# Supporting Information

Pierce et al. 10.1073/pnas.1115888108

## SI Materials and Methods

**Method for Quantification of Xylulose in Plasma.** ***Sample preparation.*** Plasma was deproteinized by mixing 300 μL of plasma with 1,000 μL of ice-cold acetone (HPLC grade; Burdick & Jackson), vortexing for 20 s, and incubating on ice for 20 min. After centrifugation at 14,000 × *g* for 20 min at 4 °C, the supernatant was removed, lyophilized, and reconstituted in 100 μL 99.9%/0.1% water/formic acid. After centrifugation at 14,000 × *g* for 10 min, 10 μL of the resulting supernatant was mixed with 90 μL of acetonitrile (HPLC grade; Burdick & Jackson).

***Liquid chromatography-mass spectrometry.*** Hydrophilic interaction liquid chromatography (HILIC) mass spectrometry (MS) was performed using an Agilent 1100 HPLC system coupled to a dual pressure linear ion trap (LTQ-Velos) mass spectrometer (Agilent). Five microliters of sample were injected into a 100-μL loop and flushed onto a commercial 2.0-mm × 10.0-cm column packed with 5 μm Amide 80 stationary phase (Tosoh Biosciences) at a constant flow-rate of 200 μL/min. Solvents A and B consisted of 100% water and 100% acetonitrile, respectively. Xylulose was eluted using a 10-min analysis time in which the solvent composition remained constant (5% A) for 1 min, ramped to 20% A over 4 min, held constant for 0.5 min (20% A), and re-equilibrated at initial conditions from 5.5 to 10 min (5% A). Atmospheric-pressure chemical ionization (APCI) was performed using the commercial ion max source from ThermoFisher operating in negative ion mode. Source parameters were as follows: (*i*) current set to 10 μA, (*ii*) vaporizer temperature of 475 °C, and (*iii*) sheath and auxiliary gas flow-rate set to 30 and 10 arbitrary units, respectively. Chloroform was infused at 5 μL/min post-column and promoted negative ionization of xylulose via chlorine adduction, significantly increasing sensitivity. For detection by mass spectrometry, a selected ion monitoring scan was centered on the nominal mass of chlorine adducted xylulose with a total scan range of 10 *m*/*z* (185 *m*/*z* ± 5 *m*/*z*). Xcalibur version 2.1.0.1139 was used for peak detection and integration.

***Absolute abundance determination.*** Method of standard addition was used to quantitatively measure xylulose in plasma. L-Xylulose (95% purity; Sigma Aldrich) was added to control plasma, before sample cleanup, to obtain absolute amounts of 3,120, 1,040, 346, 115, and 38 pmoles on the column. A calibration curve was generated by plotting integrated peak areas as a function of pmoles on the column in logarithmic form and used to determine the absolute amounts of xylulose in the unknown samples. It is important to note that in the control and heterozygous samples, xylulose was either not detected or was at an amount well below the lowest concentration standard.

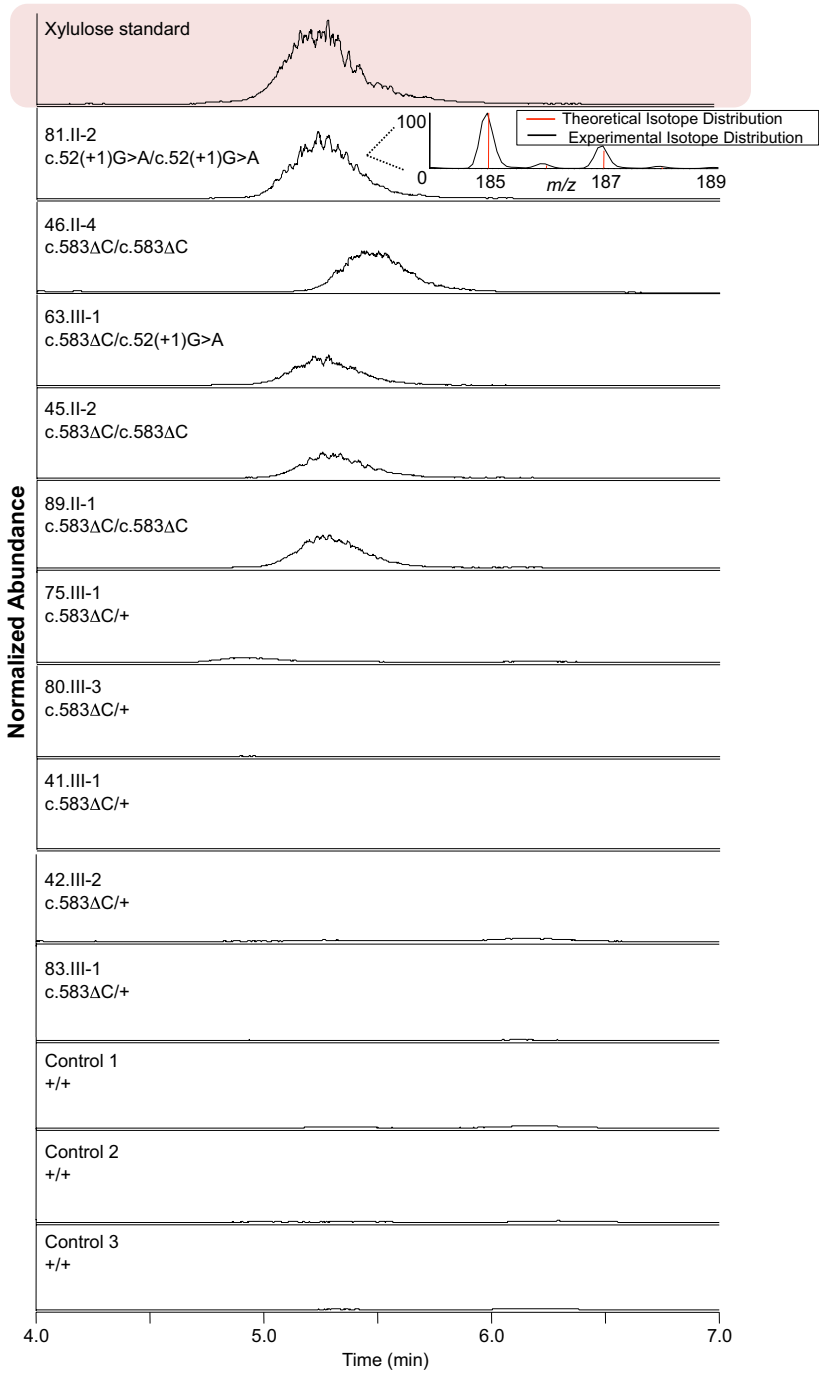


**Fig. S1.** Quantification of xylulose in blood plasma. Total ion chromatograms are shown for the analysis of the xylulose calibration standard and 13 subjects. The peak at ~5.2 min corresponds to chlorine adducted xylulose $[M+Cl]^-$. The identity of this peak was validated by mass and retention time comparisons with a commercial standard. The high degree of similarity between the experimental and theoretical isotope distribution further validates the identification (*Inset*). Chromatograms are normalized to the highest intensity peak (sample 81.II-2). All samples with high levels of xylulose ($S/n \geq 60$) are from individuals diagnosed with pentosuria and deficient in DCXR protein. In most samples from control individuals or *DCXR* mutation carriers, detection of xylulose at an acceptable noise level ($S/n > 3$) was not observed (samples 80.III-3, 41.III-1, 83.III-1, Control 2, Control 3). *DCXR* genotype is indicated below the sample number.

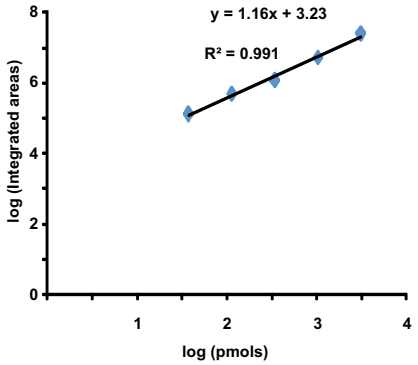


**Fig. S2.** Calibration curve for xylulose concentration in plasma. A logarithmic plot of integrated peak areas versus amounts of xylulose added to control plasma is shown. This calibration plot was used to calculate the absolute abundance of xylulose in samples from individuals of various *DCXR* genotypes. A high degree of linearity was observed across two orders of magnitude. The limit-of-detection was defined as the concentration determined from the x-intercept of the given equation ($1.7 \times 10^{-5}$ mg/dL).

**Table S1. Xylulose levels in plasma**

| Subject | *DCXR* genotype | Xylulose concentration (mg/dL) |
|---|---|---|
| 81.II-2 | c.52(+1)G > A/c.52(+1)G > A | 1.91 |
| 46.II-4 | c.583ΔC/c.583ΔC | 1.39 |
| 63.III-1 | c.583ΔC/c.52(+1)G > A | 1.00 |
| 45.II-2 | c.583ΔC/c.583ΔC | 0.87 |
| 89.II-1 | c.583ΔC/c.583ΔC | 1.08 |
| 75.III-1 | c.583ΔC/+ | 0.23* |
| 80.III-3 | c.583ΔC/+ | ND |
| 41.III-1 | c.583ΔC/+ | ND |
| 42.III-2 | c.583ΔC/+ | 0.06* |
| 83.III-1 | c.583ΔC/+ | ND |
| Control 1 | +/+ | 0.01* |
| Control 2 | +/+ | ND |
| Control 3 | +/+ | ND |

*Extrapolated from the calibration curve. ND, none detected.

**Table S2. Test of Hardy-Weinberg Equilibrium of *DCXR* mutations**

| *DCXR* genotype | c.538ΔC /c.538ΔC | c.538ΔC /c.52(+1)G > A | c.52(+1)G > A /c.52(+1)G > A | Both alleles |
|---|---|---|---|---|
| Expected frequency of genotype | 1.85E-04 | 1.01E-04 | 1.41E-05 | 2.99E-04 |
| Expected among nine unrelated pentosurics | 5.56 | 3.03 | 0.41 | 9.00 |
| Observed among nine unrelated pentosurics | 6 | 2 | 1 | 9 |
| X^2 (*P* value) | | | | 1.50 (0.47) |

**Table S3. European origins of families with *DCXR* mutations, as reported to Margaret Lasker (by the pentosuria families) or to the authors (by the controls)**

| Lasker family ID | Source of genotype | *DCXR* genotype | European origins of the family |
|---|---|---|---|
| Pentosuria families | | | |
| 41 | Children | Heterozygous c.583ΔC | Russia; Bessarabia |
| 42 | Child | Heterozygous c.583ΔC | Rozdil, Ukraine (near Lvov) |
| 45 | Proband | Homozygous c.583ΔC | Germany; France |
| 46 | Proband | Homozygous c.583ΔC | Kovel, Ukraine; Kashivka, Ukraine |
| 55 | Proband | Compound heterozygote c.583ΔC/c.52(+1)G > A | Dubno, Ukraine; Odessa, Ukraine; Kiev, Ukraine |
| 63 | Proband | Compound heterozygote c.583ΔC/c.52(+1)G > A | Poland; Russia; Suvalkai region, Lithuania |
| 64 | Proband | Homozygous c.583ΔC | Dvinsk, Latvia; Russia |
| 72 | Child | Heterozygous c.583ΔC | Russia; Germany |
| 75 | Child | Heterozygous c.583ΔC | Lodz, Poland; Kovel, Ukraine |
| 80 | Children | Heterozygous c.583ΔC | Russia |
| 81 | Proband | Homozygous c.52(+1)G > A | Giessen, Germany; Grebenhain, Germany |
| 83 | Child | Heterozygous c.583ΔC | Kovel, Ukraine |
| 88 | Proband | Homozygous c.583ΔC | Lvov, Ukraine; Russia; Bessarabia |
| 89 | Proband | Homozygous c.583ΔC | Vienna, Austria; Ganserndorf, Austria; Husiatyn, Ukraine; Buchach, Ukraine |
| 94 | Probands | Homozygous c.583ΔC | Lvov, Ukraine |
| Unrelated AJ controls heterozygous for *DCXR* pentosuria alleles (1,067 controls genotyped) | | | |
| 1 | | Heterozygous c.583ΔC | Bohemia (Czech Republic) |
| 2 | | Heterozygous c.583ΔC | Klaipeda, Lithuania |
| 3 | | Heterozygous c.583ΔC | Warsaw, Poland |
| 4 | | Heterozygous c.583ΔC | Drohobych, Ukraine; Sambir, Ukraine |
| 5 | | Heterozygous c.583ΔC | Poland; Russia |
| 6 | | Heterozygous c.583ΔC | Hungary |
| 7 | | Heterozygous c.583ΔC | Riga, Latvia |
| 8 | | Heterozygous c.583ΔC | Polack, Belarus |
| 9 | | Heterozygous c.583ΔC | Odessa, Ukraine |
| 10 | | Heterozygous c.583ΔC | Austria |
| 11 | | Heterozygous c.583ΔC | Vilnius, Lithuania; Romania |
| 12 | | Heterozygous c.583ΔC | Hungary |
| 13 | | Heterozygous c.583ΔC | Vilnius, Lithuania |
| 14 | | Heterozygous c.583ΔC | Jekabpils, Latvia; Romania |
| 15 | | Heterozygous c.583ΔC | Bialystok, Poland; Lithuania; Hungary |
| 16 | | Heterozygous c.583ΔC | Odessa, Ukraine; Czech Republic |
| 17 | | Heterozygous c.583ΔC | Budapest, Hungary; Vac, Hungary |
| 18 | | Heterozygous c.583ΔC | Minsk, Belarus |
| 19 | | Heterozygous c.583ΔC | Vilnius, Lithuania |
| 20 | | Heterozygous c.583ΔC | Warsaw, Poland; Riga, Latvia |
| 21 | | Heterozygous c.583ΔC | Eastern Europe |
| 22 | | Heterozygous c.583ΔC | Minsk, Belarus |
| 23 | | Heterozygous c.583ΔC | Hungary, Poland |
| 24 | | Heterozygous c.583ΔC | Germany |
| 25 | | Heterozygous c.583ΔC | Germany; Poland |
| 26 | | Heterozygous c.583ΔC | Germany |
| 27 | | Heterozygous c.583ΔC | Austria; Russia |
| 28 | | Heterozygous c.583ΔC | Lithuania; Russia |
| 29 | | Heterozygous c.583ΔC | Lithuania; Poland; Ukraine |
| 30 | | Heterozygous c.52(+1)G > A | Lithuania; Russia |
| 31 | | Heterozygous c.52(+1)G > A | Vaukavysk, Belarus; Odessa, Ukraine |
| 32 | | Heterozygous c.52(+1)G > A | Lodz, Poland; Czlopa, Poland |
| 33 | | Heterozygous c.52(+1)G > A | Warsaw, Poland; Estonia |
| 34 | | Heterozygous c.52(+1)G > A | Germany; Russia |
| 35 | | Heterozygous c.52(+1)G > A | Kozova, Ukraine; Zolachiv, Ukraine |
| 36 | | Heterozygous c.52(+1)G > A | Poland |
| 37 | | Heterozygous c.52(+1)G > A | Russia |

European regions of origin reported by our participants may only partially reflect the origins of *DCXR* alleles for several reasons: not all lineages of a family harbor a *DCXR* mutation, family information may be incomplete, or the origin of the mutation may precede the family's residence in the area.

**Table S4. Homozygous regions on chr 17q25.3 for pentosuric subjects and Ashkenazi Jewish controls**

| Subject | *DCXR* genotype | start (hg19) | end (hg19) | Length (bp) | Start SNP | End SNP |
|---|---|---|---|---|---|---|
| | *DCXR* locus | 79,993,758 | 79,995,573 | 1,815 | | |
| Control 101 | Wild-type | 79,993,212 | 80,505,809 | 512,597 | rs4969481 | rs4614761 |
| Control 102 | Wild-type | 79,614,154 | 80,509,676 | 895,522 | rs9905639 | rs10852797 |
| Control 103 | Wild-type | 79,993,212 | 80,505,809 | 512,597 | rs4969481 | rs4614761 |
| Control 104 | Wild-type | 79,954,544 | 80,020,812 | 66,268 | rs8074498 | rs11542332 |
| Control 105 | Wild-type | 79,993,212 | 80,421,870 | 428,658 | rs4969481 | rs4789693 |
| Control 106 | Wild-type | 79,869,593 | 80,509,676 | 640,083 | rs115208976 | rs10852797 |
| Control 107 | Wild-type | 79,923,718 | 80,008,392 | 84,674 | rs4239275 | rs9916764 |
| Control 108 | Wild-type | 79,993,212 | 80,505,809 | 512,597 | rs4969481 | rs4614761 |
| Control 109 | Wild-type | 79,923,718 | 80,008,392 | 84,674 | rs4239275 | rs9916764 |
| Control 110 | Wild-type | 79,974,731 | 80,037,191 | 62,460 | rs13087 | rs4075080 |
| Control 111 | Wild-type | 79,959,659 | 80,008,392 | 48,733 | rs9912335 | rs9916764 |
| 45.II-2 | 79,994,115 del C | 78,368,051 | 81,006,629 | 2,638,578 | rs8359 | rs7406119 |
| 89.II-1 | 79,994,115 del C | 79,574,124 | 80,796,236 | 1,222,112 | rs11652797 | rs4986113 |
| 94.II-3 | 79,994,115 del C | 79,059,230 | 80,439,733 | 1,380,503 | rs8080815 | rs9909476 |
| 81.II-2 | 79,995,506 G > A | 77,232,086 | 81,006,629 | 3,774,543 | rs7223911 | rs7406119 |

**Table S5. *DCXR* primers**

| Exon | Forward primer (5′ to 3′) | Reverse primer (5′ to 3′) |
|---|---|---|
| 1 | ggacaaggctgtcaagatcc | atcgtggccagagcttca |
| 2 | gagtcgcgcacagaggta | gggaaaggtgggcaaagg |
| 3 | agcccctccctgaggttc | ctggtaactgccccctgata |
| 4 and 5 | agctggggcactcacagtag | tctctgaccagggtctgtcc |
| 6 and 7 | ctacccagggcacacacag | ccagcttgggttcctcttc |
| 8 | gggcagcagaatcaggttta | cttcctctatgcccctgacc |