

## Digital Gene Expression - Tag Profiling FAQ

Q: What are the sizes of the reads/tags?

A: For *Nla*III tag profiling you read 17 bp but analyze using a 21 bp tag. For *Dpn*II tag profiling you read 16 bp but analyze using 20 bp tag. The extra bases come from knowing the binding site sequence for *Nla*III or *Dpn*II. For example *Nla*III is a 17 bp read plus your knowledge of the 4 bases that constitute the restriction site = 21 bp tag.

Q: What programs do you use to analyze the data from a Digital Gene Expression – Tag Profiling run?

A: We will be releasing a short description on how to analyze Digital Gene Expression - Tag Profiling data with kit release. Essentially we use the standard pipeline to determine sequence and then count the number of times each unique sequence appears. We also offer a complete set of annotation for all the human and mouse tags generated with this method.

Q: Would the PCR step introduce RNA subpopulation biases?

A: For DGE - Tag Profiling we are amplifying an 85 bp fragment for only 15 cycles. In addition only 16-17 bp vary in the amplicon. With these factors we feel we are not introducing any PCR bias in any transcript levels with PCR

Q: What are the chances that mutations can be introduced by the PCR step?

A: We do very limited amplification over very short amplicons to limit this.

Q: What is the recommended depth of coverage?

A: You do not really recommend depth of coverage with DGE - Tag Profiling. We usually recover 4 million tags per flow cell lane. At that depth of coverage a transcript that is present in the cell only once will appear approximately 12 times. This should be enough to call tags as not present with confidence. If the end user wants to further push the sensitivity they can run two lanes on a flow cell.

Q: Can this technology be used to analyze bacterial or viral transcripts?

A: The current Digital Gene Expression – Tag Profiling sample prep manuals assume eukaryotic total RNA is the starting material. We take advantage of the poly A tail in the first step of the protocol. .

Q: Can this technology distinguish different transcript isoforms?

A: Unless the isoform difference generates two distinct tags we cannot distinguish splice variants with these protocols.

Q: Do you have any publications using this technology for gene expression analysis?

A: Not at the time of release

Q: What if my mRNA does not contain any (*Nla*III or *Dpn*II) restriction enzyme sites?

A: In this rare case you will not see a tag for your transcript of interest. We expect a *Nla*II cut site in 95% of all human genes. If your gene does not show up try *Dpn*II method. Between the two we should hit roughly 99% of genes. We suggest the end user use *Nla*III as the first pass, and *Dpn*II method is our back up. We do not support using both enzymes in single digest. Protocol is such that this will not work. *Nla*III and *Dpn*II methods must also be run if different channels.

Q: How much total RNA input is required?

A: For Digital Gene Expression – Tag Profiling we suggest 1-2 ug.

Q: Can we sequence random primed cDNA libraries? Do you have such a protocol?

A: At this point no.

Q: Are the adapters for gene expression the same as the sequencing adapters?

A: The sequencing adapters for gene expression and small RNA are not identical to each other or to the oligos from genomic DNA sample prep. Also note that they are added sequentially so we can control the orientation and position of the adapter. For genomic DNA sample prep, we add the adapters in one step and need them to be specially modified to get them in the right orientation and position.

Q: What is the accuracy of the digital gene expression measurements? How is accuracy measured? How comparable is the data to qRT-PCR such as Taqman?

A: We are measuring the actual number of times the sequence is present. With a pool of 4 million we are very confident in the counts of low and high expressors and can say with confidence the transcript is not present. We have MAQC data showing Tag Profiling giving results closer to those of TaqMan than any array on the market. .

Q: Do we need to run a control DNA lane?

A: At this point, all DGE SBS runs need a lane of control DNA be it BAC or PhiX. This is mandatory to generate the sequencing matrix.

Q: Can we analyze partially degraded transcripts (ie from FFPE samples) using this technology?

A: For FFPE we recommend Illumina's DASL product line

Q: Is there a special protocol/recommendation for using smaller total RNA input? What is the lower limit? Can this technology be used to profile samples from laser dissection?

A: At this point no. We do know of some end users who are using less starting total RNA but we are not offering a supported method at this time.

Q: Can I analyze transcripts from organisms for which the genome has not been fully assembled/sequenced?

A: You can absolutely get Digital Gene Expression - Tag Profiling data from organisms that are not sequenced. This will just make annotation more difficult but you get the data. Comparisons to similar databases can help identify annotation questions

Q: Do you have a tool for mapping the tags to the reference sequence?

A: We provide clear instructions for collecting the sequencing data and then mapping to genome browser data bases.

Q: If I were to run total RNA from different treatments, how would I normalize the results?

A: You do not need to normalize the data from Digital Gene Expression - Tag Profiling the way you would for microarrays.

Q: Do you provide technical support for gene expression data analysis?

A: We train customers during pipeline training

Q: MmeI cuts with a 2 bp overhang at 18 and 20 bp away from binding site. How do we deal with that?

A: Our GEX Adapter 2 is not just a simple double stranded DNA oligo pair. It is actually a mixture of many double stranded DNA oligo pairs covering all the potential bases in the overhang.

Q: Does the end user need to provide any reagents for tag profiling?

A: Yes. The kit covers 35 items required but we do require the end user purchase an Invitrogen Superscript II kit and supply typical lab elements like PAGE gels and buffers, phenol, chloroform, isoamyl alcohol, and ethanol.