

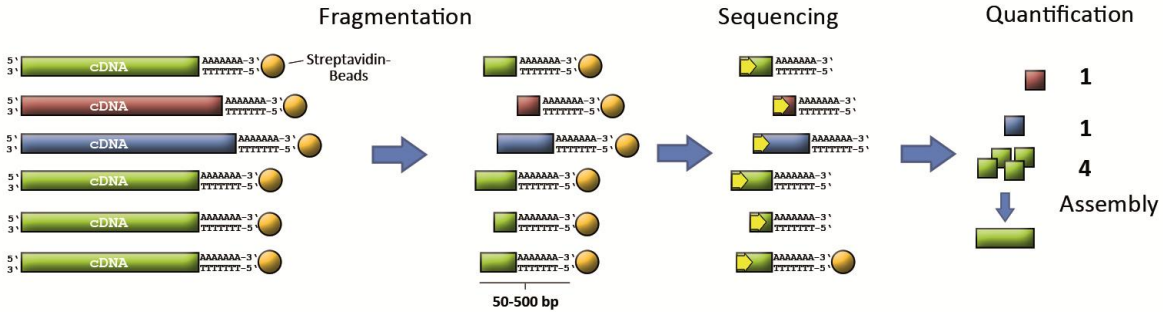


## **Massive Analysis of cDNA Ends (MACE)**

### **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 costs are the hallmarks that distinguish this transcription profiling technology from conventional approaches such as RNA-seq.

**Scheme of MACE:**



**Description:** 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.

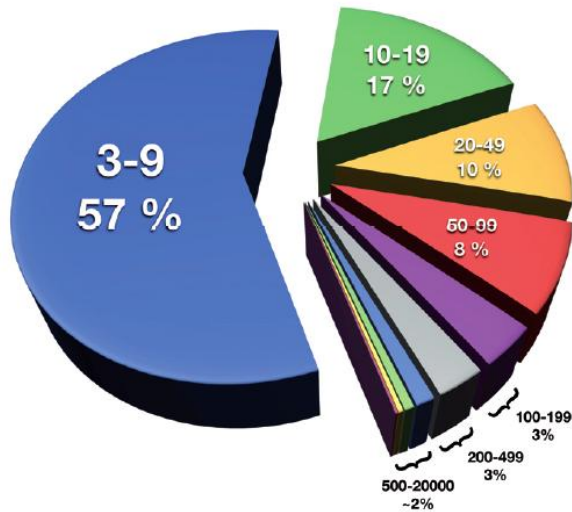
*Some transcripts in many copies, many transcripts in few copies.*

In a typical transcriptome of a cell or tissue, only a handful of transcript-species-copies can make up 80% of all transcripts. On the other hand, the vast majority of transcripts are only present in 1-20 copies. Among them, receptors or transcription factors with crucial functions. In order to adequately identify rare and medium expressed transcripts by sequencing, it is recommended to analyze at least 10 Mio transcripts.

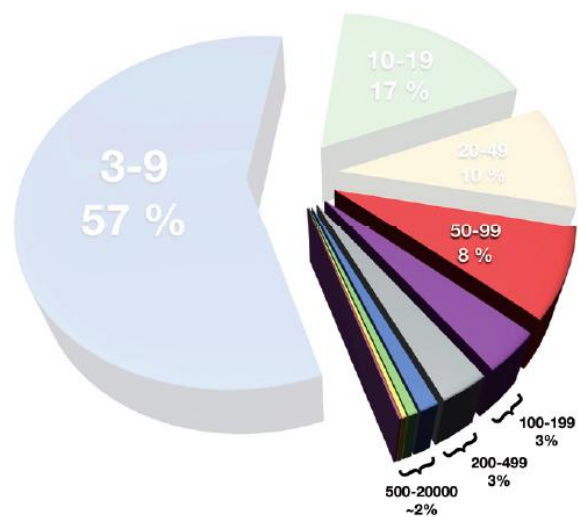
In MACE, each transcript is represented by only one single, highly specific cDNA fragment. In consequence, hundreds of millions of transcript-molecules are analyzed simultaneously on a single Illumina HiSeq sequencing lane, providing the required resolution even after multiplexing up to 12 samples. In RNA Seq, each transcript is represented by 10-30 fragments, and hence 10-30 x more sequencing is required to obtain a similar resolution.

Consequently, MACE finally allows analyzing the mRNA transcriptome at an unprecedented low costs, depth and accuracy, which cannot at all 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. MACE, however, counts the transcripts and therefore truly quantifies gene expression.

**Transcripts visible with MACE**



**Transcripts visible with Microarrays**

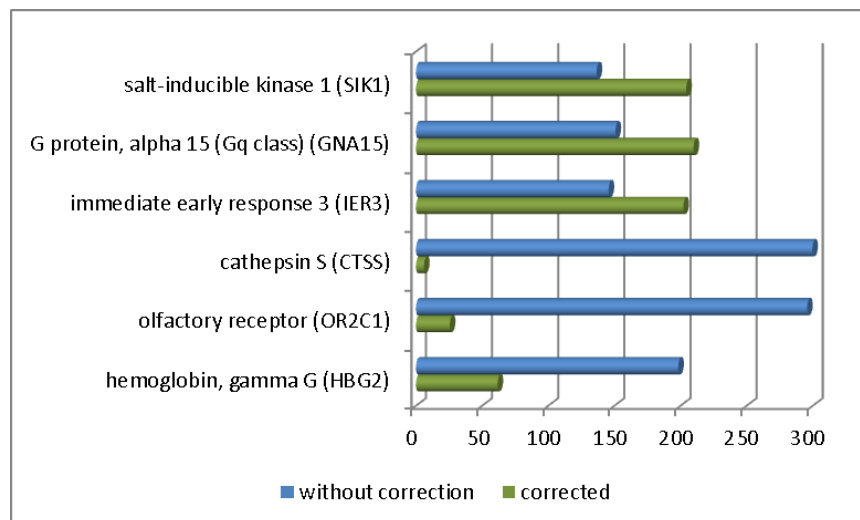


**Comparison between MACE- und Micorarray- Data**

About 70% of the transcriptome consisting of medium and low- level expressed transcripts remains invisible on microarrays.

**“TrueQuant“: PCR-Bias Free Data!**

All second- generation sequencing-based data are prone to PCR-bias, because the different DNA fragments are amplified with different amplification efficiency. GenXPro has developed a method to eliminate this bias using its “TrueQuant” technique and can therefore- as the single provider world wide- offer PCR-bias free sequencing data (Fig.2).



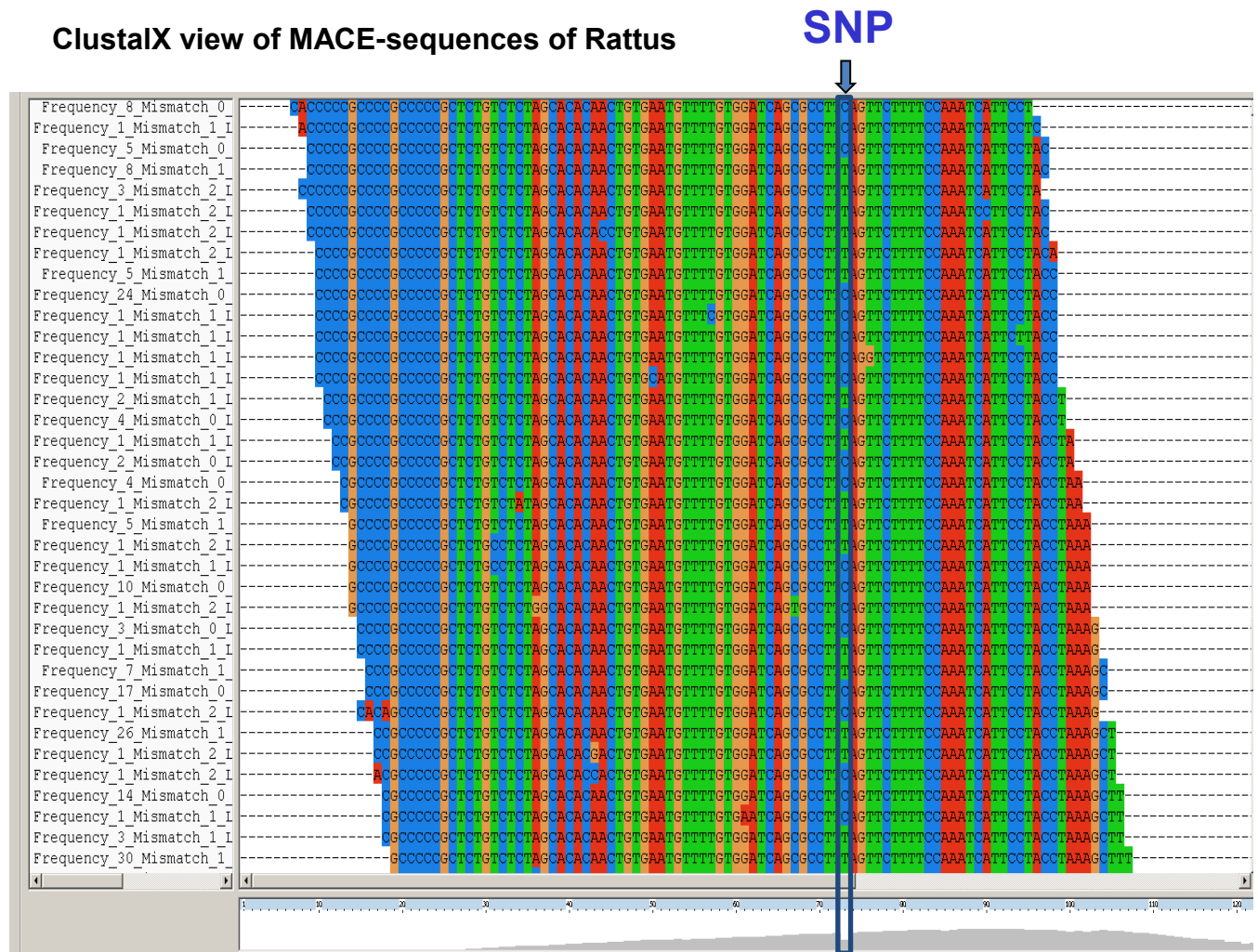
**Fig.2:** Negative common logarithm of the p-value for differential expression of gene expression comparisons (Audic & Claverie; 1997). If not corrected for PCR bias (TrueQuant; blue), the p-value changes dramatically as a consequence of biased data.

## TransNiPtomics using MACE

The analyzed 3' ends of transcripts are mostly consisting of 3' untranslated regions (3'-UTRs) which are under lower evolutionary pressure. As a consequence, these regions contain many sequence polymorphisms such as SNPs or Indels. Since these are directly located in genes, they represent highly valuable genomic markers, often directly linked to specific traits.

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

The figure below shows an excerpt of MACE-data from a laboratory strain of *Rattus Norvegicus*, analyzed by ClustalX. Even in very closely related laboratory rats, SNPs are common in the analyzed 3'UTRs.



## **Robustness: Partly degraded material e.g. of paraffin embedded samples**

As only the 3' ends of each cDNA molecule are sequenced, even partly degraded material such as from formalin-fixed paraffin-embedded specimen can be reliably analyzed

## **Accuracy / Coverage**

MACE analysis of 3'ends distinguishes between ~20.000 different transcripts in mammals. While it misses certain splice variants, it discovers and quantifies rare transcripts which are often not detected by RNAseq and microarrays.

## **High throughput SNPs and Gene Expression Analysis**

We can analyze 96 samples simultaneously for only a fraction of the costs for RNA seq- forget the financial restrains of RNAseq for large sample amounts!

Analyze hundreds of genotypes and obtain SNPs and quantitative gene expression simultaneously!

## **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!

**Stay competitive! Exploit the turnkey solutions of**

