BEP Barley Epigenome Platform/ Epigenom-Plattform für Gerste

- **Principal investigators and institutions:**
  - Prof. Dr. Klaus Humbeck, Martin Luther University,
  - Prof. Dr. Nils Stein and Dr. Martin Mascher, both Leibniz Institute of Plant Genetics and Crop Plant Research
- **Names of PhD students:** Charlotte Ost
- **Project duration:** 01.10.2016 – 30.09.2019
- **Funding amount:** 213.500 €
- **Industry partner:** Saaten-Union Biotec GmbH
- **Other research collaborations:** Dr. Hieu Cao, Research Focus Program "Molecular Biosciences as a Motor for a Knowledge-Based Economy"
- **Number of Bachelor and Master students working on the project and title and date of the thesis':** One Bachelor student working on the project, title: "Analyses of factors involved in epigenetic regulation of barley plants under drought stress", 2018

Die Fähigkeit von Nutzpflanzen, Schutzmechanismen gegen Stressbedingungen auszubilden, bestimmt im Endeffekt den Ertrag. Wir wollen von der Natur lernen, wie Pflanzen sich schützen können. Da die pflanzlichen Stressantworten von der koordinierten Expression der richtigen Gene zur richtigen Zeit abhängen, ist das Verständnis dieser Mechanismen eine Grundvoraussetzung für gezielte Züchtungsansätze. Neuere Forschungsarbeiten haben gezeigt, dass die Genexpression gerade auch unter widrigen Umweltbedingungen einer epigenetischen Kontrolle unterliegt. Das Projekt hat sich zum Ziel gesetzt, eine Epigenom-Plattform für Getreidepflanzen zu etablieren, die die genomweite und räumliche Identifizierung epigenetischer Regulatoren ermöglicht.

I. Project Parameters

1. Topic, open questions and purpose

Stress-induced, premature leaf senescence in crop plants severely impairs yield and causes massive economic losses. To cope with this problem, we need to understand the underlying molecular mechanisms. Recent research revealed that stress responses and leaf senescence are controlled by higher order epigenetic mechanisms. Our innovative goal is to establish at the Science Campus Halle an epigenome platform for crop plants which enables genome wide analyses of epigenetic histone and DNA modifications. Using this platform, in a first step we want to identify the epigenetic key regulators of leaf senescence. This will be very important for targeted breeding approaches aiming at crops which are more tolerant to stress.

2. Interdisciplinarity of the project

The joined project combines the expertise of three groups, two of them from Leibniz Institute of Plant Genetics and Crop Plant Research, and one from Martin-Luther-University Halle-Wittenberg. Its scientific background covers topics from Plant Biology, Plant Physiology, Plant Biochemistry and Plant Genetics. The outcome of the project will be the establishment of a barley epigenome platform at the Science Campus Halle which will allow identification of epigenetic key regulatory processes of plant stress responses and plant development. This knowledge is important for targeted breeding approaches. The project is primarily located in the first WCH research pillar, the primary plant production, but reaches out in the second pillar, processing/conversion.

3. Innovativity of the research

Only recently it was realized that plants performance in a changing and challenging environment is regulated via epigenetic mechanisms which control gene expression. While intensive epigenetic research in medicine already lead to highly interesting findings which were implemented in industrial production of therapeutics, research on plant epigenetics is still in the beginning. Furthermore, most of the work so far was performed with the model plant *Arabidopsis thaliana*. However, knowledge about epigenetic control of plant development and abiotic/biotic stress responses is highly needed, to be implemented in breeding efforts for better crops. Now, the methodology is available for this kind of research. In this project the expertise in analyzing epigenetic reprogramming in barley (Humbeck lab), in barley genome next generation sequencing (Stein lab) and in bioinformatics tools to analyze the ChIP-seq data (Mascher lab) is combined.

4. Application-oriented research

The knowledge of epigenetic key players in regulation of plants performance (development and stress responses) is central for future targeted breeding approaches. A barley epigenome platform at Science Campus Halle will allow application-oriented research in this direction. The techniques established already allowed to start a focused project together with Saaten Union Biotec GmbH, in which the epigenetic basis for low efficient double haploid production in breeding industry is analyzed. This project is funded by the MLU-based Research Focus Program "*Molecular Biosciences as a Motor for a Knowledge-Based Economy*".

5. Relevance of the project

The project is a combined effort of three groups of two members of the Science Campus (Leibniz Institute of Plant Genetics and Crop Plant Research and Martin-Luther-University Halle-Wittenberg, Faculty I of Natural Science). It is concentrating on the molecular basis of primary plant production via investigating higher-order control of plant genes during growth and development of plants, including their interaction with abiotic and biotic stressors. This

knowledge is relevant for plant-based bioeconomy, as also proven by the interest of breeding industry. The barley epigenome platform located at the Science campus Halle is intended to be open for researchers in the research focus “Plant-based Bioeconomy” of Saxony-Anhalt and for cooperation with industry located in Saxony-Anhalt (e.g. Saaten Union Biotech GmbH in Gatersleben).

II. Theory and methods

Leaf senescence, the latest phase in leaf development, is an efficient recycling program. Valuable resources are degraded in the senescing leaves, recycled within the plant and used for young, growing leaves or finally for the developing seeds (reviewed in Jing et al. 2003). This is achieved by complex reprogramming of gene expression and senescence-associated genes could be identified in several recent transcriptome studies (e.g. Breeze et al. 2011; Buchanan-Wollaston et al. 2005). Leaf senescence can be induced either by internal or by external factors. One major problem of stress-induced senescence is premature leaf senescence which leads to inefficient recycling causing massive losses in yield (reviewed in Gregersen et al. 2013).

In recent years it became obvious that there is, in addition to the trans-acting factors, a higher order, epigenetic control level of plant senescence. Epigenetic regulation is characterized through changes in the chromatin status manifested via modifications at DNA or associated histones, but not in the DNA sequence. The chromatin status is determined by different epigenetic mechanisms. Our project has its focus on histone modifications and DNA methylation. Histone modifications are characterized by post-translationally modifying the N-terminal tails of histones by specific enzymes which are adding different side groups (e.g. methyl, acetyl or phosphate) or removing them. Thereby the chromatin status is changing and genes which are associated with these marks get regulated. The methylation of N-terminal histone tails can be associated with either activating (e.g. H3K4me3 and H3K9ac) or repressing (e.g. H3K9me2) genes (Berger, 2007; He et al. 2011; Kouzarides, 2007). A further mechanism of epigenetic modification is DNA methylation which is characterized by adding a methyl group to a nucleotide, mainly cytosine (Richards 1997, Law et al. 2010). One of the main functions of DNA methylation is the maintenance of genome stability due to silencing of transposons (Yoder et al. 1997, Chan et al. 2005). The second important function is regulating gene expression by methylation of the promotor and other gene regions (Zhang X et al. 2006, Zilberman et al. 2007). Since the first insights of the distribution of DNA methylation at the *HvS40* gene in barley (Ay et al. 2014) there is a demand for a genome-wide Bisulfite sequencing. The group of Klaus Humbeck could show for the first time that leaf senescence in *Arabidopsis thaliana* is under control of epigenetic mechanisms (Humbeck 2013; Ay et al. 2009; Ay et al., 2014a; Ay et al. 2014b). Most of the reports about epigenetic processes in plants were done with *Arabidopsis thaliana*. However, there is an increasing demand to analyze the epigenome also in economically important crop plants. Meanwhile, the experimental techniques to analyze epigenetic processes in the group of Klaus Humbeck

were successfully transferred to barley (*Hordeum vulgare*), e.g. in Ay et al. (2015). In addition, the recent deciphering of the barley genome by the groups of Nils Stein and Martin Mascher and colleagues was a milestone in barley genome research and is a prerequisite for genome wide epigenetic studies also in barley (Beier et al. 2017, Mascher, 2019). The methodology to analyze the epigenome in barley includes defined growth of barley plants under standardized control and stress conditions, exact mapping of the kinetics of the senescence program via measurement of chlorophyll content and photosynthetic parameters, chromatin immunoprecipitation of cross-linked histone-DNA fragments (ChIP), deep sequencing of immunoprecipitated fragments (ChIPseq) and bioinformatic analyses of sequence data. In addition, we aim at analyzing the transcriptome via RNA-sequencing and the methylome via Bisulfite sequencing. Currently, the ChIP- and RNA-sequencing is finished and the Bisulfite sequencing is in process by the IPK Gatersleben.

III. Results and perspectives

1. Results

The development of control and drought stress treated primary barley leaves was documented by measuring PS II efficiency and relative chlorophyll content (Fig.1A).

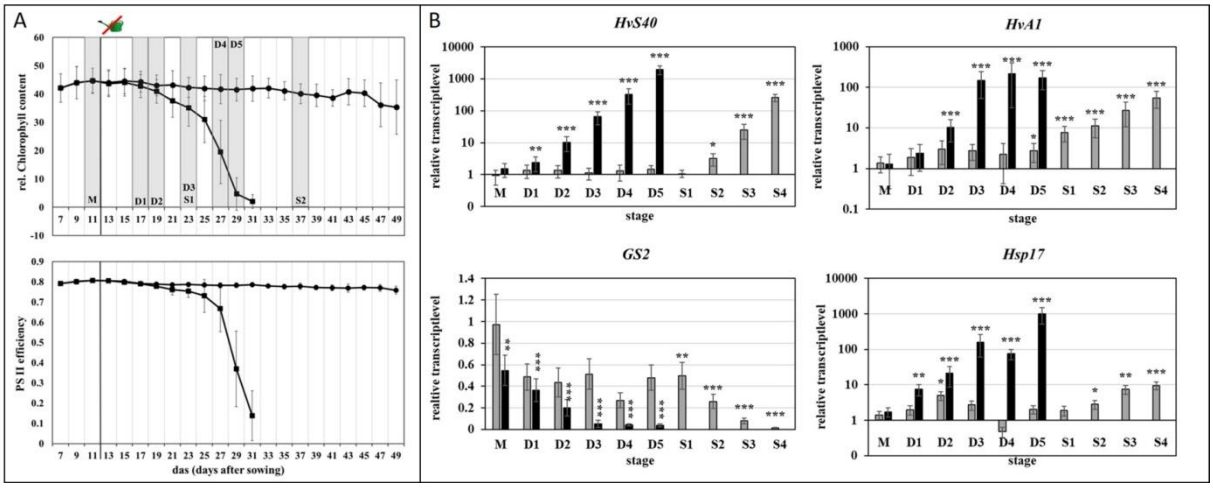
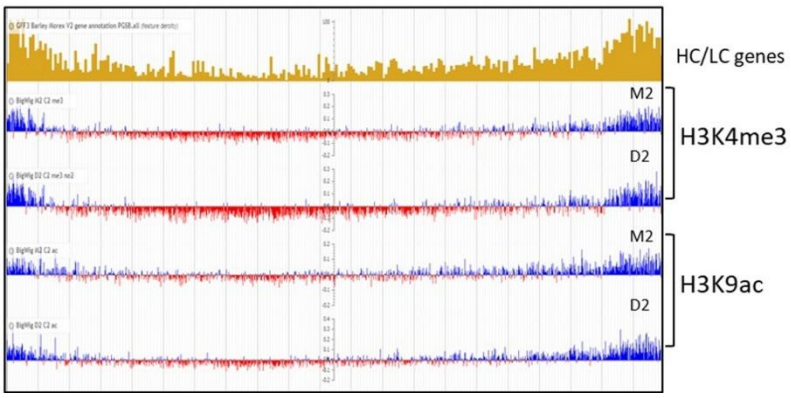


Figure **Fehler! Kein Text mit angegebener Formatvorlage im Dokument.**-1: Physiological and molecular characterization of barley plants under developmental and drought stress conditions. (A) Top: Relative chlorophyll content of control (●) and drought stress (■) samples with the defined drought stress (D) and developmental senescence (S) stages. Bottom: PS II efficiency of control and drought stress samples (n=3). (B) Molecular characterization of the marker gene expression. The diagrams show the relative transcript level of control (grey) and drought stress (black) samples for *HvS40*, the putative key regulator of leaf senescence in barley, *HvA1*, which belongs to the group III LEA proteins, Glutamine synthetase II (*GS2*) and *Hsp17*, a heat shock protein. (*) indicate statistically significant differences (t test: *p < 0.05; **p < 0.01; ***p < 0.001). Data represent the mean (±SD) from at least three biological replicates.

Whereas PS II efficiency and chlorophyll content stay high under control conditions, drought stress causes a premature decrease of these parameters reflecting onset of stress-dependent leaf senescence. To pinpoint the reprogramming of gene expression during drought stress induced leaf senescence, transcript levels of marker genes for leaf senescence and drought stress were analyzed via qRT-PCR (Fig. 1B). Opposite to the three senescence-

associated genes (SAGs), *HvA1*, *HvS40* and *Hsp17*, which are induced during the stress treatment, *HvGS2* is a senescence down regulated gene (SDG). The ChIP analyses were carried out after the protocols of Gendrel (2005) and Ay et al. (2009). Figure 2A shows the read distribution of the M2 (control) and D2 (drought treated) samples over the chromosome 1H for H3K4me3 and H3K9ac.

A



B

H3K4me3	Peaks Ø	Genes
M2	210,844	30,735
D2	193,424	26,595
H3K9ac	Peaks Ø	Genes
M2	214,370	21,805
D2	200,077	23,232

Figure **Fehler! Kein Text mit angegebener Formatvorlage im Dokument.**-2: A) Read coverage of the M2 and D2 samples of chromosome 1H for H3K4me3 and H3K9ac. (B) Computed peaks (enriched read regions) and identified genes. For the peaks the average value of the four replicates were calculated. Genes which are in at least three of four replicates were taken.

Enrichment of reads at specific genes and time points for a particular histone modification were marked as a peak and genes were identified among these peaks (Fig. 2B). Splitting the gene lists in the regions, where the peak appears, revealed that 27 % of the genes showed a H3K4me3 mark at the Promotor and ATG (PROATG) region only in M2. 14% of the genes showed a mark for H3K4me3 only in D2 in the PROATG region. Most of the genes (59 %) were shared between the control and the drought sample. For the H3K4me3 mark in the body region 74 % of the genes are shared between M2 and D2 and only 6 % are unique for D2 and 20 % are specific for M2. For the acetylation of H3K9Ac, 14 % genes are specific for M2 and 33 % are specific for D2 (Fig.3).

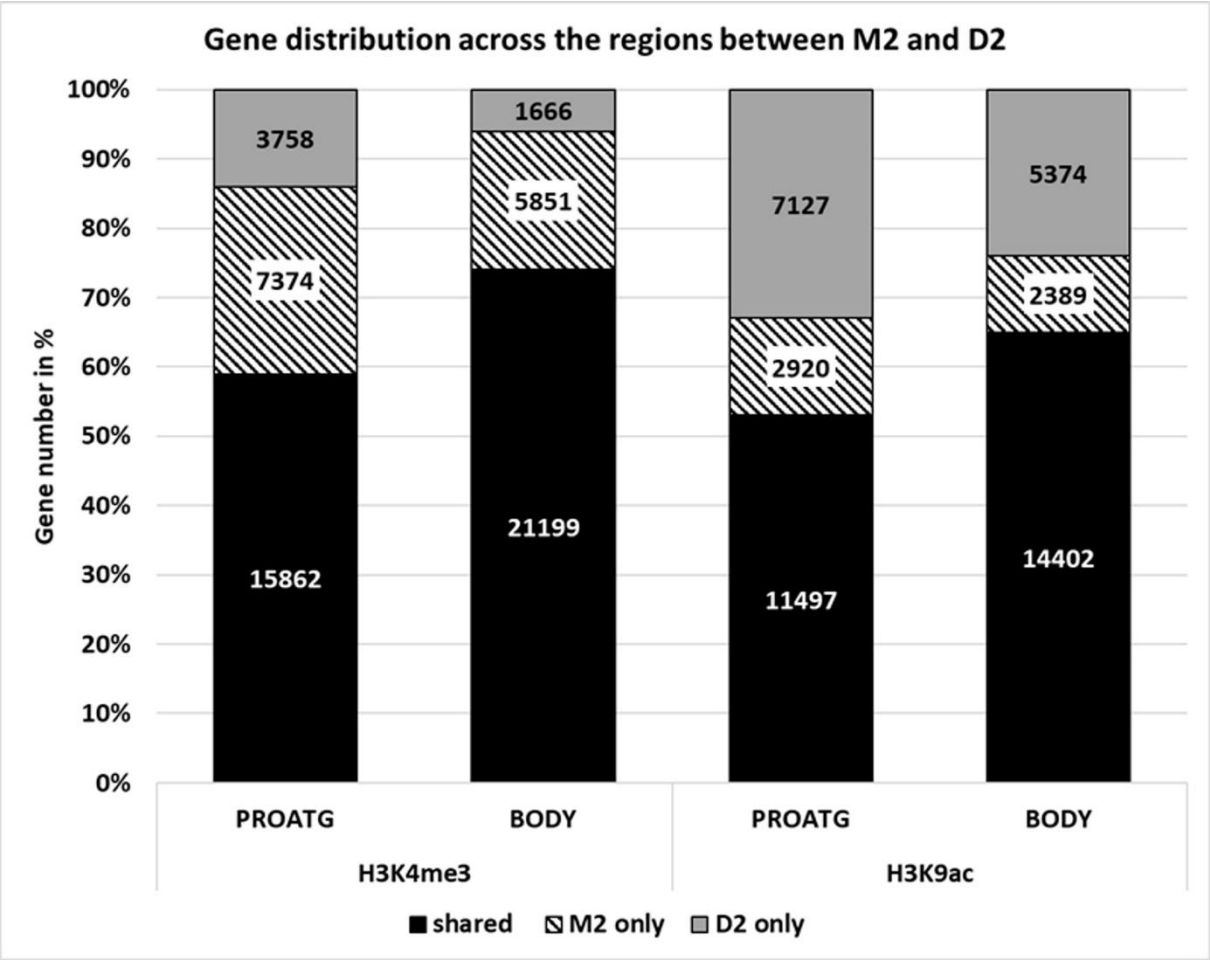
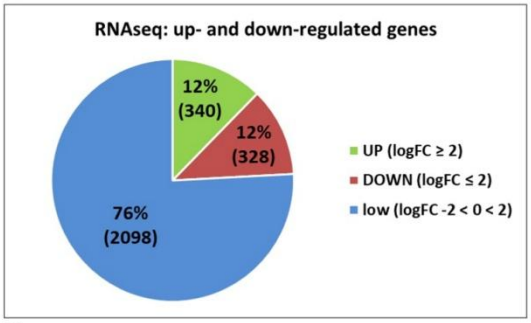


Figure **Fehler! Kein Text mit angegebener Formatvorlage im Dokument.**-3: Distribution of unique and shared H3K4me3/H3K9ac modified genes between control (M2) and drought stress (D2) samples. The numbers represent the total count of genes.

For transcriptome studies, RNA-sequencing was performed with four replicates. In total, a list of 22,038 differentially expressed genes (DEGs) was generated. After removing all DEGs with p-Values above 0.05 and only focusing on significantly up- or downregulated genes, 340 up- and 328 down-regulated genes were identified (Fig. 4A).

A



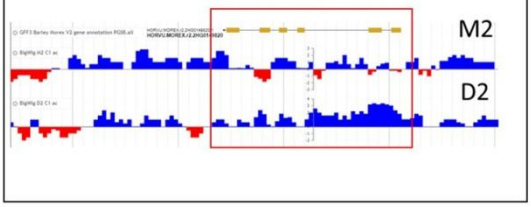
B

UP-regulated genes		
GO_ID	Name	p-Value
GO:0009698	phenylpropanoid metabolic process	6.167E-09
GO:0046271	phenylpropanoid catabolic process	1.68E-08
GO:0046274	lignin catabolic process	1.68E-08
GO:0052716	hydroquinone:oxygen oxidoreductase activity	2.18E-08
GO:0016491	oxidoreductase activity	1.235E-07
GO:0055114	oxidation-reduction process	1.417E-07
GO:0009808	lignin metabolic process	2.053E-07
GO:0016682	oxidoreductase activity, acting on diphenols and related substances as donors, oxygen as acceptor	2.246E-07
GO:0004722	protein serine/threonine phosphatase activity	3.283E-07
GO:0016679	oxidoreductase activity, acting on diphenols and related substances as donors	1.71E-06

DOWN-regulated genes		
GO_ID	Name	p-Value
GO:0009765	photosynthesis, light harvesting	4.10E-21
GO:0015979	photosynthesis	1.97E-19
GO:0018298	protein-chromophore linkage	9.00E-14
GO:0019684	photosynthesis, light reaction	1.24E-12
GO:0009853	photorespiration	7.45E-12
GO:0015977	carbon fixation	2.93E-09
GO:0043094	cellular metabolic compound salvage	4.15E-09
GO:0015035	protein disulfide oxidoreductase activity	9.27E-09
GO:0006091	generation of precursor metabolites and energy	1.10E-08
GO:0015036	disulfide oxidoreductase activity	3.31E-08

C

gene	annotation	logFC	TPM M2	TPM D2
HORVU.MOREX.r2.2HG0149820	Caleosin	5.55	2.03	129.80



D

H3K4me3	UP	DOWN	# of genes overlap	# of all genes
M2	40	66	275	5563
D2	19	7	73	1423
H3K9ac				
M2	31	42	161	3594
D2	73	51	331	5021

Figure **Fehler! Kein Text mit angegebener Formatvorlage im Dokument.**-4: Analysis of the RNAseq. (A) Distribution of up- and downregulated genes. (B) List of top ten significantly enriched GO terms for the up- and downregulated genes. (C) Example of the gene Caleosin which is upregulated during drought and shows a mark for H3K9ac specific in the drought stress sample. (D) Comparison of genes loaded with a histone mark and the list of DEGs.

76 % (2098) of the genes of the DEG-list show zero to low expression. The GO term Enrichment analysis with WEADE (Fig. 4B) presents enriched biological processes, e.g. the up-regulated genes show enrichment in the phenylpropanoid process and in the oxidoreductase activity. For the down-regulated genes there is enrichment in photosynthetic related processes such as light reaction or carbon fixation. In a next step, the gene lists from the ChIPseq were compared with the up- and downregulated genes from the RNAseq to identify a possible correlation between histone modifications and gene expression in drought stress (Fig. 4D). Interestingly, only a low proportions of the genes loaded with a histone mark seem to be expressed differentially during drought. Furthermore, genes with a histone mark specific for M2 are more downregulated. Figure 4C presents exemplary the caleosin gene, which is upregulated during drought (Aubert et al. 2010). Caleosin is loaded with the H3K9ac mark at the PROATG region specific in the drought stress sample. The view of JBrowse shows the chromosomal region of the gene and its read coverage for the control and drought stress sample.

IV. Index

Aubert et al. 2010. *Plant and Cell Physiology*. *51. Jg., Nr. 12, S. 1975-1987*

Ay et al. 2009. *Plant Journal*. *58*, 333-346.

Ay et al. 2015. *Plant Molecular Biology* 89, 127–41.

Ay et al. 2014. *Journal of Experimental Botany 65*, 3875-3887.

Beier et al. 2017. *Sci Data4*, doi: 10.1038/sdata.2017.44

Berger 2007. *Nature* 447 (7143): 407-412

Bernatavichute et al. 2008. *PLoS ONE* 3 (9), doi: 10.1371/journal.pone.0003156

Breeze et al. 2011.*The Plant Cell* 23 (3): 873–94.

Buchanan-Wollaston et al. 2005. *Plant Journal* 42 (4): 567–85.

Chan et al. 2005. *Nature Reviews Genetics*, *6*(5), 351.

Gendrel et al. 2005. *Nature Methods* 2 (3): 213–18.

Gregersen et al. 2013. *Plant Molecular Biology*. doi:10.1007/s11103-013-0013-8.

He et al. 2011. *Annu. Rev. Plant Biol* 62: 411–35.

Humbeck K. 2013. *Plant Molecular Biology 82*, 529-537.

Jing et al. 2003. *Plant Biology* 5 (5): 455–64.

Kouzarides 2007. *Cell* 128 (4): 693-705.

Law et al. 2010. *Nature Reviews Genetics*, *11*(3), 204-220.

Mascher et al. 2017. *Nature* 544 (7651): 427–33.

Mascher 2019.*e!DAL*. doi: 10.5447/IPK/2019/8

Richards, E. J. 1997 *Trends in Genetics*, *13*(8), 319-323.

Yoder et al. 1997 *Trends in genetics*, *13*(8), 335-340.

Zhang X et al. 2006. *Cold Spring Harbor symposia on quant. biology* (Vol. 71, pp. 439-447)

Zhang et al. 2009. *Genome Biology* 10 (6): 1-14

Zilberman et al. 2007. *Nature genetics*, *39*(1), 61-69.

Programs

Trost et al. WEADE: a workflow for enrichment analysis and data exploration. *PloS one*, 2018, 13. Jg., Nr. 9.

Skinner et al. JBrowse: a next-generation genome browser. *Genome research*, 2009, 19. Jg., Nr. 9, S. 1630-1638.