

A vibrant field of sunflowers in full bloom, with a large sunflower in the foreground on the right. The background shows a green field and a blue sky with scattered white clouds.

*See why.*

**Massive Analysis of  
cDNA Ends (MACE)**

GenXPro

## Gene Expression and SNP detection in high throughput and high resolution

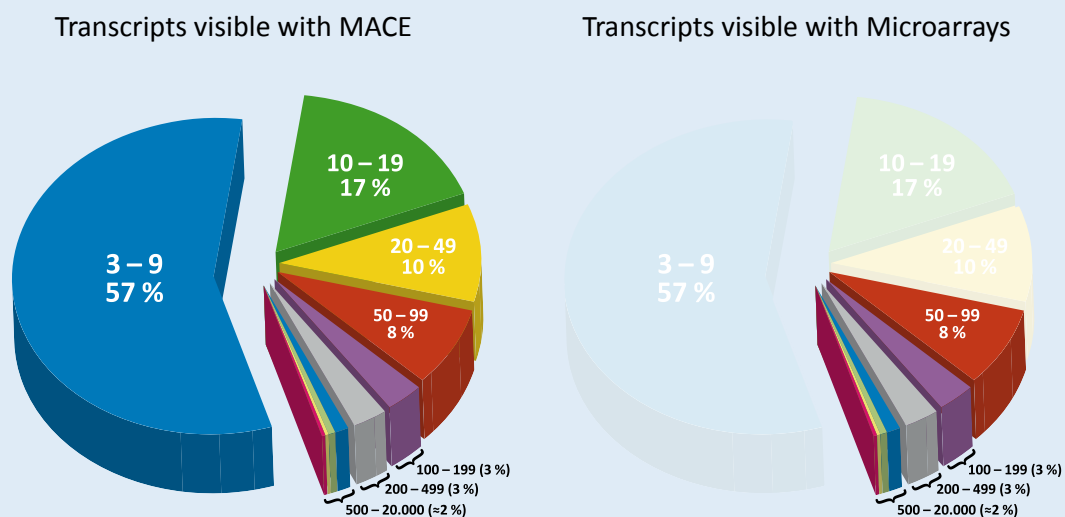
*MACE (Massive Analysis of cDNA Ends) is the ideal deep sequencing method for high-resolution gene expression analysis of any biological material. Deep coverage, excellent quantification and highly reliable SNP detection at low cost are the hallmarks that distinguish this transcription profiling technology from conventional approaches such as RNA-seq.*

### Some transcripts in many copies, many transcripts in few copies

In a typical transcriptome, only a handful of transcript-species can make up 80 % of all transcripts. On the other hand, the vast majority of transcripts are only present in 1 – 20 copies (Fig. 1). Among them are receptors or transcription factors with crucial functions. To adequately identify rarely and moderately expressed transcripts by sequencing, one has to analyze at least 10 million transcripts.

In MACE, each transcript is represented by only one single, highly specific cDNA fragment originating from the 3' end of each transcript-molecule (Fig. 2). In consequence, hundreds of millions of transcript molecules are analyzed simultaneously by next generation sequencing, providing the required resolution even after multiplexing. In RNA Seq, each transcript is represented by 10 – 30 fragments, and hence 10 – 30 x more sequencing is required to obtain a similar resolution.

#### Comparison between MACE- and microarray-data



*Fig. 1: Pie chart of the distribution of transcript species – abundancies: 57 % of the transcript species are present in only 3 – 9 copies/million. The medium and low-level transcripts which make up about 70 % of the transcript-species remain invisible on microarrays.*

Consequently, MACE allows to analyze the mRNA transcriptome at unprecedented low cost, depth and accuracy, which neither can be reached by microarrays nor regular RNA-Seq. Since microarrays measure transcripts via semi-quantitative light-signal intensities, rare transcripts are obscured in the microarray's background signal noise. MACE, however, counts the transcripts and therefore truly quantifies gene expression.

### “TrueQuant“: PCR-bias free data

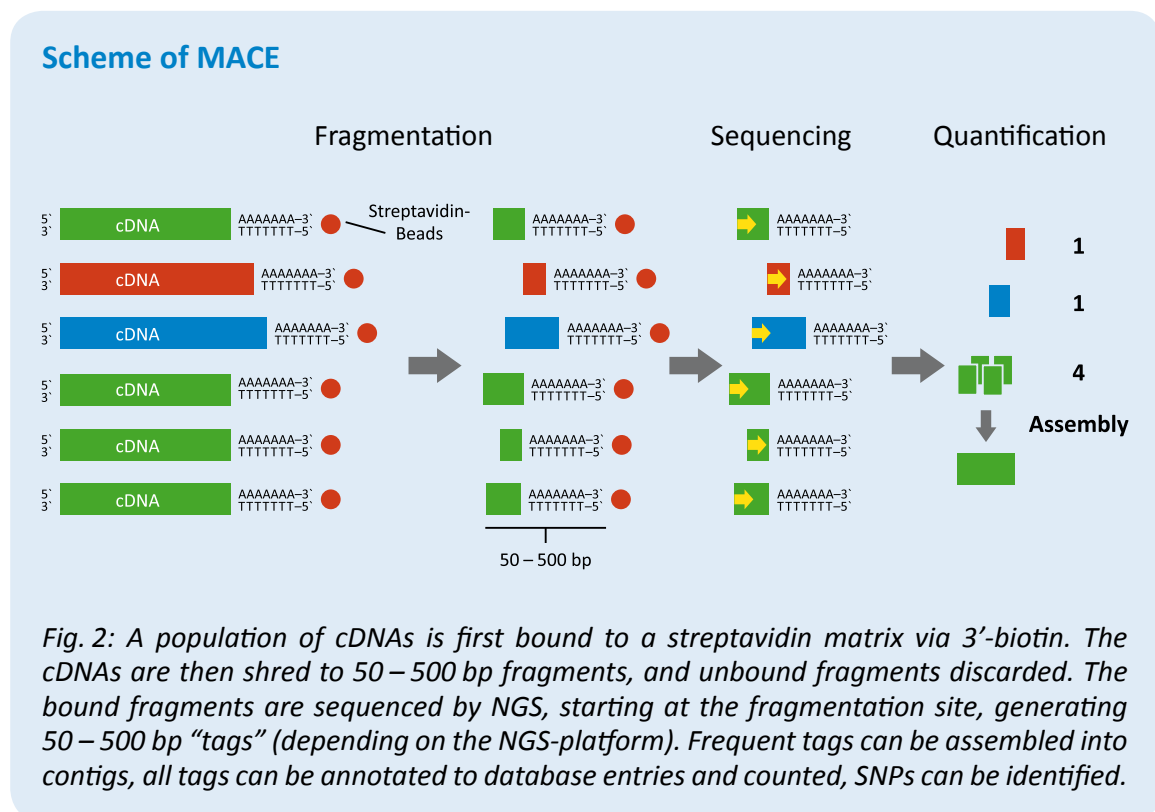
All second-generation sequencing-based data is prone to PCR-bias, because the different DNA fragments are amplified with different efficiency. GenXPro has developed a method to eliminate this bias using its unique “TrueQuant” technique and can therefore – as the single provider world wide – offer PCR-bias free sequencing data (see fig. 3).

### TranSNiPtomics using MACE

The analyzed 3'ends of transcripts are mostly consisting of 3' untranslated regions (3'-UTRs) which are under lower evolutionary pressure. Therefore, these regions contain many sequence polymorphisms such as SNPs or Indels (see fig. 4). Since these are located in genes and not somewhere in the genome they represent highly valuable genomic markers, often directly linked to specific traits.

Since MACE focuses only on the 3' ends of each transcript, the required coverage to define an SNP, even in low-abundant transcripts is warranted!

Figure 5 shows an excerpt of MACE-data from a laboratory rat. Even in very closely related laboratory rats, SNPs are detected by MACE.



*Fig. 2: A population of cDNAs is first bound to a streptavidin matrix via 3'-biotin. The cDNAs are then shredded to 50–500 bp fragments, and unbound fragments discarded. The bound fragments are sequenced by NGS, starting at the fragmentation site, generating 50–500 bp “tags” (depending on the NGS-platform). Frequent tags can be assembled into contigs, all tags can be annotated to database entries and counted, SNPs can be identified.*

### Robustness: partly degraded material

As only the 3' end of each cDNA molecule is sequenced, even partly degraded material such as from formalin-fixed paraffin-embedded (FFPE) specimen can be reliably analyzed.

### High throughput SNP – and gene expression analysis

We can analyze 96 samples simultaneously for only a fraction of the cost for RNA seq. Therefore forget the financial restraints of RNAseq for large sample amounts!

### Bioinformatics, data output

Our service includes annotation, assembly and GO-enrichment analysis. Our tools for data analysis allow for a very easy data handling – no bioinformatics know-how is required!

### Your advantages using MACE

#### High-resolution

- low-level transcripts as e.g. of transcription factors and receptors are exactly quantified

#### TrueQuant technology

- no PCR-introduced bias

#### High SNP-coverage

- highly polymorphic 3' ends are sequenced
- high coverage for reliable SNP detection and genotyping is warranted

#### High throughput & low costs

- we can analyze hundreds of samples simultaneously

### Negative common logarithm of the p-value for differential gene-expression

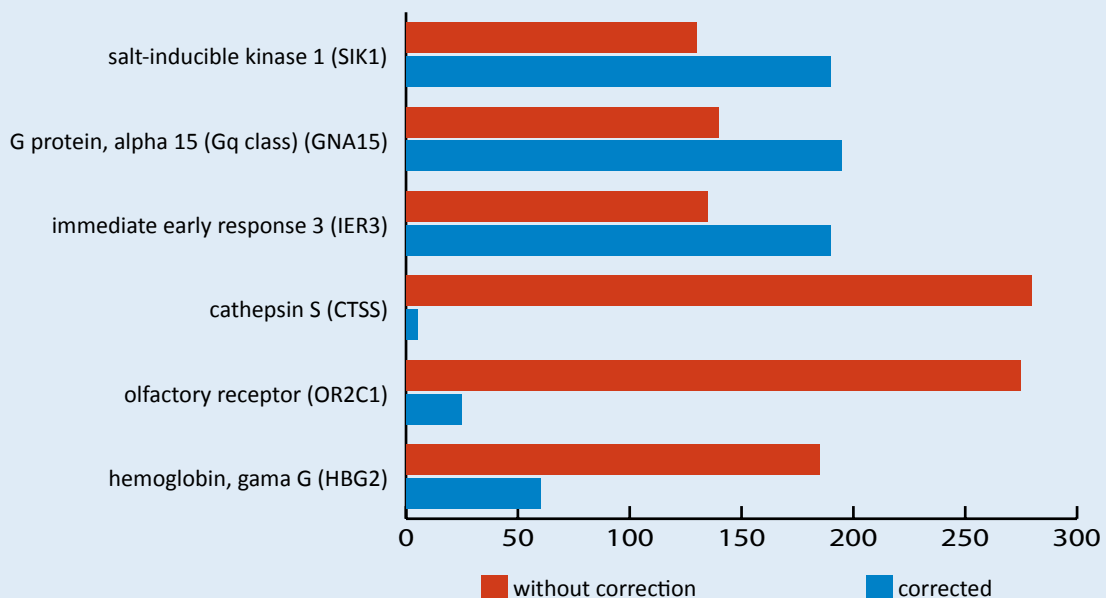


Fig 3: If not corrected for PCR bias (TrueQuant, blue), the p-value is dramatically different as a consequence of biased data (red).



### TranSNiPtomics

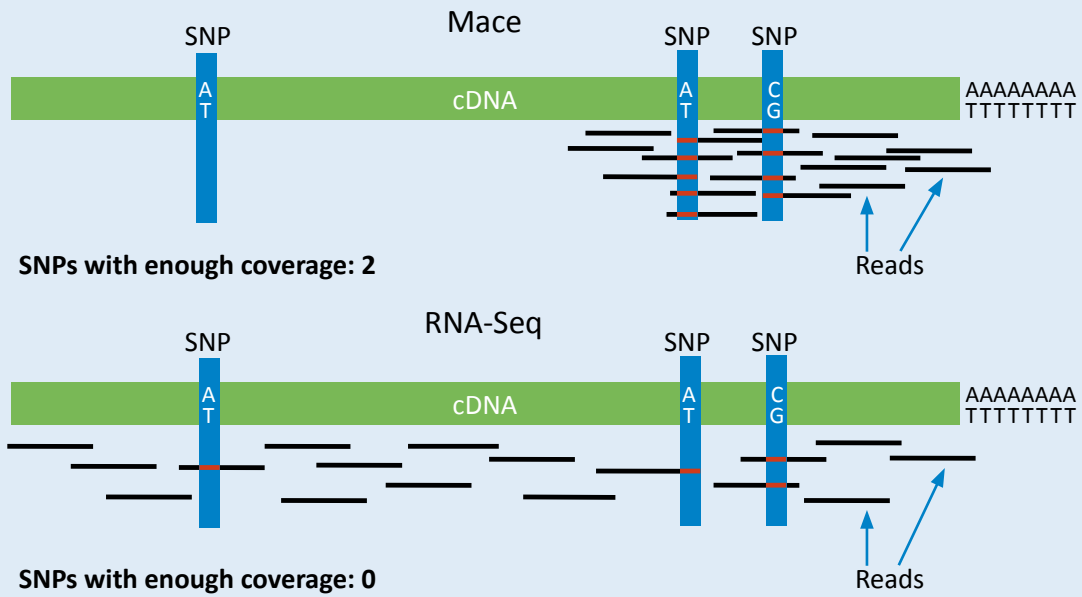


Fig. 4: Concentration on the SNP-rich 3' UTRs provides the required coverage for SNP detection. RNAseq at similar sequencing depth will not reveal the SNPs.

### *Rattus norvegicus*, laboratory strain, $\alpha$ actin MACE tags

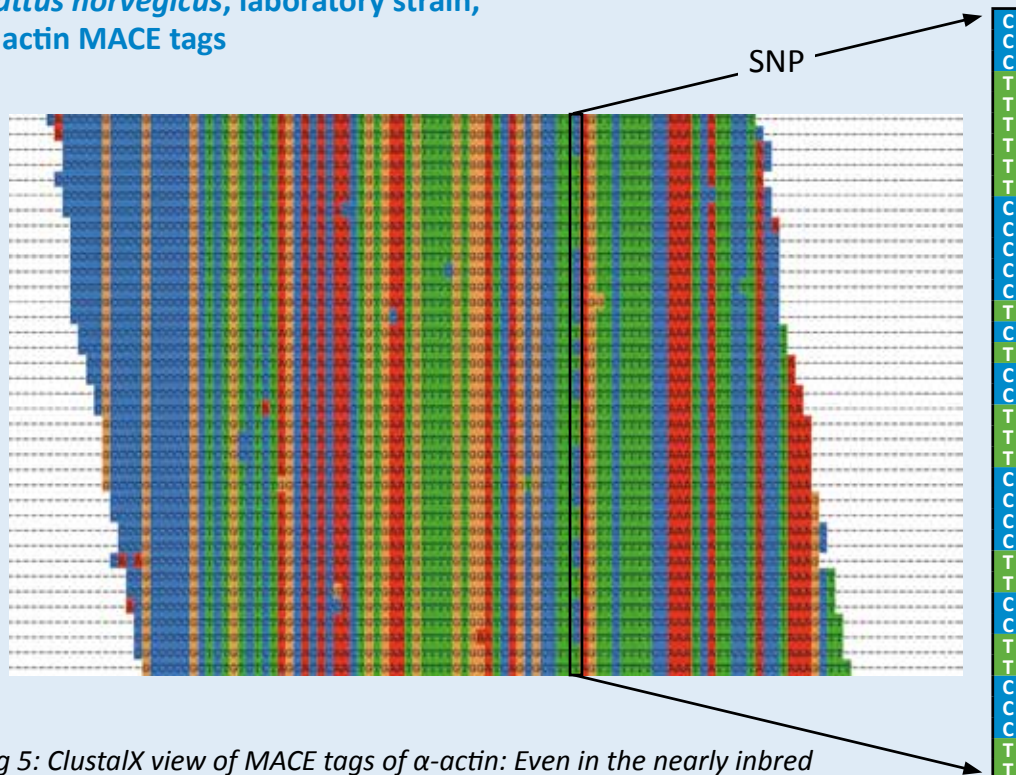


Fig 5: ClustalX view of MACE tags of  $\alpha$ -actin: Even in the nearly inbred laboratory rat, the different alleles of this highly conserved transcript are visible.



# GenXPro

*Genes never seen before...*

**Transcriptomics**

**Genomics**

**Epigenomics**

**Bioinformatics**

**GenXPro GmbH**  
**Altenhöferallee 3, 60438 Frankfurt am Main, Germany**  
**Phone: +49 (0)69-95 73 97 10**  
**[www.genxpro.de](http://www.genxpro.de) | [info@genxpro.de](mailto:info@genxpro.de)**