RUHR UNIVERSITÄT BOCHUM

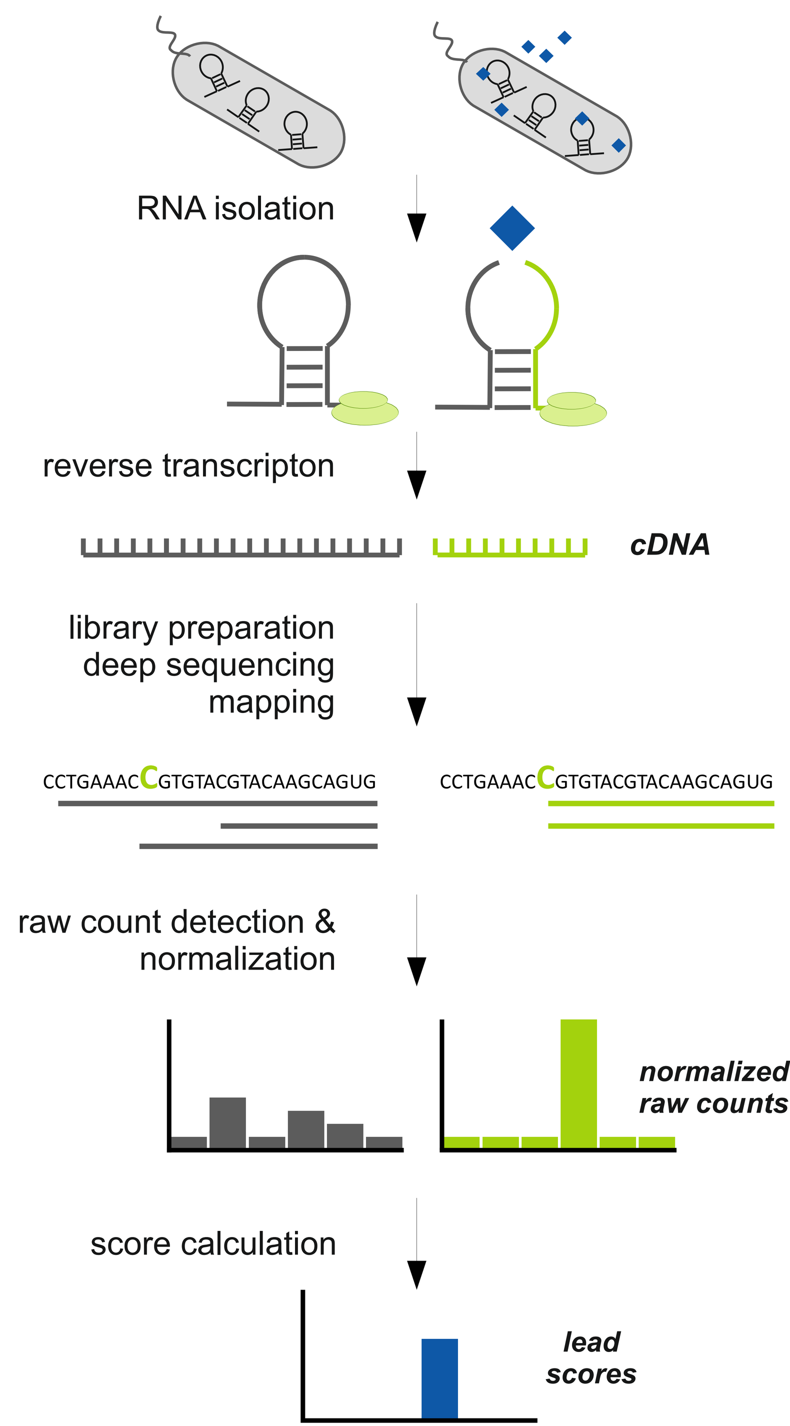
RUB

# Lead-seq: *in vivo* RNA structure probing on a transcriptome-wide scale

Vivian B Brandenburg[a†], Christian Twittenhoff[a†], Francesco Righetti[a†], Aaron M Nuss[b], Axel Mosig[c], Petra Dersch[b,d], Franz Narberhaus[a1]

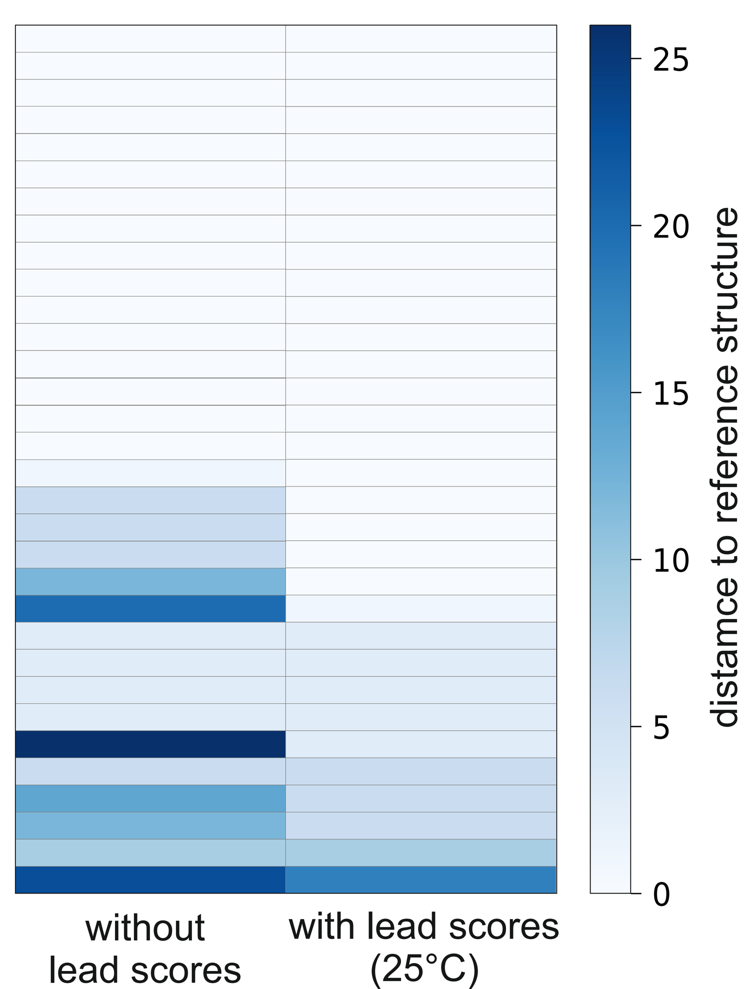
a Microbial Biology, Ruhr University Bochum | b Molecular Infection Biology, Helmholtz Centre for Infection Research | c Biophysics, Ruhr University Bochum | d Infectiology, Center for Molecular Biology of Inflammation, University of Münster | † authors contributet equally to this work

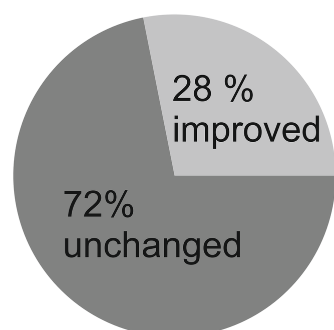
## overview




**Lead-seq provides information about single-stranded nt in the transcriptome of *Yersinia pseudotuberculosis* at 25°C and 37°C.** For each temperature, two bacterial cultures were grown. One of them was treated with lead(II)-acetate, which induced strand breaks in single stranded regions of the RNA. From both samples, RNA was isolated. The reverse transcription stopped at strand breaks. After library preparation and deep sequencing, the resulting reads were mapped to the transcriptome. The number of 5'-ends of reads mapping to each nt (raw counts) was identified and normalized. These counts either refer to both, single stranded nt and spontaneous Reverse Transcriptase drop-offs (lead-treated sample) or exclusively to the latter (control sample). In a last step, the counts are combined into lead scores.

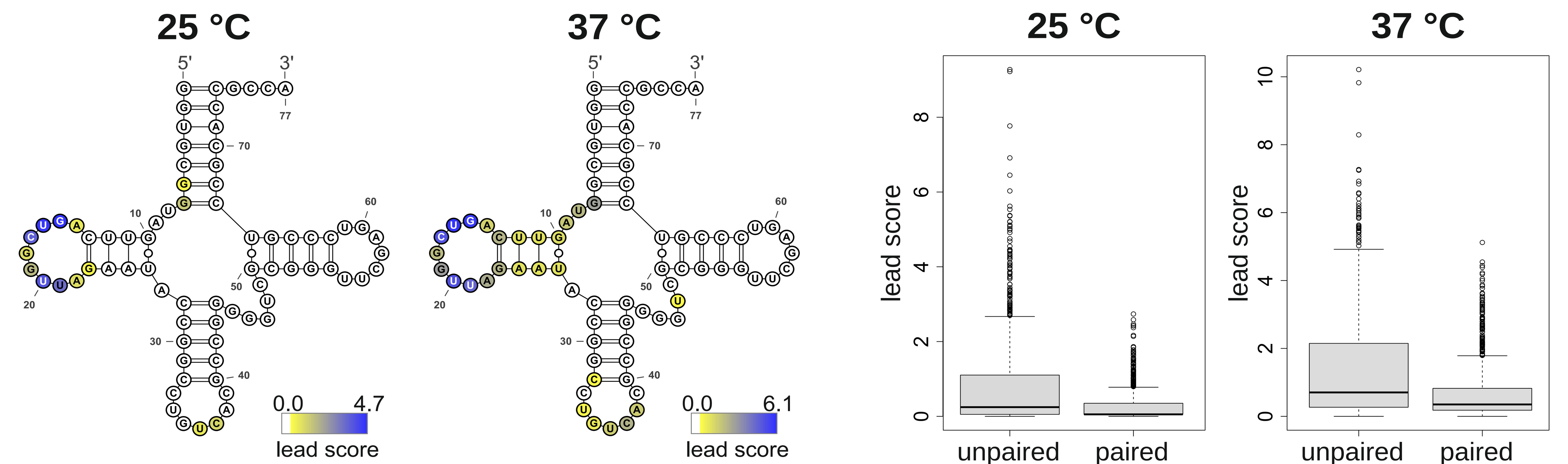
## structure prediction



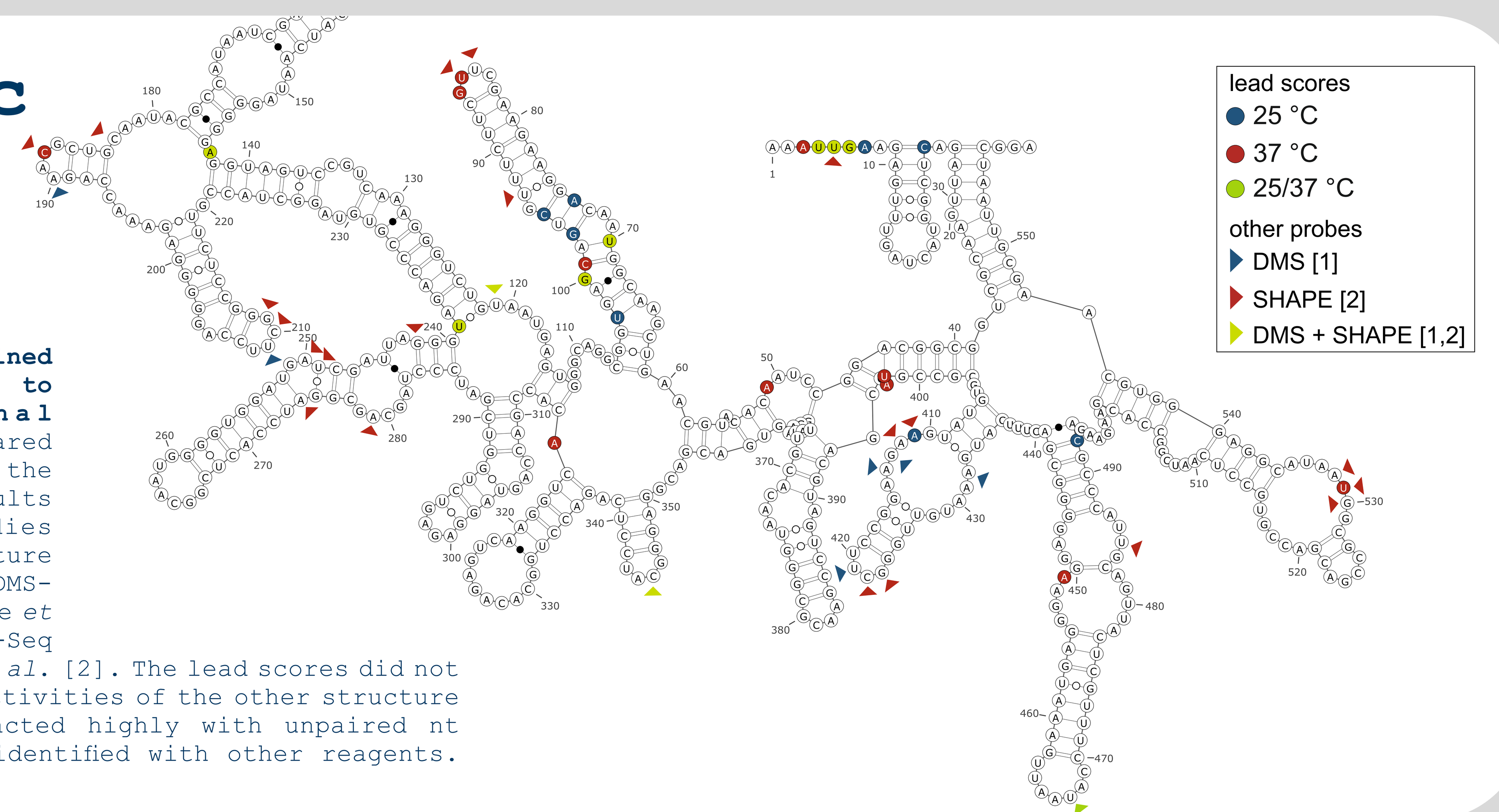

**Lead scores can improve RNA secondary structure prediction.** The distance between the structures predicted by RNAfold [4] and the reference structures of 32 tRNAs was calculated. Subsequently, lead scores deriving from experiments at 25°C were included into the prediction with the method from Washietl *et al.* [5]. Particularly the tRNAs with poorer prediction were improved by incorporation of experimental data. In total, the addition of lead scores improved the secondary structure prediction of 28% of the 32 tRNAs.

## lead scores of tRNAs




**Lead scores are informative about the pairing state of nt at 25°C and 37°C.** The examination of single tRNAs as the $tRNA^{His}$ (left) revealed that high lead scores map to single stranded nt. We also investigated the distribution of lead scores in paired and unpaired nt of 32 tRNAs (right). They differ significantly at both temperatures (unpaired two-sided t-test 25°C/37°C: $p < 2.2e{-16}$).

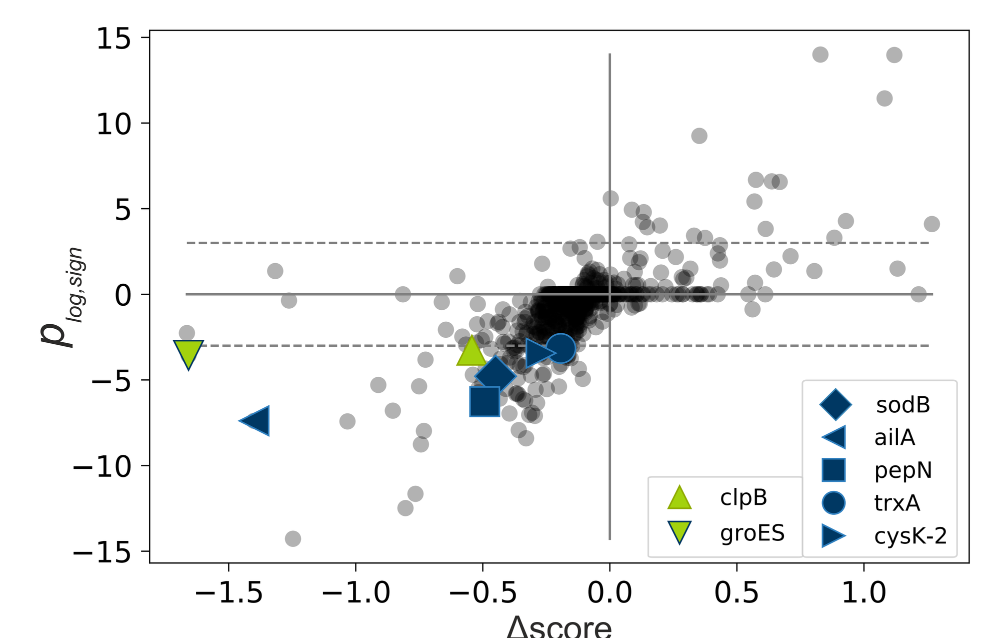
## methodic synergy




**Lead-seq can be combined with other methods to create additional versatility.** We compared the lead scores from the 16S rRNA with the results from previous studies using different structure probing agents: a DMS-based study from Mustoe *et al.* [1], and a SHAPE-Seq study from McGinnis *et al.* [2]. The lead scores did not only overlap with reactivities of the other structure probes, but also reacted highly with unpaired nt different from those identified with other reagents.

## RNA thermometer identification

| gene | gene description | Δscore | $p_{log,sign}$ |
|---|---|---|---|
| ***ailA*** | outer membrane protein | -1.4 | -7.4 |
| ***pepN*** | aminopeptidase | -0.5 | -6.3 |
| ***sodB*** | superoxide dismutase | -0.5 | -4.8 |
| ***cysK-2*** | cysteine synthetase | -0.3 | -3.4 |
| ***trxA*** | thioredoxin | -0.2 | -3.2 |
| ***groES*** | chaperonin | -1.7 | -3.6 |
| ***clpB*** | ATP-dependent chaperone | -0.5 | -3.3 |




**Lead-seq can be used to detect thermosensing RNA structures.** We investigated the temperature-induced change of lead scores from 25°C to 37°C (Δscores) around the Shine-Dalgarno region. We also compared the temperature-dependent behavior of this region with the behavior of the rest of the transcript ($p_{log,sign}$). RNA thermometers unfold around the Shine-Dalgarno region in general (negative Δscore) and are significantly more temperature-responsive in this region than the rest of the transcript ($p_{log,sign} \leq -3$). With this method, we found RNA thermometers which were previously verified by Righetti *et al.* [3] (marked in blue), as well as new RNA thermometers (marked in green), subsequently verified through reporter gene assays.

## summary

**Lead-seq probes RNA structures globally.**
**Lead reactivity identifies unpaired nt in small and large RNAs.**
**Lead-seq combined with other structure probes can give high-resolution insights on structuromes.**
**Lead-seq is a powerful method for identifying RNA thermometers.**
**Lead scores can improve RNA secondary structure prediction.**

DFG Deutsche Forschungsgemeinschaft

This work was funded by the German Research Foundation (DFG NA 240/10-2).

## references

[1] Mustoe *et al.*, 2019: RNA base-pairing complexity in living cells visualized by correlated chemical probing. *PNAS* 116, 24574–82.
[2] McGinnis *et al.*, 2015: In-cell SHAPE reveals that free 30S ribosome subunits. *PNAS* 112, 2425–30.
[3] Righetti *et al.*, 2016: Temperature-responsive *in vitro* RNA structurome of *Yersinia pseudotuberculosis*. PNAS 113, 7237–42
[4] Lorenz *et al.*, 2011: ViennaRNA Package 2.0. *Algorithms Mol Biol*, 6, 26.
[5] Washietl *et al.*, 2012: RNA folding with soft constraints: reconciliation of probing data and thermodynamic secondary structure prediction. *NAR* 40, 4261–72.