Supplementary Material

The functional readthrough extension of malate dehydrogenase reveals a modification of the genetic code

Julia Hofhuis¹, Fabian Schueren¹, Christopher Nötzel¹, Christopher Nötzel¹, Fabian Schueren¹, Thomas Lingner², Jutta Gärtner¹, Olaf Jahn³, Sven Thoms¹,¹

¹ Department of Pediatrics and Adolescent Medicine, University Medical Center Göttingen, University of Göttingen, 37075 Göttingen, Germany
² Microarray and Deep Sequencing Core Facility, University Medical Center Göttingen, University of Göttingen, 37077 Göttingen, Germany
³ Proteomics Group, Max-Planck-Institute of Experimental Medicine, 37075 Göttingen, Germany
⁴ Current Address: Program in Biochemistry and Structural Biology, Cell and Developmental Biology, and Molecular Biology, Weill Cornell Graduate School of Medical Sciences, New York, NY, United States of America

¹ Correspondence: Email: sven.thoms@med.uni-goettingen.de

Supplementary Figures S1 – S6
Supplementary Tables S1 – S3

Supplementary Figure S1. Genomic distribution of readthrough propensity. Positively scaled and normalized LinIter scores (RTP⁺) of the 31,000 unique SCCs are plotted versus the transcript rank. LinIter: regression coefficients for the iterative model considering all 12 stop context positions.
Supplementary Figure S2. Lack of conservation in RNA structural elements in the readthrough extension of vertebrate MDH1x. (A) mRNA structure prediction was performed with RNAfold Web Server using nucleotides from position -10 up to the second stop codon. MDH1 extensions shown: Human (*Homo sapiens*), Mouse (*Mus musculus*), and Zebra fish (*Danio rerio*). (B) mRNA structure prediction of MDH1-SCCx and MDH1-SCCx with scrambled extension (MDH1-SCCx_scr). (C) Conserved predicted secondary structure based on alignment of the mammalian MDH1 3’UTRs (nucleotides -10 up to +180, see also alignment in Supplementary Fig. S3) using RNAfold Web
Mammalian MDH1 3'UTRs (nucleotides -10 to +180) were aligned using Multiple Sequence Alignment Tool Clustal Omega and the alignment was then used for prediction of consensus secondary structures (RNAfold Web Server). Dot-bracket notation shows the optimal secondary structure with a minimum free energy.
Supplementary Figure S4. Genomic distribution of product score. Product score of scaled and normalized LINiter readthrough propensity scores (RTP\(^+\)) and PTS1 targeting probabilities plotted versus the rank of the transcript for the first 4000 out of 31,000 unique SCCs.
**Supplementary Figure S5.** MDH1x targets to the peroxisome by translational readthrough and a hidden peroxisomal targeting signal type1 (PTS1). (A) Direct immunofluorescence microscopy of transfected HeLa cells. YFP-MDH1_UGG localizes to the cytoplasm and peroxisomes. Removal of cytosol (-CYT) after digitonin treatment allows visualization of the peroxisomal localization of MDH1. (B) Peroxisomal targeting of MDH1 is dependent on the hidden PTS1 CRL in the extension. Deletion of leucine (∆L) in CRL blocks import of MDH1 into peroxisomes. Scale bar: 10 µm.

**Supplementary Figure S6.** Endogenous MDH1 localizes to peroxisomes in glioblastoma cells and murine cardiomyocytes. Immunofluorescence with anti-MDH1 and anti-Pex14 or anti-PMP70 antibodies. (A) Endogenous MDH1 shows mainly cytosolic localization in U118 and U373 cells. Detection of colocalization of MDH1 with the peroxisomal marker PMP70 after removal of cytosol (-CYT). (B) In murine cardiomyocytes MDH1 colocalizes with Pex14 on top of the cytosolic localization of MDH1. Arrows indicate peroxisomal localization of MDH1. Scale bars: 10 µm.
**Supplementary Table S1. Mass spectrometry detection of amino acids encoded by stop codon.**

Intensity (in arbitrary units, a. u.), relative precursor mass deviation (ΔM), fragment count (b-y ions) and fragment mass deviation (RMS error) of peptides detected in our study and table of amino acids identified in other studies that used eukaryotic translation systems.

| PST number | Name | Source |
|------------|------|--------|
| pEYFP-C1   |      | Thermo Scientific |
| pcDNA3.1/Myc-His(-)A | | Thermo Scientific |
| pEXP-N-Venus | | Anja Muntau lab |
| pENTR/D-TOPO | | Thermo Scientific |
| 1226 | pEXP-N-Venus-PTS1-ACOX3 | (Schueren et al., 2014) |
| 1360 | pDRVL | (Schueren et al., 2014) |
| 1385 | pDRVL-LDHB-SCC | (Schueren et al., 2014) |
| 1435 | pDRVL-MDH1-SCC | (Schueren et al., 2014) |
| 1436 | pEYFP-C1-MDH1x | This study |
| 1443 | pcDNA3.1-HA-MDH1x-myc | This study |
| 1473 | pDRVL-MDH1 [TAA CTA] | This study |
| 1474 | pDRVL-MDH1 [TAG CTA] | This study |
| 1475 | pDRVL-MDH1 [TAA TTA] | This study |
| 1476 | pDRVL-MDH1 [TGA TTA] | This study |
| 1479 | pDRVL-MDH1 [TGA CCA] | This study |
| Oligo   | Name                              | Sequence 5'-3'                                                                 |
|---------|-----------------------------------|-------------------------------------------------------------------------------|
| JH77    | DR-MDH1-[TGAATA]_for             | GTCACGTTTCTCTCTGCTGAGCAGACATGATGTTACTAAATGCTTCAGC                             |
| JH78    | DR-MDH1-[TGAATA]_rev             | CCGGACATTGTCTATTCCGACGAGGAACG                                                |
| JH102   | MDH1TGG_\_L_rev                  | GCGCGGATCTCTAAGCAGACATTTTAGATTCTTCAGC                                        |
| JH103   | MDH1TAA_for                      | GTGCTTTTGAATTCTTTCTCTCGCACTAGCAGATGTTACTAAATGCTTCAGC                         |
| JH104   | MDH1TAA_rev                      | GAAGCATTAGTAAACATCTACATGTTAGCGAGGAAAGAAATCTACAGC                             |
| JH105   | DR-MDH1x_for                     | GTACCGTTTCTCTCCTGACTAGACATGTTACTAAATGCTTCAGC                                 |
| JH106   | DR-MDH1x_rev                     | CCGGACAGCAGCAGACATTTTAGATTCTTCAGCAGCAGGAAAGAAATCTACAGC                       |
| JH109   | DR-MDH1TGGx_for                  | GTCACCGTTTCTCTCCTGACTAGACATGTTACTAAATGCTTCAGC                                 |
| JH110   | DR-MDH1TGGx_rev                  | CCGGACAGCAGCAGCAGACATTTTAGATTCTTCAGCAGCAGGAAAGAAATCTACAGC                     |
| JH113   | DR-MDH1x\_\_L_26_for            | GTCACCGTTTCTCTCCTGACTAGACATGTTACTAAATGCTTCAGC                                 |

Supplementary Table S2. DNA constructs used in this study.
| Name       | Oligonucleotide                        | Sequence                                                                 |
|------------|----------------------------------------|--------------------------------------------------------------------------|
| JH114      | DR-MDH1xΔ26_rev                        | CCGGATGAAGCATTATAGTAAACATCTATTGGTCTAGTCAGGCAAGAGGACG                      |
| JH137      | DR-MDH1x-scr_for                       | GCCGACTGCCGTAGTAATTACTTGACTATAAA2AAGTACCTTATAATATATTATCAGGACG            |
| JH138      | DR-MDH1x-scr_rev                       | CCGGACTGCCGTAGTAATTACTTGACTATAAA2AAGTACCTTATAATATATTATCAGGACG            |
| JH139      | DR-MDH1TGGx-scr_for                    | GTCACCGTTCCTCCTTGTTAGCTAGCTACTGACTCTAGGCAACATTTAGGCGACGACATTG            |
| JH140      | DR-MDH1TGGx-scr_rev                    | CCGGACTGCCGTAGTAATTACTTGACTATAAA2AAGTACCTTATAATATATTATCAGGACG            |
| OST1192    | EcoRI-MDH1x_for                        | GCAGGATCTCTTTGCTCTGGAACAAATTGAGAGACG                                    |
| OST1193    | MDH1x-BamHI_rev                        | GCCGCGATCTCTCAAAGACGACATTGAGACG                                        |
| OST1204    | Nhel-HA tag-MDH1                        | GCCGCTAGTACGATGCACTCCAGATTACGACTCGTCTGAACCAATTTAGGTTCTC                  |
| OST1205    | BamHI-MDH1x-wo_second_STOP             | GCCGCGATCTCTCAAAGACGACATTGAGACG                                        |
| OST1231    | MDH1TGG-mut_for                        | GCTTTTGAATTCTTCTCTGCTCTGCTAGACATTGATGTTAC                               |
| OST1232    | MDH1TGG-mut_rev                        | TCAAGAGACGACATTGAGACATTGAGACG                                          |
| OST1245    | MDH1-human-PTS1_for                    | CACCTTCAGGCTAGACATTGAGACATTGAGACG                                        |
| OST1246    | MDH1-human-PTS1_rev                    | TCAAGAGACGACATTGAGACATTGAGACG                                          |
| OST1247    | MDH1-zebra finch-PTS1_for              | CACCTTCAGGCTAGACATTGAGACATTGAGACG                                        |
| OST1248    | MDH1-zebra finch-PTS1_rev              | TCAAGAGACGACATTGAGACATTGAGACG                                          |
| OST1463    | DR-MDH1-[TAA CTA]_for                  | GTCACCGTTCTCTGCCTAATGACGATATG                                          |
| OST1464    | DR-MDH1-[TAA CTA]_rev                  | CCGGACATTTGCTCTTTGCTCTACGTCAGGAGAAGCG                                  |
| OST1465    | DR-MDH1-[TAG CTA]_for                  | GTCACCGTTCTCTGCCTACGTCAGGAGAAGCG                                      |
| OST1466    | DR-MDH1-[TAG CTA]_rev                  | CCGGACATTTGCTCTTTGCTCTACGTCAGGAGAAGCG                                 |
| OST1467    | DR-MDH1-[TAA TTA]_for                  | GTCACCGTTCTCTCTGCTCATTGACGTCAGGAGAAGCG                                 |
| OST1468    | DR-MDH1-[TAA TTA]_rev                  | CCGGACATTTGCTCTTTGCTCTACGTCAGGAGAAGCG                                 |
| OST1469    | DR-MDH1-[TGA TTA]_for                  | GTCACCGTTCTCTTCTGCTCATTGACGTCAGGAGAAGCG                                 |
|OST1470     | DR-MDH1-[TGA TTA]_rev                  | CCGGACATTTGCTCTTTGCTCTACGTCAGGAGAAGCG                                 |
|OST1475     | DR-MDH1-[TGA CCA]_for                  | GTCACCGTTCTCTTCTGCTCATTGACGTCAGGAGAAGCG                                 |
|OST1476     | DR-MDH1-[TGA CCA]_rev                  | CCGGACATTTGCTCTTTGCTCTACGTCAGGAGAAGCG                                 |
|OST1477     | DR-MDH1-[TGA CTT]_for                  | GTCACCGTTCTCTTCTGCTCATTGACGTCAGGAGAAGCG                                 |
|OST1478     | DR-MDH1-[TGA CTT]_rev                  | CCGGACATTTGCTCTTTGCTCTACGTCAGGAGAAGCG                                 |
|OST1479     | DR-MDH1-[TGA CTG]_for                  | GTCACCGTTCTCTTCTGCTCATTGACGTCAGGAGAAGCG                                 |
|OST1480     | DR-MDH1-[TGA CTG]_rev                  | CCGGACATTTGCTCTTTGCTCTACGTCAGGAGAAGCG                                 |

**Supplementary Table S3.** Oligonucleotides used in this study. DR, dual reporter.