**Supplementary Data**

**Figure S1.** *GSTZ1* mutations in MHSA individuals. The figure shows a *GSTZ1* cDNA including the boundaries of the 9 exons. Above the cDNA are shown the mutations described in this article. Asterisks indicate the 3 codons used to haplotype the individuals described in this article (94G (Glu32), 124G (Gly42), 245T (Met82) in the D haplotype). Below the cDNA, blue circles indicate the positions of GSH binding residues and red circles indicate MAA binding residues.

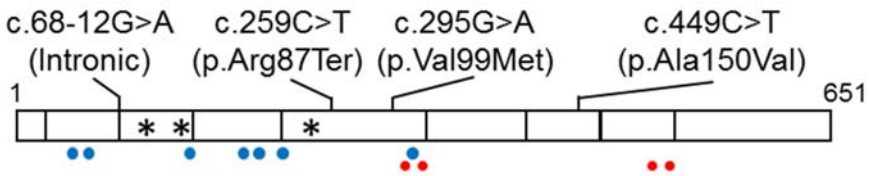
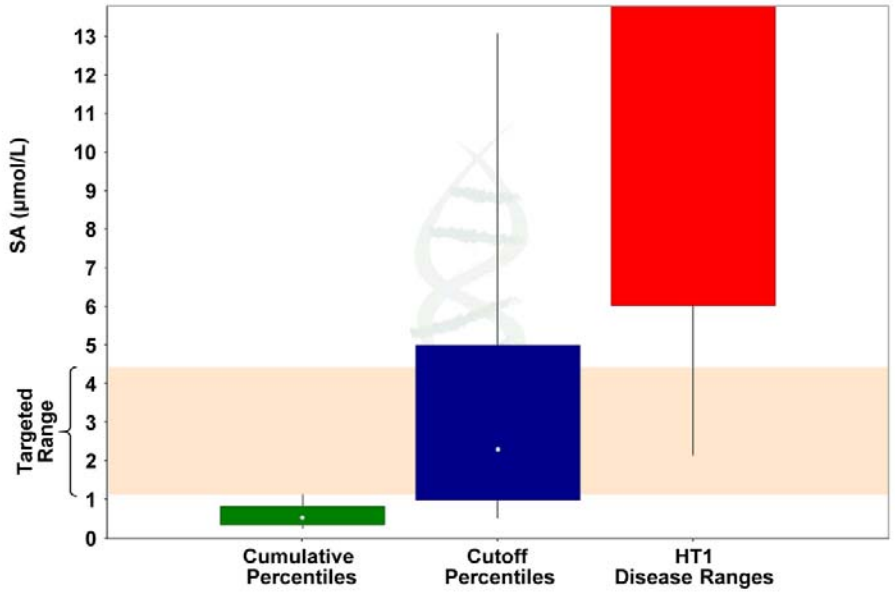


**Figure S2.** Succinylacetone (SA) concentrations in newborn screening bloodspots; graphs obtained from the Region 4S MS/MS Collaborative Project website (http://www.clir-r4s.org/), accessed 6th October 2016.
(A) Cumulative parameters, compiled from data contributed by 54 participating newborn screening programs. Control population data are shown in dark green, cutoff values in dark blue, values from newborns with tyrosinemia type 1 (HT1) in red. Within each group, the upper whisker end corresponds to the 99th percentile; top of box, 90th percentile; white circle within box, median; bottom of box, 10th percentile; lower whisker end, 1st percentile. The 'targeted range', shown as an orange rectangle, is defined as the interval between the cumulative 99th percentile of SA in normal populations and the 5th percentile of SA observed in newborns with tyrosinemia type 1.

(B) Data from each of the 54 newborn screening programs, presented separately. Their normal control populations in light green (bars) and their selected cutoff values in blue (diamonds). The areas of the blue diamonds are proportional to the number of laboratories using a given cutoff value.

**A**


SA (µmol/L)
13
12
11
10
9
8
7
6
5
4
3
2
1
0
Targeted Range
Cumulative Percentiles
Cutoff Percentiles
HT1 Disease Ranges

**B**

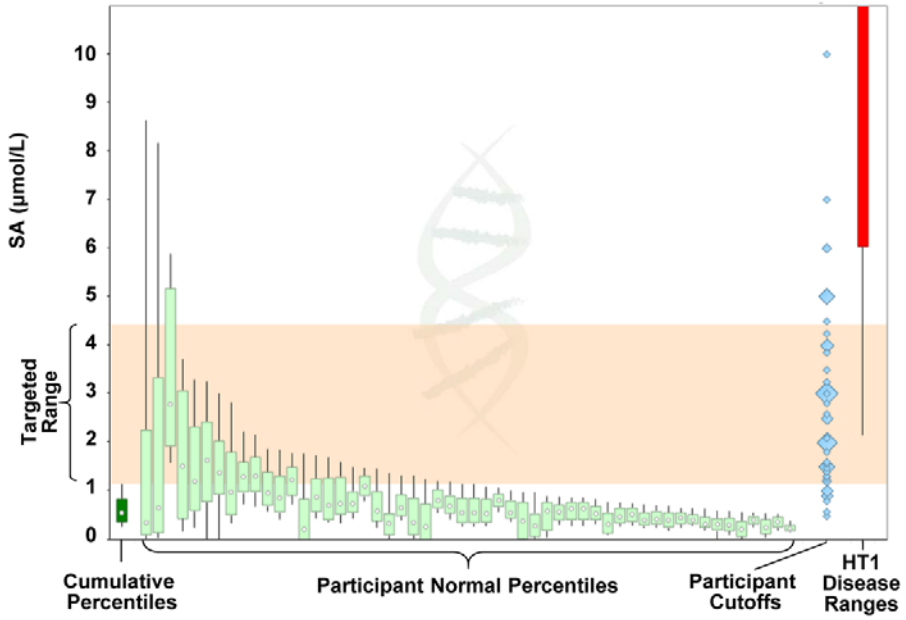

SA (µmol/L)
10
9
8
7
6
5
4
3
2
1
0
Targeted Range
Cumulative Percentiles
Participant Normal Percentiles
Participant Cutoffs
HT1 Disease Ranges

**Table S1**. *GSTZ1* mutations in MHSA individuals.

| **Mutation Designations** | | | | **In silico prediction for coding mutations** | | | **ExAC allele Prevalence** | **Predicted functional category [1]** | **Comments** |
|---|---|---|---|---|---|---|---|---|---|
| **cDNA** | **Protein** | **Genomic** | **db SNP** | **SIFT** | **Polyphen** | **Mutation Taster** | | | |
| c.259C>T | Arg87Ter | 14:77794297 C/T | Not found | NA | NA | Disease causing | 0.00002482 | Pathogenic | Pathogenic |
| c.68-12G>A | NA | 14:77793169 G/A | Not found | NA | NA | NA | 0.00001933 | VUS | Predicted to disrupt normal splicing [2] |
| c.449C>T | Ala150Val | 14:77796125 C/T | rs199552988 | Not Tolerated | Probably damaging | Disease causing | 0.00007415 | VUS | Decreased activity (Figure 2) |
| c.295G>A | Val99Met | 14:77794333 G/A | rs140540096 | Not Tolerated | Probably damaging | Disease causing | 0.0005329* | VUS | Decreased activity (Figure 2) |

The designations of each mutation with respect to mRNA transcript, mature protein and genomic position are shown in the four left columns. The predicted effects of coding point mutations on protein function are shown in the next columns, followed by the prevalence of each mutation in the ExAc database, the predicted functional category of the mutation[1] and additional comments. The reference cDNA sequence for GSTZ1/MAAI was NM_145870.2. The programs used to predict the effects of point mutations are described in Methods. Abbreviations: NA, not applicable; VUS, variant of unknown significance. *, one homozygous individual, described in the text.

## References


1. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, Committee ALQA. 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;17(5):405-24.

2. Human Splice Finder 3.0 (http://www.umd.be/HSF3/HSF.html).