Supplementary Material

Supplementary Materials and Methods

Isolation and mapping of *fas* mutants

Mutations were induced by feeding EMS (25 mM) to males carrying a second chromosome marked with *btl*-Gal4, UAS-GFP and UAS-Verm-RFP transgenes. Mutagenized chromosomes were balanced over the *CyO Dfd-GMR-nvYFP* chromosome (Le *et al*., 2006). Living F3 embryos were analyzed for tracheal morphology and the distribution of Verm-RFP protein (Förster *et al*., 2010). Mutations were classified according to their phenotypes. Within each phenotypic class, mutants were crossed to each other to test for complementation of lethality.

The lethality of *H124* and *P218* mutations was mapped to the cytological interval 49C1-50D2 by non-complementation of the chromosomal deficiency *Df(2R)CX1*. Within this interval, 14 overlapping deficiencies (FlyBase) complemented *H124* and *P218* mutants. A small interval (50B4-B6) comprising six annotated genes was not covered by these deficiencies and contained a lethal P-element insertion (*P{PZ}05488*), which was allelic to the *fas* locus (Lekven *et al*., 1998) and failed to complement *H124* and *P218* mutants.

Sequencing of *fas* alleles

Genomic DNA was extracted from homozygous fas^{H124}, fas^{P218} and fas^{1} embryos selected by the absence of the *CyO Dfd-YFP* balancer chromosome (Le *et al*., 2006). Coding sequences of *CG6197* and *CG17716* were amplified and sequenced using oligonucleotides listed in Supplementary Table 3.

Statistics and reproducibility

For phenotypic analyses, sample size (*n*) was not defined using statistical methods, but was determined by taking into account the variability of a given phenotype. Investigators were not

blinded to allocation during experiments and samples were not randomised for the allocation of samples to experimental and control groups.

RNASeq data analysis

For differential gene expression, kallisto (v 0.42.1; Bray *et al*., 2016) was used to estimate the read count for each annotated transcript in each sample, and the counts were then summarized on the gene level. Genes with an estimated CPM (counts per million) exceeding one for at least three samples were retained for differential expression analysis and were considered as 'expressed'. The GLM framework of the edgeR R package (v 3.10.2; McCarthy *et al*., 2012; Robinson *et al*., 2010) was applied to test for differential gene expression between the control and mutant groups, using the sample preparation batch as a confounding factor in the model.

DEXSeq (Anders *et al*., 2012) was used to investigate the extent of intron retention. The reads were aligned to the *Drosophila melanogaster* genome (Ensembl v70) with STAR (v2.4.2a) (Dobin *et al*., 2013), retaining only reads aligning to a unique location. Based on the kallisto results obtained as described above, transcripts contributing less than 5% to the total abundance (TPM) of the corresponding gene in all samples were removed from the gtf file (Soneson *et al*., 2016). The reduced gtf file was processed with the python scripts provided with DEXSeq to "flatten" the annotation file by generating disjoint exon bins. Finally, the flattened annotation file was extended with intronic bins, defined as intragenic regions that did not overlap with any exon of the retained transcripts. Using this extended annotation, DEXSeq was used to quantify the abundance of each exonic and intronic bin, and to test each bin for differential inclusion between the control and mutant groups. Only results for intronic bins were retained for interpretation.

FPKMs were estimated for each exonic and intronic bin by dividing the normalized counts obtained from DEXSeq (adding 1 to avoid taking the log of 0) by the width of the bin and the total number of counts (exonic and intronic) for the corresponding sample.

The RNA-seq data has been deposited in ArrayExpress under accession number E-MTAB-5069.

Supplementary References

Anders, S., Reyes, A. and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017.

Bray, N. L., Pimentel, H., Melsted, P. and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.

Förster, D., Armbruster, K. and Luschnig, S. (2010). Sec24-dependent secretion drives cell-autonomous expansion of tracheal tubes in Drosophila. *Curr. Biol.* **20**, 62–68.

Le, T., Liang, Z., Patel, H., Yu, M. H., Sivasubramaniam, G., Slovitt, M., Tanentzapf, G., Mohanty, N., Paul, S. M., Wu, V. M., et al. (2006). A new family of Drosophila balancer chromosomes with a w-dfd-GMR yellow fluorescent protein marker. *Genetics* **174**, 2255–2257.

Lekven, A. C., Tepass, U., Keshmeshian, M. and Hartenstein, V. (1998). faint sausage encodes a novel extracellular protein of the immunoglobulin superfamily required for cell migration and the establishment of normal axonal pathways in the Drosophila nervous system. *Development* **125**, 2747–2758.

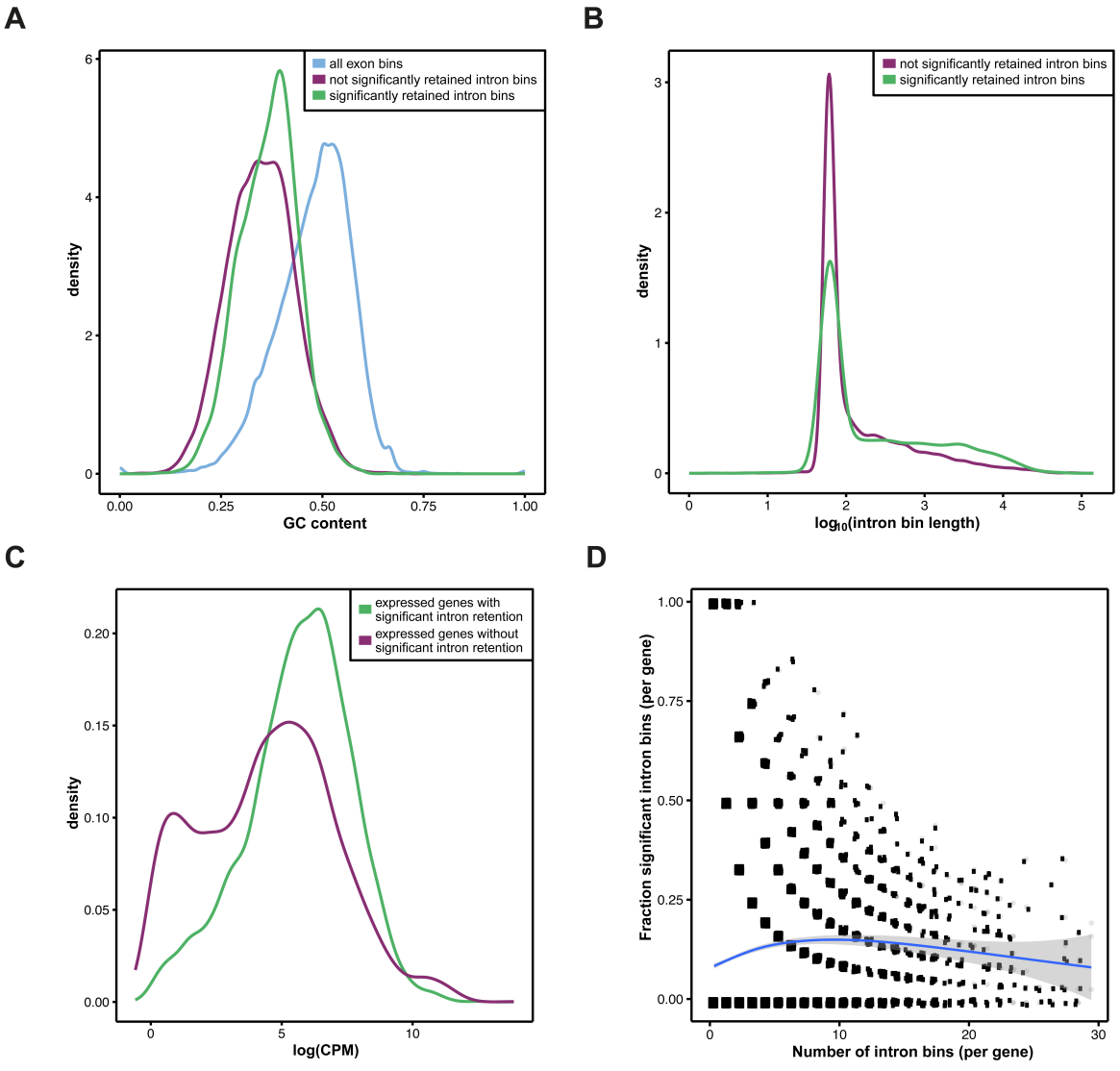
McCarthy, D. J., Chen, Y. and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**, 4288–4297.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140.

Soneson, C., Matthes, K. L., Nowicka, M., Law, C. W. and Robinson, M. D. (2016). Isoform prefiltering improves performance of count-based methods for analysis of differential transcript usage. *Genome Biol.* **17**, 248.

Supplementary Figures

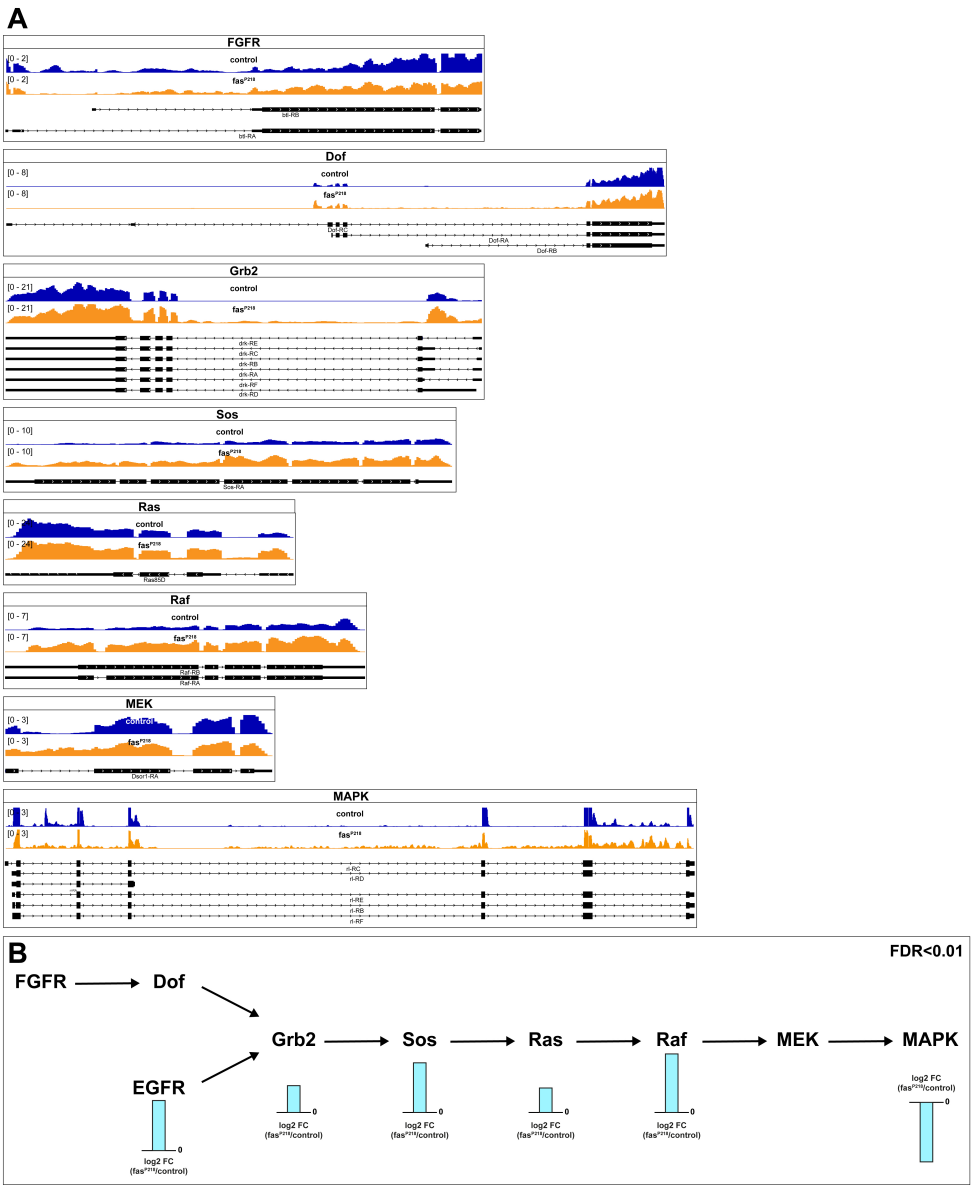
Sauerwald_Supplementary_figure_1



Supplementary Figure 1

Characterization of significantly retained intron bins and the genes containing them. (**A**) Distribution of GC content for the significantly retained intron bins (adjusted $p < 0.01$, log2FC > 1.5), not significantly retained intron bins (adjusted $p > 0.1$) and all exon bins. Significantly retained intron bins show slightly increased GC content compared to not significantly retained intron bins. (**B**) Distribution of the bin lengths of significantly retained intron bins (adjusted $p < 0.01$, log2FC > 1.5) and not significantly retained intron bins (adjusted $p > 0.1$). The significantly retained intron bins are slightly enriched with long bins. This association may potentially be explained by the higher power of count-based methods to give significant test results for features with large read counts, since longer bins give rise to more fragments than short bins, and thus a higher number of reads. (**C**) Distribution of expression values for expressed genes (i.e., genes not filtered out in the differential gene expression analysis) containing any of the significantly retained introns or not. Similarly to (B), the difference may be partly explained by the higher power of count-based methods to detect differences in features with high numbers of counts. (**D**) The association between the number of intron bins per gene and the fraction of these showing significant retention (adjusted $p < 0.01$, log2FC > 1.5). Each dot represents a gene, and dots have been jittered slightly to decrease overplotting. The blue line represents a smooth fit to the points. No strong association between the expected fraction of significant bins and the number of bins is apparent.

Sauerwald_Supplementary_Figure_2



Supplementary Figure 2

Abnormal transcript processing of receptor tyrosine kinase signaling components.

(**A**) Coverage plots of transcripts encoding components of EGF and FGF signaling. (**B**) Significant (FDR < 0.01) changes in transcript levels of EGF and FGF signaling components.

Supplementary Table 1

Analysis of intron retention on intron-bin-level using DEXSeq. Columns include the gene identifiers, gene names, details about the corresponding intronic bins, fold changes (FC) for intron retention and statistics relating to DEXSeq analysis.

Click here to Download Table S1

Supplementary Table 2

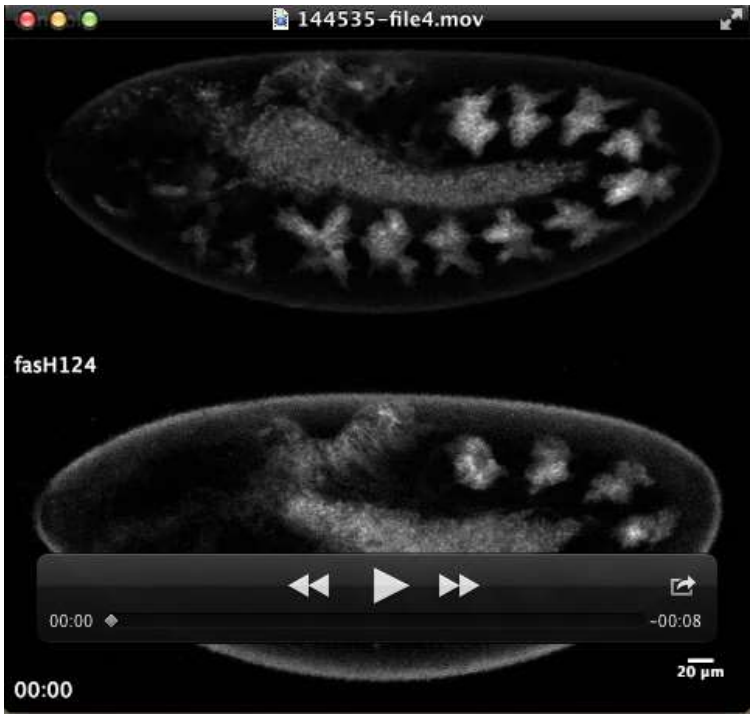
Gene-level differential expression analysis using kallisto and edgeR. Columns include the gene identifiers, gene names, fold changes (FC) and statistics relating to differential expression analysis.

Click here to Download Table S2

Supplementary Table 3

List of oligonucleotides used for sequencing of *CG6197* and *CG17716* coding region.

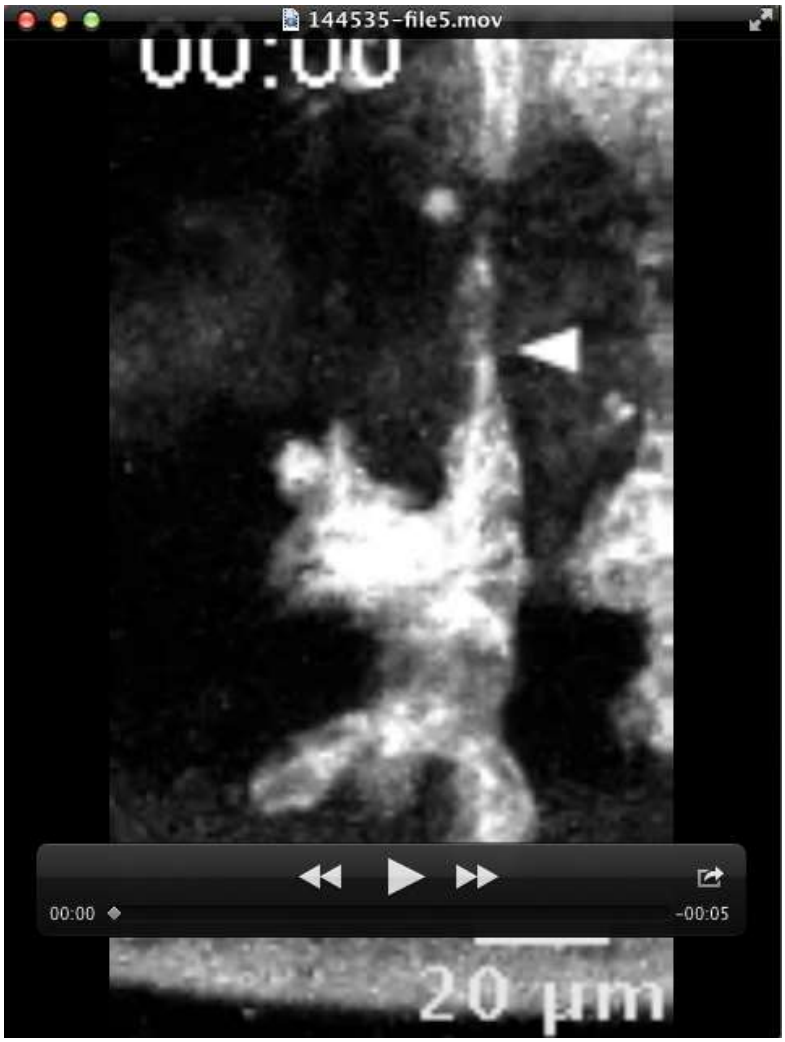
primer	sequence	description
JuS15	ACTTTGGTATTAGCTGCGGTA	sequencing *CG6197* exon 1-2
JuS16	CGCACAATTCGTTCCACAGC	sequencing *CG6197* exon 1-2
JuS17	CTGCGAGTCTACCGTCGATA	sequencing *CG6197* exon 2
JuS18	GCACCGTTTCCGTATCATCA	sequencing *CG6197* exon 2
JuS19	GGAGTTCGCCAAGTTCTACG	sequencing *CG6197* exon 2
JuS20	CACTTCGCCCAACTTCGTTT	sequencing *CG6197* exon 2
JuS21	AAGGCAGCCGAGATCTATGG	sequencing *CG6197* exon 2-3
JuS22	TCTTGGATTGCGCGTAAAAGC	sequencing *CG6197* exon 2-3
SL79	GTCTTTGTTTGCCCTGCACT	sequencing *CG17716* exon 3
SL80	ACCTGCCTGACTGCGATTAT	sequencing *CG17716* exon 3
JuS9	GGGACCAAGGAATTGCAGTG	sequencing *CG17716* exon 4
JuS10	TGAGACAGACGGAGGGAAAA	sequencing *CG17716* exon 4
JuS7	TTGCCACATTACCAGTGGAT	sequencing *CG17716* exons 5-6
JuS8	GCTCGAGTGCAAGACAAACA	sequencing *CG17716* exons 5-6
SL71	TACACCCCCGATGATTGATT	sequencing *CG17716* exon 7
SL72	GATACCCAAGCGCACTGTTT	sequencing *CG17716* exon 7
SL73	CACTCTGCGAGCCTAGCATT	sequencing *CG17716* exon 8-9
SL74	ATGTGCGGGTGTCGTAAAAT	sequencing *CG17716* exon 8-9
SL75	CAACAGCCCTCGGATTTG	sequencing *CG17716* exon 10-11
SL76	CTGATTCATCGATTGGGTTG	sequencing *CG17716* exon 10-11
SL77	ATGCGAGGCGAAACAAACT	sequencing *CG17716* exon 12-13
SL78	GACTAGGGCAATTTGGATGC	sequencing *CG17716* exon 12-13



Supplementary Movie 1

Tracheal branch outgrowth is affected in *fas*^{H124} mutant embryo.

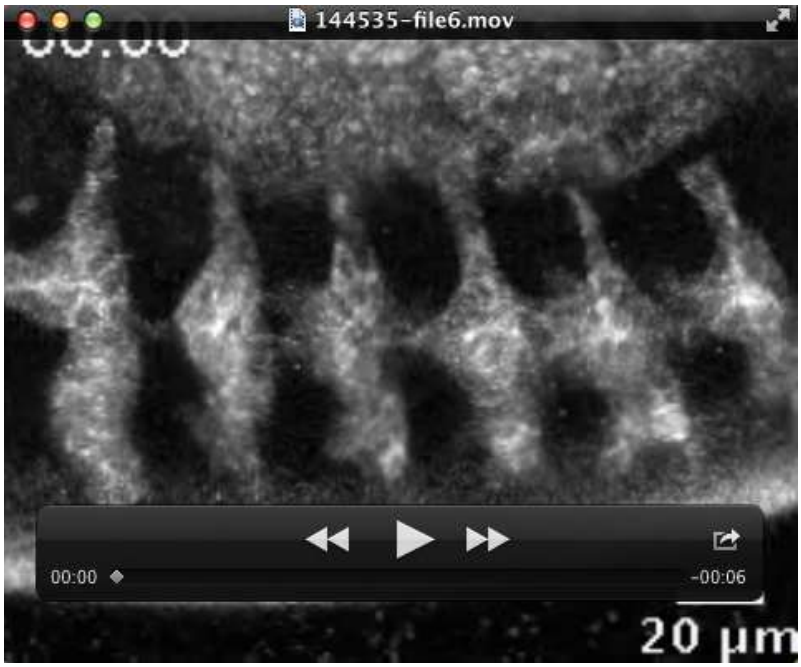
Time-lapse movies of wild-type (top) and *fas*^{H124} homozygous mutant (bottom) embryos expressing palmitoylated mNeonGreen under the control of *btl*-Gal4. The movies were acquired with a 20x/0.75 NA objective and a frame rate of 5 min.



Supplementary Movie 2

Dorsal branch separating from the tracheal primordium in *fas*^{H124} mutant embryo.

Time-lapse movie of tracheal metamere one in a *fas*^{H124} homozygous mutant embryo expressing palmitoylated NeonGreen under the control of *btl*-Gal4. The movie was acquired with a 20x/0.75 NA objective and a frame rate of 5 min.



Supplementary Movie 3

Formation of dorsal trunk connections in *fas*^{H124} mutant embryo.

Time-lapse movie of tracheal metameres forming DT connections in a *fas*^{H124} homozygous mutant embryo expressing palmitoylated NeonGreen under the control of *btl*-Gal4. The movie was acquired with a 20x/0.75 NA objective and a frame rate of 5 min.