nanoCAGE-XL FAQs

Q1: Do you recommend to try the protocol from exp. 2 or 3, and for what reasons?

A1: Experiment 2 tested the presence of a common linker sequence for each sample (in combination with barcodes (as in Tang et al., Nucleic Acids Res. **2013** Feb 1;41(3):e44. PubMed: 23180801), which reduced individual sample biases as well as the total number of artefactual reads by using non-G bases immediately upstream of the 'GGG' in the 3' end of the template switching primer. Experiment 3 tested Illumina-recommended barcodes combinations, which improved sequence diversity at the start of reads and thus sequencing. We currently use barcoding as per Exp. 3 plus the common linker as per Exp. 2.

Q2: What was the input of total RNA for each experiment? What is the minimum input for total RNA?

A2: It is really important to remove as much rRNA as possible. For Arabidopsis roots, we start with 5 ug total RNA which we subject to one round purification with RiboMinus Plant kit (Invitrogen) followed by one round purification with RiboZero Root Seed kit (Illumina). For Arabidopsis leaf, we start with 10 ug total RNA and run 2 rounds with RiboMinus Plant kit (Invitrogen) and a round with RiboZero Leaf kit (Illumina).

Q3: How much RNA was used to start the nanoCAGE-XL protocol after the rRNA depletion?

A3: We use ~50 ng rRNA-depleted RNA per a 10 ul RT reaction as in the paper. In our latest runs, we have used up to 30 ul reaction volumes with ~150 ng rRNA-depleted RNA. This comes handy as this is usually the yield from 5 ug root total RNA and 10 ul leaf total RNA. One could obtain more reads with 40 ul reactions, but this makes it very expensive at the RNA purification step. We haven't tried RT reaction volumes larger than 40 ul.

Q4: Did you try to add some phiX control library in order to improve the low alignment rate observed in exp. 2?

A4: We decided against spiking in phiX for the purpose of improving sequence diversity. The nanoCAGE libraries are sequenced with a custom primer, and phiX is sequenced with the Illumina standard sequencing primer. Because at present Illumina does not disclose the exact sequence of their Seq. primer, we decided against mixing an unknown primer with our sequencing primer. We are currently using Illumina-recommended barcode combinations plus the common linker, which improves the read alignment.

Q5: Did you consider adding N bases instead of GCTATA in exp. 2?

A5: We were considering using N bases for a while, but decided on the GCTATA linker because of its low 'G' content immediately adjacent to the 'GGG' tail of the RT oligo, which reduces the frequency of artificial annealing of the oligo and premature RT template switching.

Q6: Why did you use 50 ng mRNA + 100 ng total RNA in exp. 3? Is this a strategy you recommend?

A6: We tested the consequence of adding a defined amount of total RNA contamination to the sample, as opposed to using only rRNA-depleted sample. Based on our results, we recommend removal of as much rRNA as possible in order to minimize the number of (useless) rRNA reads.

Q7: Why did you change the PCR enzyme from *Ex* Taq to Phusion for the second PCR?

A7: We changed to Phusion in the second PCR in order to generate blunt-end cDNAs - we wanted blunt ends because we introduced an Exo I digestion step to remove single-stranded

primer contamination. We highly recommend removing the primer contamination before sequencing.

Q8: Why did you change the Reverse Library PCR Primer sequence from experiment 1 to a different sequence in experiments 2 and 3?

A8: We changed to a different primer sequence for the barcoded libraries in experiments 2 and 3 based on personal communication with Dr. C. Plessy (RIKEN Center for Life Science Technologies).