# Chapter 8

# Using Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE) to Identify Functional Regulatory DNA in Insect Genomes

## Daniel J. McKay

## Abstract

Differential regulation of gene expression determines cell-type-specific function, making identification of the *cis*-regulatory elements that control gene expression a central goal of developmental biology. In addition, changes in the sequence of *cis*-regulatory elements are thought to drive changes in gene expression patterns between species, making comparisons of *cis*-regulatory element usage important for evolutionary biology as well. Due to the number of extant species and the incredible morphological diversity that they exhibit, insects are favorite model organisms for both developmental and evolutionary biologists alike. However, identifying *cis*-regulatory elements in insect genomes is challenging. Here, I describe a method termed FAIRE-seq (Formaldehyde-Assisted Isolation of Regulatory Elements, followed by high-throughput sequencing) that can be used to identify functional DNA regulatory elements from developing insect tissues, genome-wide.

**Key words** Open chromatin, *Cis*-regulatory element, Enhancer, FAIRE, *Drosophila* development

## 1 Introduction

Identifying *cis*-regulatory elements in insect genomes faces multiple challenges. As in other organisms, *cis*-regulatory elements such as enhancers and promoters are small, typically around 1000 base pairs in length. Enhancers in particular can also function over long distances and in a position-independent manner, meaning that they lack a stereotypical location relative to the genes they regulate. And in contrast to protein-coding genes, *cis*-regulatory elements lack a universal genetic code. For these reasons, accurate prediction of *cis*-regulatory elements based on DNA sequence alone is difficult.

A variety of experimental methods have been employed to directly identify *cis*-regulatory elements in target cells. Due to their small size, collecting sufficient amounts of input cells for genomic assays can be challenging in insects, especially if cell-

type-specific information is desired. Likewise, many insects undergo rapid changes during development, limiting the time available to identify *cis*-regulatory elements with temporal-specific activity. Here, I present a method termed FAIRE-seq (Formaldehyde-Assisted Isolation of Regulatory Elements) [1] that has successfully been used to identify functional *cis*-regulatory elements in insects, genome-wide. FAIRE-seq is a simple method that is sensitive enough to be used with limiting amounts of input cells. It also does not rely on enzymes or antibodies, thus simplifying the ability to identify and compare *cis*-regulatory element usage between species. Finally, since FAIRE-seq relies on formaldehyde-fixed samples, it allows time for precise tissue dissections and collection of short-lived developmental stages.

FAIRE-seq identifies sites of "open" chromatin in the genome [1–6]. Due to the competition between transcription factors and nucleosomes for DNA binding, active *cis*-regulatory elements, which are typically bound by transcription factors, are often depleted of nucleosomes or contain nucleosomes that are remodeled in such a way as to decrease their interactions with DNA [7]. In the FAIRE-seq protocol, samples are briefly fixed with formaldehyde, sonicated to shear chromatin into smaller fragments, and then subjected to a phenol-chloroform extraction to isolate water-soluble DNA, which is then prepared for high-throughput sequencing. FAIRE is thought to work by exploiting the differences between histones and transcription factors for being crosslinked to DNA. Due to the large number of amino acid residues that react with formaldehyde, and because of the extensive contacts that they make with DNA, histones are more likely to be crosslinked to DNA than transcription factors. As a result, genomic sites bound by transcription factors will preferentially segregate to the aqueous phase during the phenol-chloroform extraction, whereas DNA crosslinked to protein will be trapped at the interface between the aqueous and organic phases.

The ease of performing FAIRE-seq has led to its use with a range of developmental samples, including whole embryos [8], dissected tissues [9], and FACS-purified cell populations [10], as well as with a variety of insects, including fruit flies [8, 11], mosquitoes [12], aphids [13], silkworms [14], and beetles [15]. FAIRE-enriched DNA has been found to correspond to functional and developmentally relevant *cis*-regulatory elements, including transcriptional enhancers [8–10]. And when combined with a developmental time course, FAIRE-seq can successfully identify enhancers with temporal-specific activity because the timing of regulatory element accessibility often coincides with the timing of its activity. Details of the protocol are described below. The protocol is written for performing FAIRE on *Drosophila* wing imaginal discs. Guidelines and suggestions for other sample types are contained in the "**Notes**" Subheading 4.

## 2  Materials

| | |
|---|---|
| ***2.1  Buffers and Reagents*** | 1. Apple-agar plates: Flask 1: 200 mL apple juice, 300 mL dH$_2$O. Flask 2: 25 g BactoAgar, 500 mL dH$_2$O, stir bar. Autoclave Flask 1 and Flask 2 for 30 min. Pour contents of Flask 1 into Flask 2. Mix on magnetic stirrer until cool enough to handle. Pour into 100 × 15 mm or 60 × 15 mm plastic petri dishes. Allow to harden overnight and store at 4 °C. |

2. 1 × PBS: 130 mM NaCl, 7 mM Na$_2$HPO$_4$, 3 mM NaH$_2$PO$_4$, pH to 7.4 with HCl. Autoclave to sterilize.

3. FA Fix Buffer: 4% paraformaldehyde, 50 mM HEPES pH 8.0, 100 mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0. This buffer must be prepared fresh.

4. Quench Buffer: 1 × PBS, 125 mM Glycine, 0.01% Triton X-100.

5. Buffer A: 10 mM HEPES pH 8.0, 10 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.25% Triton X-100, 1 mM PMSF.

6. FAIRE Lysis Buffer: 2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris–Cl pH 8.0, 1 mM EDTA pH 8.0.

7. TE: 10 mM Tris–Cl pH 8.0, 1 mM EDTA pH 8.0.

***2.2  Equipment and Consumables***

1. Screw cap tubes with 2.38 mm tungsten beads for tissue homogenization, MoBio, catalogue # 13117-50.

2. Mini-BeadBeater-8, Biospec products.

3. Branson model 450-digital sonifier.

4. QIAquick PCR purification kit (Qiagen).

5. QuBit dsDNA HS Assay kit (Invitrogen).

## 3  Methods

***3.1  Dissection and Fixation of Drosophila Wing Imaginal Discs***

1. Collect larvae of the target developmental stage. For precise developmental staging, set up an embryo-collection cage of the desired genotype. Let adults lay eggs for several hours on an apple-agar plate, then transfer the plate off the cage and incubate the plate at 25 °C. After 20–24 h, use forceps to clear all larvae from the plate, and return the plate to 25 °C. After an additional 2 h, use forceps to pick newly hatched larvae. Transfer larvae in batches of ~35 to a fresh vial containing standard *Drosophila* food. Incubate at 25 °C until the target developmental stage (*see* **Note 1**). We typically use 40 wing imaginal discs from wandering third instar larvae for a single FAIRE-seq replicate. Since each wing disc contains approximately 50,000 cells, this corresponds to about 2-million cells per replicate.

However, we have also successfully performed FAIRE-seq from 40 haltere imaginal discs, which contain approximately 10,000 cells per disc. Scale culture sizes for each tissue of interest accordingly (*see* **Note 2**).

2. Wash larvae extensively with $1 \times$ PBS to remove any traces of food. Transfer larvae to wells containing $1 \times$ PBS in a 9-well dissecting dish. Screen through larvae to select females (*see* **Note 3**). Dissect and invert larvae, leaving the wing discs attached to the cuticle (*see* **Note 4**). Transfer inverted larvae to cold $1 \times$ PBS in a 9-well dissecting dish on ice.

3. Remove $1 \times$ PBS and fix inverted larvae in 1 mL of FA Fix Buffer for 10 min at room temperature on an orbital shaker (*see* **Note 5**).

4. Stop the crosslinking reaction. Remove the FA Fix Buffer and replace with 1 mL Quench Buffer. Incubate with shaking at room temperature for 5 min. Repeat wash, for a total of 10 min of Quench incubation.

5. Remove Quench Buffer, and replace with 1 mL of ice-cold Buffer A. Keep fixed inverted larvae on ice in Buffer A until all larvae have been dissected, fixed, and quenched.

6. Dissect wing discs off cuticle in Buffer A, working in small batches to keep solution cold. Transfer discs to 1.5 mL tube on ice containing Buffer A (*see* **Note 6**).

7. Once all discs have been dissected off cuticle, pellet discs at 1500 RCF for 5 min at 4 °C. Remove Buffer A (*see* **Note 7**).

*3.2 Chromatin Preparation*

1. Add 1 mL of FAIRE Lysis Buffer to discs. Transfer discs to screw-cap tube containing 2.38 mm tungsten beads. Homogenize in a bead beater for six cycles of 1 min on, 2 min off at 4 °C, power setting high (*see* **Note 8**).

2. Transfer the lysate to a fresh 15 mL tube. Add FAIRE Lysis Buffer to a total volume of 1.8 mL. Sonicate to a DNA fragment size of 100 bp to 2 kb (*see* **Note 9**). We use a Branson 450-digital Sonifier and perform five 30-s cycles, with a duty cycle of 1 s on, 0.5 s off at 18% amplitude. Keep tube on ice during sonication to prevent sample heating. Let tube rest on ice for at least 2 min between sonication cycles.

3. Verify sonicated DNA fragment size by transferring 50 μL of lysate to a fresh 1.5 mL tube. Add 50 μL FAIRE Lysis Buffer and incubate the sample at 65 °C overnight to reverse crosslinks. Then add RNaseA to a final concentration of 100 μg/mL and incubate for 30 min at 37 °C. Then add Proteinase-K to a final concentration of 200 μg/mL and incubate for 2 h at 37 °C. Purify the DNA by phenol-chloroform extraction and ethanol precipitation. Resuspend precipitated DNA in 25 μL
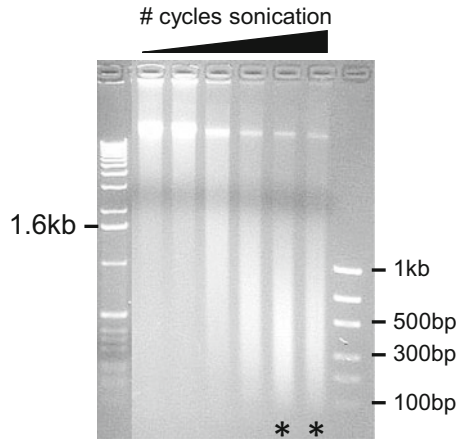
# cycles sonication



1.6kb —

— 1kb
— 500bp
— 300bp
— 100bp

* *

**Fig. 1** The effects of increasing sonication cycle number on DNA fragment size. Agarose gel stained with ethidium bromide. Lanes 1 and 8: DNA ladders. Lanes 2–7: A fixed amount of *Drosophila* S2 cell chromatin was sonicated for an increasing number of cycles. The DNA fragment sizes in Lanes 6 and 7 (asterisks) are suitable for FAIRE-seq

TE and run on a 1.2% agarose gel. The desired fragment size is 100 bp to 2 kb, with the average size being approximately 500 bp. If the fragment size is too big, additional rounds of sonication are necessary (Fig. 1).

4. Once the desired fragment size has been achieved, pellet insoluble chromatin at 15,000 RCF for 5 min at 4 °C. Transfer the supernatant to a fresh tube, leaving behind ~30 μL to avoid carryover of DNA in the pellet. Insoluble chromatin can be discarded.

**3.3 Preparation of FAIRE DNA**

1. Remove a 50 μL aliquot of soluble chromatin for use as an Input sample. Reverse crosslinks, treat with RNase and Proteinase K, and purify Input DNA as described in Subheading 3.2, **step 3**. Store Input DNA at −20 °C (*see* **Note 10**).

2. Perform FAIRE on the remaining soluble chromatin lysate. Transfer 500 μL aliquots of lysate to fresh 1.5 mL tubes. Add an equal volume of phenol-chloroform to the chromatin lysate. Vortex 10-s. Centrifuge at 15,000 RCF for 5 min at 4 °C. Transfer aqueous (upper) phase to fresh 1.5 mL tube. Avoid transferring any debris that may be trapped at the interface between aqueous and organic phases.

3. Perform a second phenol-chloroform extraction. Add an equal volume of phenol-chloroform to the FAIRE DNA-containing aqueous phase. Vortex 10-s. Centrifuge at 15,000 RCF for 5 min at 4 °C. Transfer the aqueous (upper) phase to a fresh

1.5 mL tube. If the interface is not clear and free of debris, perform a third phenol-chloroform extraction.

4. Perform a chloroform extraction. Add an equal volume of chloroform/isoamyl alcohol to each tube. Vortex 10-s. Centrifuge at 15,000 RCF for 5 min at 4 °C. Transfer the aqueous (upper) phase to a fresh tube. Pool aqueous phases from samples that were divided across multiple tubes.

5. Precipitate the FAIRE-enriched DNA. Add 1/10th volume of 3 M sodium acetate (pH 5.2) and mix well. Add 10–20 µg glycogen and mix well. Add 2-volumes of 95% ethanol and mix well by inverting the capped tubes three to four times. Incubate the tubes at −80 °C for at least 30 min (*see* **Note 11**).

6. Recover tubes from −80 °C and centrifuge at 15,000 RCF for 20 min at 4 °C. Remove the supernatant without disturbing the DNA pellet. Wash the pellet with 500 µL of cold 70% ethanol. Centrifuge at 15,000 RCF for 20 min at 4 °C. Remove all traces of ethanol and let pellet dry. Resuspend FAIRE-enriched DNA pellet in 100 µL TE.

7. Heat overnight at 65 °C to reverse any remaining crosslinks. Then add 100 µg/mL RNaseA. Incubate for 30 min at 37 °C.

8. Purify DNA with QIAquick PCR purification spin column. Add 500 µL Buffer PB. Mix well by pipetting. Transfer to column. Centrifuge at 10,000 RCF for 1 min at room temperature. Empty the collection tube. Add 750 µL Buffer PE to column. Incubate for 2 min at room temperature. Centrifuge at 10,000 RCF for 1 min. Empty the collection tube. Centrifuge at 10,000 RCF for 1 min. Transfer the column to a fresh 1.5 mL tube. Centrifuge at 15,000 RCF for 1 min to remove all residual ethanol. Transfer the column to a fresh 1.5 mL tube. Add 38 µL Buffer EB directly to silica at bottom of column (*see* **Note 12**). Incubate for 1 min at room temperature. Centrifuge at 10,000 RCF for 1 min at room temperature. The eluted buffer contains the FAIRE-enriched DNA.

9. Measure DNA concentration of FAIRE-enriched DNA and Input DNA using a fluorometer (e.g., QuBit High Sensitivity dsDNA kit). The quantity of FAIRE-enriched DNA should be approximately 1–10% of the total DNA in the original sample (*see* **Note 13**).

10. FAIRE-enriched DNA and Input/genomic DNA can be stored at −20 °C indefinitely, and prepared for single-end high-throughput sequencing using a variety of commercial kits. For *Drosophila*, 5–10 million aligned reads per replicate is often sufficient for subsequent bioinformatics analysis.

## 4 Notes

1. When possible, take advantage of behavioral milestones or distinct morphological features to improve precision of developmental staging. For example, wandering third instar larvae can be synchronized by clearing all larvae from the walls of the vial, waiting 2 h, and then collecting newly wandering larvae for dissection. Similarly, pupal stages can be synchronized collecting animals at the white prepupal stage, which lasts only 1 h at 25 °C.

2. When possible, process experimental and control samples in parallel to avoid batch effects, which are technical in origin, from being interpreted as having a biological basis. This is preferred to processing experimental samples separately from control samples. Batch effects can arise from multiple steps in the protocol, including culture conditions, chromatin preparation, FAIRE, and high-throughput sequencing library preparation.

3. Performing FAIRE-seq only on females simplifies subsequent bioinformatics analysis by equalizing the number of sex chromosomes and autosomes.

4. Depending on how quickly the dissections can be performed, it is preferred to work in small batches to minimize the time between dissection and fixation. Fixed larvae can be left in Buffer A on ice after quenching (Subheading 3.1, **step 5**) until all larvae have been dissected.

5. These conditions are designed for *Drosophila* imaginal discs. Other conditions may be required for different tissues, developmental stages, or insect species. The objective is to achieve a light crosslinking of proteins to DNA. Therefore, as a guide, we suggest using fixation conditions that have been established for performing immunofluorescence on the target sample, with the exception of decreasing the crosslinking time by half. For example, fixation of *Drosophila* embryos for immunofluorescence is typically performed in a mixture of heptane and fixation buffer for 20 min at room temperature. For FAIRE-seq on *Drosophila* embryos, we used the same mixture of heptane and fixation buffer for 10 min at room temperature [8].

6. Since open chromatin profiles are often cell type specific, it is important to remove extraneous tissues such as trachea. Finer-detailed dissections can also be performed at this time to increase cell type specificity of FAIRE-seq data.

7. Stopping point. Dissected discs can be frozen in Liquid Nitrogen and stored at −80 °C.

8. We have noticed that bead-beating the sample improves tissue dissociation and increases reproducibility between replicates. It is possible that other means of homogenization will sufficiently dissociate the tissue prior to sonication [5].

9. It is often necessary to optimize sonication conditions for each type of sample. For example, tissues with a high amount of cuticle proteins may require additional sonication to achieve the desired DNA fragment size. Similarly, different ages, models, or brands of sonicator may vary in their performance. We suggest performing pilot experiments to empirically determine the optimal sonication conditions for a given sample, as well as the reproducibility of these results.

10. As an alternative to Input DNA, genomic DNA can be prepared from each genotype used in the experiment and sonicated to a fragment size of 100 bp to 2 kb. Sequencing Input or genomic DNA sample is essential for interpreting FAIRE-seq data because it controls for genomic features such as repetitive DNA and copy number variations.

11. Stopping point. DNA/ethanol can be stored at −80 °C indefinitely.

12. Water can be used instead of Buffer EB to elute the DNA from the column. Using water allows for the DNA to be concentrated in a speed-vac if necessary for high-throughput sequencing library preparation.

13. FAIRE-enriched DNA that is substantially greater than 10% may be indicative of samples that have been under-crosslinked. Suboptimal crosslinking of proteins to DNA may result in experimental failure due to poor signal to noise. To quantify FAIRE enrichment, use the Input sample to calculate the amount of DNA in the original sample and compare this amount to the quantity of DNA recovered in the FAIRE sample. For example, if 50 μL of soluble chromatin lysate was used for the Input sample, and 500 μL of soluble chromatin lysate was used for the FAIRE sample, the amount of DNA in the Input sample represents 1/10th of the total DNA that was present in the FAIRE sample. If the Input sample yielded 10 μg of DNA, and the FAIRE sample yielded 1 μg of DNA, then the FAIRE enrichment would be 1% of the total DNA in the original sample.

## References

1. Giresi PG et al (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res 17(6):877–885

2. Giresi PG, Lieb JD (2009) Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). Methods 48 (3):233–239

3. Gaulton KJ et al (2010) A map of open chromatin in human pancreatic islets. Nat Genet 42 (3):255–259

4. Song L et al (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome Res 21(10):1757–1767

5. Simon JM et al (2012) Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. Nat Protoc 7(2):256–267

6. Simon JM et al (2013) A detailed protocol for formaldehyde-assisted isolation of regulatory elements (FAIRE). Curr Protoc Mol Biol Chapter 21:Unit21.26

7. Mueller B et al (2017) Widespread changes in nucleosome accessibility without changes in nucleosome occupancy during a rapid transcriptional induction. Genes Dev 31 (5):451–462

8. McKay DJ, Lieb JD (2013) A common set of DNA regulatory elements shapes Drosophila appendages. Dev Cell 27(3):306–318

9. Uyehara CM et al (2017) Hormone-dependent control of developmental timing through regulation of chromatin accessibility. Genes Dev 31(9):862–875

10. Pearson JC et al (2016) Chromatin profiling of Drosophila CNS subpopulations identifies active transcriptional enhancers. Development 143(20):3723–3732

11. Davie K et al (2015) Discovery of transcription factors and regulatory regions driving in vivo tumor development by ATAC-seq and FAIRE-seq open chromatin profiling. PLoS Genet 11 (2):e1004994

12. Behura SK et al (2016) High-throughput cis-regulatory element discovery in the vector mosquito Aedes aegypti. BMC Genomics 17:341

13. Richard G et al (2017) Dosage compensation and sex-specific epigenetic landscape of the X chromosome in the pea aphid. Epigenetics Chromatin 10:30

14. Zhang Q et al (2017) Genome-wide open chromatin regions and their effects on the regulation of silk protein genes in Bombyx mori. Sci Rep 7(1):12919

15. Lai YT, Deem KD, Borràs-Castells F, Sambrani N, Rudolf H, Suryamohan K, El-Sherif E, Halfon MS, McKay DJ, Tomoyasu Y (2018) Enhancer identification and activity evaluation in the red flour beetle, *Tribolium castaneum*. Development 145(7). https://doi.org/10. 1242/dev.160663 PMID: 29540499