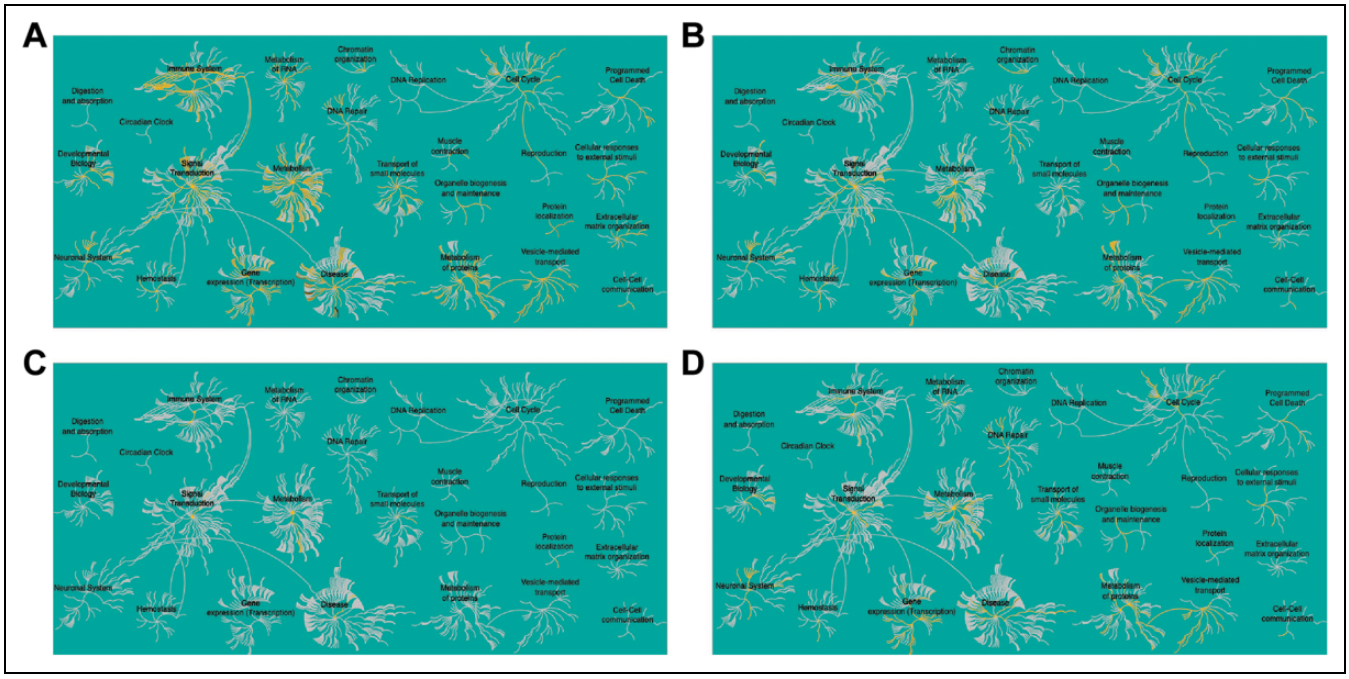

Figure 3. Reactome pathway analysis of genetic white matter disorders (GWMD) genes. Analysis is arranged in a hierarchy, with the center of each circular "burst" as the root of one top-level pathway. Each step away from center represents the next level lower in the pathway hierarchy. Yellow-coded pathways are significantly overrepresented; light gray signifies pathways not significantly overrepresented. A, Reactome pathway analysis of entire revised GWMD gene set. B, Reactome pathway analysis of 30 canonical leukodystrophy genes. C, Reactome pathway analysis of contrast-enhancing genes. D, Reactome pathway analysis of hypomyelinating gene set.

An analysis of the 25 most significantly represented biological pathways revealed that the majority of GWMD genes were involved in just 2 general categories: metabolism (metabolism, diseases of metabolism, metabolism of amino acids, biotin metabolism, defects in vitamin and cofactor metabolism, metabolism of water soluble vitamins and cofactors, biotin transport) and respiratory electron transport/mitochondrial function (respiratory electron transport; respiratory electron transport, ATP synthesis, and heat production; citric acid cycle; complex I biogenesis) (Figure 3).

We also manually evaluated the biological roles of GWMD genes, to confirm the GO and Reactome classifications, as well as to evaluate in greater details gene functions. Genes with roles in the mitochondria or mitochondrial function (*COX7*, *HSPD1*, *RMND1*, etc) were the single largest group. Interestingly, although as expected genes with lysosomal or peroxisomal roles were frequent, GWMD genes that are transcription factors were approximately as frequent (*MEF2C*, *SOX10*, *TAF2*, etc).

Discussion

We have identified a significantly greater number of genes than previously recognized, 399, that are associated with myelin signal changes on T2 MRI. This larger group of GWMD (leukodystrophy and leukoencephalopathy) genes was similar in GO group composition to previous more restrictive definitions of leukodystrophy genes.[7]

Of a total of 27 possible biological pathways represented in the analysis tool Reactome, GWMD genes were present in 23 of those groups, confirming the diverse potential etiologies of GWMDs. Genes involved in metabolic pathways were the most represented group of genes.

While nearly 400 genes is a significantly larger number of genes associated with GWMDs than previously considered, it is only a small proportion (1.9%) of the estimated 21 000 protein-coding genes in the entire human genome. From this perspective, given the complexities of myelin development and maintenance, and the diverse cell types that can affect myelin involved including oligodendrocytes, astrocytes, neurons, and microglia, 399 genes seem proportionate.

The definition of leukodystrophies has been a contentious and at times divisive topic. An initial organized attempt was made in 2015,[7] but already in a short period of time new data suggested potential revisions to this list of approximately 30 genes.[8]

Our approach consisted solely of inclusion based on the presence of white matter T2 signal hyperintensity on MRI and presumed/proven genetic etiology. This methodology poses certain limitations, in that there is no consistent pathophysiology. However, this limitation is also a strength in avoiding certain biases. Since T2 signal hyperintensity of the myelin is essentially a defining term of glial/myelin sheath abnormality,[22] this meets the Vanderver et al[7] inclusion criteria. Further, we avoided exclusion criteria that could be construed as arbitrary. For example, when considering inborn errors of metabolism, lysosomal sialic acid

storage disorder (Salla disease) met inclusion but the lysosomal disorder Niemann-Pick C did not.[7]

This finding of a large number of genes that can cause a white matter disorder (leukodystrophy or leukoencephalopathy) highlights that early use of an NGS approach such as whole exome sequencing or whole genome sequencing should be considered as a first-line diagnostic approach. With so many different genes that can cause similar T2 signal changes, NGS can provide lower costs and faster time to diagnosis.[23] For the clinician, this information about the many different genes that can cause GWMD further emphasize the need for early use of NGS in diagnosis.

An important and unresolved question is why this diversity of different genes all cause white matter pathology. In the undertaking of this project, we hypothesized that shared biological mechanisms and pathophysiology would be revealed. We did observe common themes, including overrepresentation of genes involved in metabolism and in mitochondrial function. This suggests, and is concordant with commonly accepted understanding, that the white matter is particularly sensitive to disturbances in metabolism and in energy homeostasis. It is possible that therapies directed toward these downstream targets (metabolic and energy homeostasis) could provide broad benefits for many different GWMD. Another interesting issue is the phenotypic variability, including age of onset and disease severity. This phenotypic diversity is seen even within the same disease, such as X-linked adrenoleukodystrophy or metachromatic leukodystrophy. Thus, while it is not currently possible to generalize about phenotypic presentation or age of onset, perhaps there are patterns of severity that could be experimentally explored. For example, whether diseases with more profound disturbances of energy homeostasis cause an earlier and more severe presentation.

Conclusions

We found 399 genes that are associated with white matter changes on T2 MR image sequences. This is approximately 10-fold higher than has been standardly considered as the number of genes responsible for leukodystrophies. There are not consistent biological differences between this revised list and previous definitions of leukodystrophy genes. This expanded understanding of the genetics of GWMDs including leukodystrophies and leukoencephalopathies can be useful in analysis and interpretation of NGS results for diagnosis and in understanding the pathophysiology of GWMDs.

Authors' Note

VMU, MS, and HS contributed equally to the manuscript. All data reported in this study are included in this publication.

Author Contributions

VMU, MS, and JLB contributed to conception and design. JLB drafted manuscript. All authors contributed to acquisition, analysis, and interpretation; critically revised manuscript; gave final approval; and agrees to be accountable for all aspects of work ensuring integrity and accuracy.

Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: M.S. is an employee of NXP Semiconductor. J.L.B. has served as a consultant to Bluebird Bio, Calico Life Sciences, Denali Inc, Enzyvant, and Neurogene; is on the board of directors of wFluidx Inc; and owns stock in Orchard Therapeutics.

Ethical Approval

The University of Utah IRB granted this work an exemption as nonhuman subjects research.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: J.L.B. was supported by the Bray Chair in Child Neurology Research and by NIH grant 3UL1TR002538-01S1. V.M.U. was supported by NIH grant T35HL007744.

ORCID iD

Joshua L. Bonkowsky, MD, PhD https://orcid.org/0000-0001-8775-147X

Supplemental Material

Supplemental material for this article is available online.

References

1. Bielschowsky MH, Henneberg R. Über familiäre diffuse Sklerose (Leukodystrophia cerebri progressiva hereditaria). *J Psychol Neurol*. 1928;36:131-181.
2. Kaye EM. Update on genetic disorders affecting white matter. *Pediatr Neurol*. 2001;24(1):11-24.
3. Maria BL, Deidrick KM, Moser H, Naidu S. Leukodystrophies: pathogenesis, diagnosis, strategies, therapies, and future research directions. *J Child Neurol*. 2003;18(9):578-590.
4. Bonkowsky JL, Nelson CR, Kingston JL, Filloux FM, Mundorff MB, Srivastava R. The burden of inherited leukodystrophies in children. *Neurology*. 2010;75(8):718-725. doi:10.1212/WNL.0b013e3181eee46b
5. van der Knaap MS, Bugiani M. Leukodystrophies: a proposed classification system based on pathological changes and pathogenetic mechanisms. *Acta Neuropathol*. 2017;134(3):351-382.
6. Vanderver A, Simons C, Helman G, et al. Whole exome sequencing in patients with white matter abnormalities. *Ann Neurol*. 2016;79(6):1031-1037.
7. Vanderver A, Prust M, Tonduti D, et al. Case definition and classification of leukodystrophies and leukoencephalopathies. *Mol Genet Metab*. 2015;114(4):494-500.
8. Kevelam S, Steenweg M, Srivastava S, et al. Update on leukodystrophies: a historical perspective and adapted definition. *Neuropediatrics*. 2016;47(6):349-354.
9. Ensembl 2018. Accessed February 15, 2019. http://apr2018.archive.ensembl.org

10. Kolodny EH. Dysmyelinating and demyelinating conditions in infancy. *Curr Opin Neurol Neurosurg*. 1993;6(3):379-386.
11. Vanderver A. Tools for diagnosis of leukodystrophies and other disorders presenting with white matter disease. *Curr Neurol Neurosci Rep*. 2005;5(2):110-118.
12. Wang X, He F, Yin F, et al. The use of targeted genomic capture and massively parallel sequencing in diagnosis of Chinese leukoencephalopathies. *Sci Rep*. 2016;6:35936.
13. Kohlschütter A, Bley A, Brockmann K, et al. Leukodystrophies and other genetic metabolic leukoencephalopathies in children and adults. *Brain Dev*. 2010;32(1):82-89.
14. Pouwels PJ, Vanderver A, Bernard G, et al. Hypomyelinating leukodystrophies: translational research progress and prospects. *Ann Neurol*. 2014;76(1):5-19.
15. Parikh S, Bernard G, Leventer RJ, et al. A clinical approach to the diagnosis of patients with leukodystrophies and genetic leukoencephelopathies. GLIA Consortium. *Mol Genet Metab*. 2015; 114(4):501-515.
16. Ashrafi MR, Tavasoli AR. Childhood leukodystrophies: a literature review of updates on new definitions, classification, diagnostic approach and management. *Brain Dev*. 2017;39(5):369-385.
17. Ji H, Li D, Wu Y, et al. Hypomyelinating disorders in China: the clinical and genetic heterogeneity in 119 patients. *PLoS One*. 2018;13(2):e0188869.
18. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25-29.
19. The Gene Ontology Consortium. The gene ontology resource: 20 years and still going strong. *Nucleic Acids Res*. 2019;47(D1): D330-D338.
20. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019;47(D1):D419-D426.
21. Fabregat A, Jupe S, Matthews L, et al. The reactome pathway knowledgebase. *Nucleic Acids Res*. 2018;46(D1):D649-D655.
22. Schiffmann R, van der Knaap MS. Invited article: an MRI-based approach to the diagnosis of white matter disorders. *Neurology*. 2009;72(8):750-759.
23. Richards J, Korgenski EK, Taft RJ, Vanderver A, Bonkowsky JL. Targeted leukodystrophy diagnosis based on charges and yields for testing. *Am J Med Genet A*. 2015;167A(11):2541-2543.