Chapter 11

Profiling Protein–DNA Interactions Cell-Type-Specifically with Targeted DamID

Owen J. Marshall and Caroline Delandre

Abstract

Targeted DamID (TaDa) is a means of profiling the binding of any DNA-associated protein cell-type specifically, including transcription factors, RNA polymerase, and chromatin-modifying proteins. The technique is highly sensitive, highly reproducible, requires no mechanical disruption, cell isolation or antibody purification, and can be performed by anyone with basic molecular biology knowledge. Here, we describe the TaDa method and downstream bioinformatics data processing.

Key words DamID, Targeted DamID, Gene transcription, Transcription factors, Protein-DNA interactions

1 Introduction

Targeted DamID (TaDa) [1, 2] is a recently developed method for profiling DNA–protein interactions cell-type specifically. The technique is based on DamID, a means of marking protein–DNA interactions by fusing a bacterial DNA adenine methylase (Dam) to any protein of interest [3, 4]. Adenine methylation—common in prokaryotes—is extremely sparse in eukaryotes [5], and Dam-fusion proteins leave enriched methylation at GATC sequences in close proximity to where they bind. Following genomic DNA isolation, methylated GATC sites are cut with methylation-sensitive restriction enzymes, enriched via ligation-mediated PCR and sequenced using next-generation sequencing. Resulting binding profiles can be used to identify transcription factor binding sites (e.g., [6]), gene transcription (using the proxy of RNA polymerase occupancy, e.g., [1]), chromatin state profiling (e.g., [7]) and chromatin accessibility profiling (e.g., [8]) in a semiquantitative manner. Notably, DamID can be used to profile the association of proteins that do not bind DNA directly (such as

Julia Horsfield and Judith Marsman (eds.), *Chromatin: Methods and Protocols*,
Methods in Molecular Biology, vol. 2458, https://doi.org/10.1007/978-1-0716-2140-0_11,

Lamin B [9]) as well as transcription factors that lack DNA-binding domains (O. Marshall, unpublished data) and noncoding RNAs [10].

As high levels of adenine methylation at GATC sites are toxic in eukaryotes [1], TaDa uses a bicistronic transcript and the phenomenon of spontaneous ribosomal reinitiation [11] to massively reduce the translation rates of Dam-fusion proteins [1]. The result is a DamID system that can be driven cell type–specifically by inducible bipartite expression systems such as GAL4/UAS [12]. Methylation marks are laid down gradually over the period of induction (which can be a period from as short as 16–48 h or longer) prior to tissue harvesting, and represent a form of profiling free of the potential artefacts associated with chemical fixation or cellular disruption. Importantly, the extremely low translation rates of Dam-fusion proteins mean that the concentration of these proteins are at effectively undetectable levels, and do not induce cell-fate changes or cause over-expression phenotypes. Despite this, the technique is highly sensitive, generating reproducible binding profiles from as few as 10,000 total driven cells.

Dam is a promiscuous adenine methylase, and will methylate all open chromatin regions in eukaryotes [13]. Dam-fusion proteins display similar behavior, and their methylation signature thus represents a combination of open chromatin methylation together with enrichment of methylation around specific binding sites of the protein of interest [13]. In order to account for this, all TaDa experiments with Dam-fusion proteins are carried out alongside a Dam-only control, and specific signal isolation is carried out in software [13]. An additional advantage of the methylation of open chromatin is that Dam-only TaDa experiments provide a simple and easy means to profile chromatin accessibility cell-type specifically (a feature that has been termed Chromatin Accessibility TaDa, or CATaDa) [8].

In *Drosophila melanogaster,* TaDa profiling is generally performed via a genetic cross between a GAL4 driver line and both an inducible Dam-fusion protein and a Dam-only control. A separate line is required to profile each DNA-associated protein, and preexisting lines are available for RNA polymerase profiling and chromatin protein profiling. New TaDa lines can be created by cloning and transgenesis using the pUAST-attB-LT3-NDam [1] or pTaDaG2 [14] base vectors. Alternatively, the FlyORF-TaDa system allows for the conversion of the extensive FlyORF library of lines to TaDa lines via a genetic cross [15].

Key to the success of TaDa is its ease of use. Cell-type-specificity is generated via genetic drivers rather than through mechanical disruption and sorting, and no antibodies are required. The technique is extraordinarily resilient and reliable, requiring only the most basic experience with genomic DNA isolation, restriction enzyme digestion and the PCR to generate highly reproducible

binding profiles. The procedure rarely if ever fails in our research group, and can be performed by undergraduate students with little previous laboratory experience.

Although this chapter has been written with experiments performed in *D. melanogaster* in mind, the technique has been adapted to both mammalian systems [16, 17] and recently to *Caenorhabditis elegans* [18], and, except for the initial tissue isolation steps, the following protocol can also be used on material generated from these systems.

2 Materials

2.1 Tissue Preparation

1. For fly head isolation, three woven wire test sieves of 150, 425, and 710 μm aperture size, respectively.
2. Phosphate Buffered Saline (PBS): place PBS tablet into water according to manufacturer's recommendations and autoclave.
3. 0.5 M Ethylenediaminetetraacetic acid (EDTA).
4. 1.5 mL homogenizing pestle: wipe and store in 100% ethanol after every use.

2.2 DNA Extraction

1. Quick-DNA Miniprep Kit (Zymo Research) or similar kit.
2. Agarose.
3. Tris Acetate EDTA (TAE) buffer.
4. SYBR™ Safe DNA Gel Stain (Thermo Fisher Scientific) or ethidium bromide.

2.3 DpnI Digestion

1. DpnI (20,000 units/mL).
2. CutSmart® Buffer (New England Biolabs), or buffer compatible with DpnI.
3. NucleoSpin Gel and PCR Cleanup Kit (Macherey-Nagel) or similar kit.

2.4 Adaptor Ligation

1. 50 μM dsAdR: combine 50 μL AdRt (100 μM) and 50 μL AdRb (100 μM) (*see* Table 1), incubate at 95 °C for 2 min in a removable heat block, and remove the heating block and allow to cool to room temperature (at least 45 min).
2. Adaptor Ligation Buffer: 2 μL of 10× T4 Ligase Reaction Buffer, 0.8 μL of dsAdR (*see* Subheading 2.4, **item 1**), 1.2 μL H_2O. Make a 100× batch and store at −20 °C in 40 μL aliquots.
3. T4 DNA Ligase (400,000 units/mL).

Table 1
Targeted DamID adaptor and primer sequences (* = phosphorothioate linkage)

Name	Sequence
AdRt	CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGA
AdRb	TCCTCGGCCG
DamID-PCR	GGTCGCGGCCGAGGATC
NGS-PCR1	AATGATACGGCGACCACCGA*G
NGS-PCR2	CAAGCAGAAGACGGCATACGA*G

2.5 DpnII Digestion

1. DpnII Digestion Buffer: 4 μL DpnII Buffer, 15 μL H_2O. Make a 100× batch and store at −20 °C in 190 μL aliquots.
2. DpnII (10,000 units/mL).

2.6 PCR Amplification

1. 50 μM DamID PCR primer (*see* Table 1).
2. PCR Buffer: 32 μL 5× MyTaq HS Buffer (Meridian Bioscience), 2.5 μL DamID-PCR primer, 83.5 μL H_2O. Make a 50× batch and store at −20 °C in 1180 μL aliquots.
3. MyTaq™ HS DNA Polymerase (Meridian Bioscience), or similar polymerase and buffer [2].
4. NucleoSpin Gel and PCR Cleanup Kit (Macherey-Nagel) or similar kit.
5. Spectrophotometer (e.g., NanoDrop™).

2.7 Sonication and Removal of DamID Adaptors

1. Sonicator.
2. 1.5 mL tubes compatible with sonicator (we use the Bioruptor® Plus sonicator and TPX microtubes from Diagenode).
3. DNA analyzer for sizing and quality control of DNA samples (e.g., TapeStation System; Agilent).
4. Genomic DNA ScreenTape and Reagents (Agilent) or equivalent for other system.
5. AlwI (10,000 units/mL).
6. CutSmart® Buffer (New England Biolabs) or buffer compatible with AlwI.

2.8 Sequencing Library Preparation

2.8.1 DNA Cleanup

1. Sera-Mag Speedbeads carboxyl magnetic beads.
2. Dilute Sera-Mag beads in polyethylene glycol (PEG) solution: in a 15 mL tube, make a solution of 3 g PEG 8000, 3 mL of 5 M NaCl, 150 μL of 1 M Tris−HCl pH 8.0, 30 μL of 0.5 M EDTA. Add H_2O to 14 mL and mix by inversion until PEG is dissolved. Mix Sera-Mag Speedbeads container well to ensure beads are in suspension and add 300 μL to the PEG solution. Mix by inversion and fill with H_2O up to 15 mL. Store at 4 °C.

3. Magnetic stand. To make a homemade magnetic stand (*see* **Note 1**): 75 × 10 × 3 mm N42 Neodymium magnet (e.g., FIRST4MAGNETS®; cat. no. F75103-2), 20 µL filter tip box insert, and 75 × 10 × 2.5 mm plastic spacer (supplied with a pair of the strip magnets). Magnetic stands can also be purchased (e.g., NEBNext® Magnetic Separation Rack, New England Biolabs).
4. Resuspension Buffer: 10 mM Tris−HCl pH 8.0, 0.1 mM EDTA. Make a 15 mL solution and store at −20 °C in 1 mL aliquots.

2.8.2 Concentration Adjustment

1. Qubit™ Fluorometer (Thermo Fisher Scientific) or similar instrument.
2. Qubit™ dsDNA HS Assay Kit (Thermo Fisher Scientific).

2.8.3 End Repair

1. End Repair Enzyme mix: 1.14 µL T4 DNA Polymerase (3000 units/mL), 0.23 µL DNA Polymerase I, Large (Klenow) Fragment (5000 units/mL), 1.14 µL T4 Polynucleotide Kinase (10,000 units/mL). Make a 50× batch and store at −20 °C.
2. End Repair Buffer: 3 µL 10× T4 Ligase Reaction Buffer, 1.2 µL 10 mM dNTPs, 3.3 µL H_2O. Make a 50× batch and store at −20 °C in 35 µL aliquots.

2.8.4 Adenylation of 3′ Ends and Adaptor Ligation

1. Klenow Fragment 3′ to 5′ exo- (5000 units/mL).
2. Quick Ligation™ Kit (New England Biolabs) or similar kit.
3. Sequencing adaptors (*see* Table 2): resuspend adaptors at 100 µM in Tris-EDTA buffer solution with 50 mM NaCl. Mix 25 µL adaptor index and 25 µL universal primer. Incubate at 95 °C for 2 min in a removable heat block, and remove the heating block and allow to cool to room temperature (it will take at least 45 min).

2.8.5 DNA Cleanup

Same reagents as Subheading 2.8.1.

2.8.6 DNA Fragment Enrichment

1. NEBNext Ultra II Q5 Master Mix (New England Biolabs), or other High Fidelity DNA Polymerase, optimized for amplification of NGS libraries.
2. PCR Primer Cocktail: mix 25 µL PCR1 primer, 25 µL PCR2 primer (*see* Table 1), 50 µL H_2O. Store at −20 °C.

2.8.7 DNA Cleanup

Same reagents as Subheading 2.8.1.

2.8.8 Library Quality Control

Same reagents as Subheading 2.7, **items 3** and **4**, and Subheading 2.8.2.

Table 2
NGS adaptor sequences ([Phos] = 5′ phosphorylation, * = phosphorothioate linkage)

Name	Barcode	Sequence
Universal	N/A	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCT TCCGATC*T
Index 1	ATCAGG	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCCGTCTTCTGCTT*G
Index 2	CGATGT	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACCGATGTATCTCGTATGCCGTCTTCTGCTT*G
Index 3	TTAGGC	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACTTAGGCATCTCGTATGCCGTCTTCTGCTT*G
Index 4	TGACCA	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACTGACCAATCTCGTATGCCGTCTTCTGCTT*G
Index 5	ACAGTG	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACACAGTGATCTCGTATGCCGTCTTCTGCTT*G
Index 6	GCCAAT	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACGCCAATATCTCGTATGCCGTCTTCTGCTT*G
Index 7	CAGATC	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACCAGATCATCTCGTATGCCGTCTTCTGCTT*G
Index 8	ACTTGA	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACACTTGAATCTCGTATGCCGTCTTCTGCTT*G
Index 9	GATCAG	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACGATCAGATCTCGTATGCCGTCTTCTGCTT*G
Index 10	TAGCTT	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACTAGCTTATCTCGTATGCCGTCTTCTGCTT*G
Index 11	GGCTAC	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACGGCTACATCTCGTATGCCGTCTTCTGCTT*G
Index 12	CTTGTA	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACCTTGTAATCTCGTATGCCGTCTTCTGCTT*G
Index 13	AGTCAA	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACAGTCAACAATCTCGTATGCCGTCTTCTGCTT*G
Index 14	AGTTCC	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACAGTTCCGTATCTCGTATGCCGTCTTCTGCTT*G
Index 15	ATGTCA	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACATGTCAGAATCTCGTATGCCGTCTTCTGCTT*G
Index 16	CCGTCC	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACCCGTCCCGATCTCGTATGCCGTCTTCTGCTT*G
Index 18	GTCCGC	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACGTCCGCACATCTCGTATGCCGTCTTCTGCTT*G
Index 19	GTGAAA	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCACGTGAAACGATCTCGTATGCCGTCTTCTGCTT*G

3 Methods

As TaDa uses PCR amplification from small amounts of starting material, avoiding cross-contamination is crucial. From Subheading 3.1 onward, wear gloves, use filter tips, and use small aliquots of autoclaved MilliQ (or nuclease-free) H_2O to avoid contamination.

3.1 Tissue Preparation

1. Set up an appropriate cross to induce the TaDa system with a tissue-specific GAL4 driver, together with tub-GAL80ts (in flies); or with a Cre-specific driver (in mammals). Consider the induction time, the tissue and the cell-type being profiled. (*See* **Note 2**).
2. (If using flies) grow the progeny at 18 °C until induction is required.
3. (If using flies) induce by shifting progeny to 29 °C for 16–48 h (*see* **Note 3**).
4. After inducing the TaDa system, collect enough material for at least 10,000 driven cells. For isolated tissues, dissect in PBS; for embryos, wash in PBS; and in both cases transfer to a 1.5 μL tube, gently pellet and remove supernatant. For adult flies, transfer into 15 μL conical tubes cooled in dry ice. Samples may be stored at −20 or −80 °C at this point (*see* **Note 4**). If isolating adult fly heads, place the three sieves on top of each other with the largest aperture size sieve on top and the smallest at the bottom. Place the tube of frozen flies on dry ice, vortex, and add the flies to the top sieve. Shake or tap the sieves vigorously and check the middle sieve; it should contain the fly heads (*see* **Note 5**). Fly heads can be stored in 1.5 μL tubes at −20 °C until ready to use.
5. Take samples in 1.5 μL tubes from the freezer, and add 75 μL H_2O and 20 μL 0.5 M EDTA.
6. If working with embryos or heads, homogenize samples with a pestle by inserting it into the tube, twisting and pushing down with the pestle ~20 times, and progress immediately to the next step.
7. If working with dissected tissues, progress immediately to the next step.

3.2 DNA Extraction

All the reagents for this section are included in the Quick-DNA Miniprep Plus Kit (Zymo Research). An equivalent DNA preparation kit can be used. Be very gentle with samples prior to DpnI digestion (Subheading 3.3, **step 2**), as any genomic DNA shearing will result in broken ends that can potentially ligate DamID adaptors. Do not vortex but mix samples gently by inversion and/or very gentle flicking.

1. Prior to Subheading 3.1, **step 2**, make a Prot K Master Mix by combining 10 μL Proteinase K and 95 μL Solid Tissue Buffer (blue buffer) per sample, and vortexing to mix well.
2. Add the Prot K Master Mix to the sample and flick gently to mix.
3. Digest at 56 °C for 1–3 h in a heat block (*see* **Note 6**).
4. Cool samples to room temperature, add 400 μL Genomic Binding Buffer, and mix by gentle inversion.
5. Add samples to spin columns (*see* **Note 7**) on a vacuum manifold (*see* **Note 8**) and draw through.
6. Add 400 μL DNA Pre-Wash Buffer and draw through.
7. Add 700 μL gDNA Wash Buffer and draw through.
8. Add 200 μL gDNA Wash Buffer and draw through.
9. Transfer spin column to a collection tube and spin at maximum speed for 2 min.
10. Transfer spin column to a new 1.5 mL tube, add 50 μL DNA Elution Buffer, and leave for at least 30 min.
11. Centrifuge at >6000 × *g* for 1 min to elute the DNA.
12. Optional: run 1 μL on a 0.8% agarose get to check quality (*see* **Note 9**).

3.3 DpnI Digestion

All the reagents for the cleanup (**step 3**) are included in the NucleoSpin Gel and PCR Cleanup Kit (Macherey-Nagel). An equivalent PCR clean up kit can be used.

1. Transfer 43.5 μL of the sample to a new 1.5 mL tube.
2. Add 5 μL CutSmart Buffer and 1.5 μL DpnI, mix very gently, and digest at 37 °C overnight (*see* **Note 10**).
3. Add 100 μL Buffer NTI to the sample, transfer to a spin column, and draw through with a vacuum manifold.
4. Add 700 μL Buffer NT3 and draw through.
5. Repeat **step 4**.
6. Transfer spin column to a collection tube and centrifuge at 11,000 × *g* for 1 min.
7. Transfer spin column onto a new 1.5 mL tube, add 32 μL H_2O, incubate for 1 min at room temperature, and centrifuge at 11,000 × *g* for 1 min to elute DNA.

3.4 Adaptor Ligation

1. Transfer 15 μL of each sample into a PCR tube (*see* **Note 11**).
2. Store remaining sample at −20 °C or lower for future use if required (*see* **Note 12**).

3. Add 4 μL Adaptor Ligation Buffer and 1 μL T4 DNA Ligase. Mix well.
4. Incubate for 2 h at 16 °C followed by 10 min at 65 °C to heat-inactivate (*see* **Note 13**).

3.5 DpnII Digestion

1. Add 19 μL DpnII Digestion Buffer and 1 μL DpnII. Mix well.
2. Incubate for 2 h at 37 °C followed by 20 min at 65 °C (*see* **Note 13**).

3.6 PCR Amplification

1. Add 118 μL PCR Buffer and 2 μL MyTaq™ HS DNA Polymerase. Mix well.
2. Split into 4 × 40 μL reactions in a PCR strip (*see* **Note 14**).
3. Run the following PCR program (*see* **Note 15**):

1×	68 °C	10 min
1×	94 °C	30 s
1×	65 °C	5 min
1×	68 °C	15 min
3×	94 °C 65 °C 68 °C	30 s 1 min 10 min
17×	94 °C 65 °C 68 °C	30 s 1 min 2 min
1×	68 °C	5 min

4. Clean up using the NucleoSpin Gel and PCR Cleanup Kit (*see* Subheading 3.3). Use 320 μL of Buffer NTI and wash column with Buffer NT3 three times (instead of two) if using MyTaq™ HS DNA Polymerase. Elute by adding 32 μL H_2O to the spin column and leaving for at least 5 min before centrifuging.
5. Measure DNA concentration on a Qubit or similar high sensitivity instrument (*see* **Note 16**).
6. Run 1 μL on a 1% agarose gel (or on a Genomic DNA ScreenTape using the TapeStation System) to check for amplification of a broad 400 bp–2 kb range of fragments (Fig. 1a).

3.7 Sonication and Removal of DamID Adaptors

1. Dilute 2 μg of sample in 90 μL H_2O in 1.5 mL Bioruptor® Plus TPX microtubes (*see* **Note 17**).
2. Add 10 μL CutSmart Buffer and mix well. Cool on ice.
3. Sonicate in Bioruptor® Plus: add 500 mL ice to the chamber and top up to the designated mark with water, cool tube holder on ice, sonicate for 5 cycles of 27 s on–27 s off on high power (*see* **Note 18**).

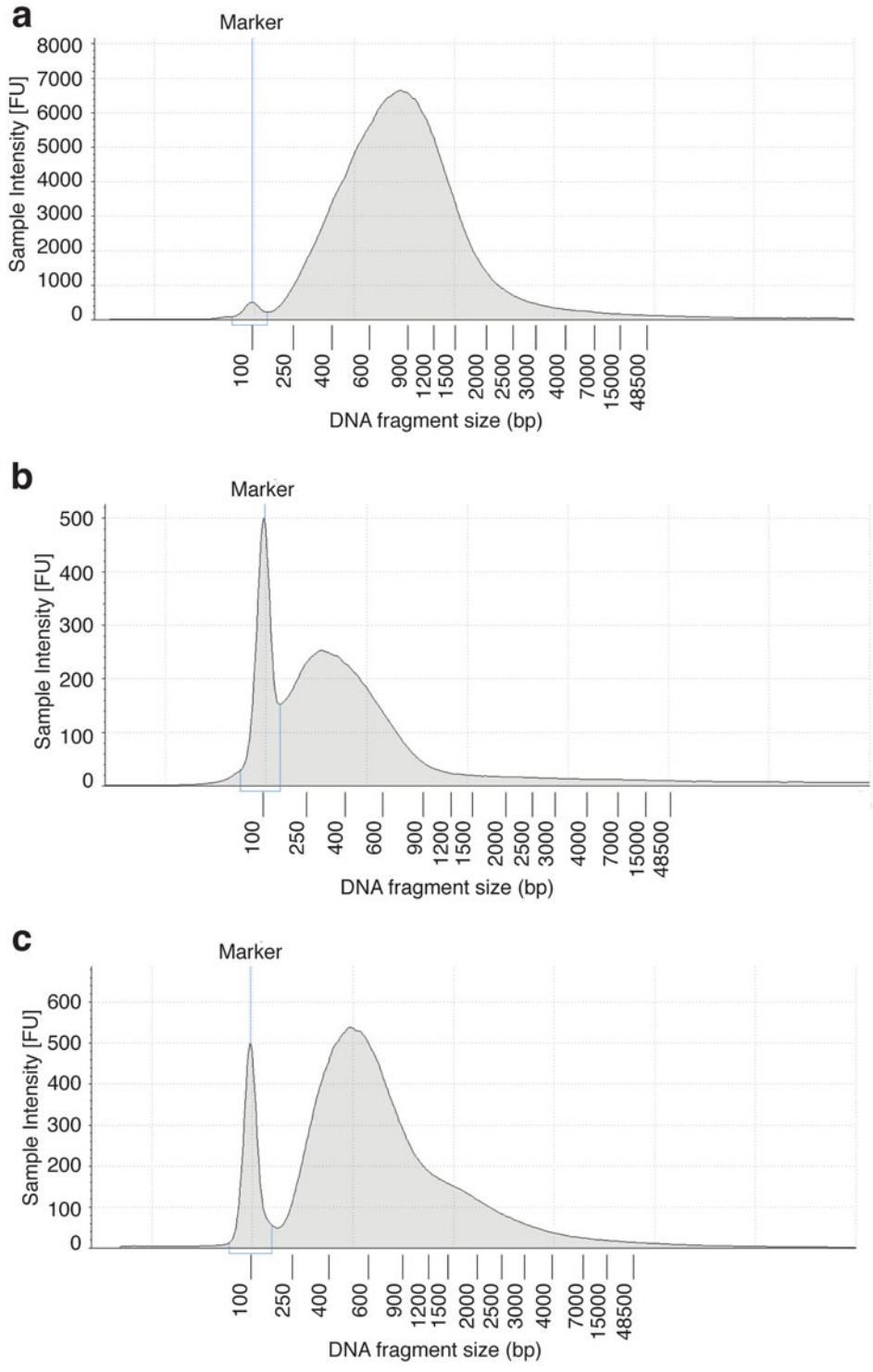

Fig. 1 Profiles of DNA fragments created by the TaDa technique. Example TapeStation plots showing the DNA size and quality for the same sample (**a**) before sonication (Subheading 3.6, **step 5**), (**b**) after sonication (Subheading 3.7, **step 4**), and (**c**) at the final library quality check (Subheading 3.8.9, **step 1**)

4. Check fragment size is ~300 bp (Fig. 1b) with a Genomic DNA ScreenTape using the TapeStation System, or equivalent (*see* **Note 19**).
5. Add 1 μL AlwI and digest at 37 °C overnight to remove DamID adaptors (*see* **Note 20**).

3.8 Sequencing Library Preparation

3.8.1 DNA Cleanup

1. Transfer 70 μL (or 90 μL, if concentration prior to sonication was low) of sample into a PCR tube and add 105 μL (or 135 to 90 μL) Sera-Mag beads. Mix well (*see* **Note 21**).
2. Incubate at room temperature for 10 min.
3. Place on magnetic stand for 10 min (or until the supernatant is clear).
4. Remove supernatant and wash twice with 190 μL 80% ethanol, 30 s each time.
5. Wait 5 min for the beads to air-dry.
6. Resuspend in 25 μL Resuspension Buffer and remove from magnetic stand. Mix well and incubate at room temperature for 2 min.
7. Place on magnetic stand for 5 min (or until the supernatant is clear).
8. Transfer 22.5 μL of the supernatant into a new PCR tube.

3.8.2 Concentration Adjustment

1. Use 1 μL of sample to measure DNA concentration using the Qubit™ dsDNA HS Assay Kit (or kit for equivalent instrument) following manufacturer's instructions.
2. Dilute samples to (no more than) 500 μg of DNA in 20 μL Resuspension Buffer in a PCR tube.

3.8.3 End Repair

1. Add 7.5 μL End Repair Buffer and 2.5 μL End Repair Enzyme mix. Mix well.
2. Incubate for 30 min at 30 °C followed by 20 min at 37 °C.

3.8.4 Adenylation of 3′ Ends

1. Add 0.75 μL Klenow Fragment 3′ to 5′ exo- and mix well.
2. Incubate for 30 min at 37 °C and proceed immediately to the adaptor ligation.

3.8.5 Adaptor Ligation

1. Add 2.5 μL Quick Ligase and 2.5 μL of the relevant adaptor (*see* **Note 22**).
2. Incubate for 10 min at 30 °C.
3. Add 5 μL 0.5 M EDTA.

3.8.6 DNA Cleanup

1. Add 40 μL Sera-Mag beads and mix well.
2. Incubate at room temperature for 10 min.

3. Place on magnetic stand for 5 min (or until the supernatant is clear).
4. Remove supernatant and wash twice with 190 µL 80% ethanol, 30 s each time.
5. Wait 5 min for the beads to air-dry.
6. Resuspend in 52.5 µL Resuspension Buffer and remove from magnetic stand. Mix well and incubate at room temperature for 2 min.
7. Place on magnetic stand for 5 min (or until the supernatant is clear).
8. Transfer 50 µL of the supernatant into a new PCR tube.
9. Add 50 µL Sera-Mag beads and mix well to start the second round of cleanup (*see* **Note 23**).
10. Incubate at room temperature for 10 min.
11. Place on magnetic stand for 5 min (or until the supernatant is clear).
12. Remove supernatant and wash twice with 190 µL 80% ethanol, 30 s each time.
13. Wait 5 min for the beads to air-dry.
14. Resuspend in 22.5 µL Resuspension Buffer and remove from magnetic stand. Mix well and incubate at room temperature for 2 min.
15. Place on magnetic stand for 5 min (or until the supernatant is clear).
16. Transfer 20 µL of the supernatant into a new PCR tube.

3.8.7 DNA Fragment Enrichment

1. Add 5 µL PCR Primer Cocktail and 25 µL NEBNext Ultra II Q5 Master Mix. Mix well.
2. Run the following PCR program:

1×	**98 °C**	**30 s**
6×	98 °C 60 °C 72 °C	10 s 30 s 30 s
1×	72 °C	5 min

3.8.8 DNA Cleanup

1. Add 50 µL Sera-Mag beads and mix well.
2. Incubate at room temperature for 10 min.
3. Place on magnetic stand for 5 min (or until the supernatant is clear).
4. Remove supernatant and wash twice with 190 µL 80% ethanol, 30 s each time.

5. Wait 5 min for the beads to air-dry.
6. Resuspend in 32.5 μL Resuspension Buffer and remove from magnetic stand. Mix well and incubate at room temperature for 2 min.
7. Place on magnetic stand for 5 min (or until the supernatant is clear).
8. Transfer 30 μL of the supernatant into a new tube.

3.8.9 Library Quality Control

1. Check DNA quality (Fig. 1c) with a Genomic DNA ScreenTape using the TapeStation System (*see* **Note 24**).
2. Measure library concentration using the Qubit™ dsDNA HS Assay Kit following manufacturer's instructions.
3. Pool samples to give a final DNA concentration of 20 nM.
4. Run on a next-generation sequencer to obtain at least 20 million reads per index (for *Drosophila melanogaster* samples) or >50 million reads per index with mammalian samples. Both single-end (SE) and paired-end (PE) sequencing data can be processed; however, we recommend PE sequencing if available cost-effectively.

3.9 Processing of Sequencing Data

1. Download sequencing data in FASTQ format (compressed fastq.gz files are fine) to a single directory. If possible, naming the files such that the sample name is at the start of the filename will simplify downstream processing. Dam-only control samples should ideally start with "Dam" (*see* **Note 25**).
2. Download and install the damidseq_pipeline software [13], freely available online from https://owenjm.github.io/damidseq_pipeline. Detailed installation and usage instructions, along with a test dataset, are available from the website.
3. Run damidseq_pipeline in the directory (use the --paired flag if using paired-end sequencing) (*see* **Note 26**). The final outputs are normalized Log_2(Dam-fusion/Dam-alone) binding profiles in BEDGRAPH format.
4. The resulting binding profiles can be viewed using browser software such as IGV (Integrative Genomics Viewer) [19], or publication-quality binding profile figures can be generated using pyGenomeTracks [20].
5. If profiling transcription factor binding, significantly enriched peaks called on false discovery rate (FDR) may be identified using find_peaks, freely available from https://github.com/owenjm/find_peaks. Peaks are outputted in GFF format, and can be visualized using the software described in Subheading 3.9, **step 4**, or used for subsequent processing and analysis.

6. If profiling gene expression through RNA polymerase occupancy, genes with significantly enriched polymerase occupancy (a proxy for gene expression) can be called from the RNA Polymerase II ratio files using the polii.gene.call Rscript, freely available from https://github.com/owenjm/polii.gene.call.
7. If comparing gene expression from different conditions, we recommend the use of NOISeq [21] on antilog-transformed occupancy data generated by polii.gene.call. An example of this workflow can be found in [22].

4 Notes

1. A homemade magnetic stand may be built by taping together the magnet with the plastic spacer under a 20 μL filter tip box insert (acting as a PCR strip rack) so the magnet will be near the tip of the PCR tubes once inserted in the stand. Use tape on the side of the PCR strip to ensure its edge remains firmly in contact with the magnet. Strength N42 works well, but there should be no harm in using more powerful magnets (resulting in a faster purification process).
2. We strongly recommend testing driver specificity under experimental conditions before proceeding with TaDa. If using the original TaDa constructs, we recommend testing with a UAS-inducible marker (e.g., membrane-bound GFP or nuclear RFP) inserted in the same targeted insertion site as the TaDa construct. If using the TaDaG2 vectors [14] induced cells are marked with membrane-bound GFP during TaDa induction, and a proportion of the experimental collection can be sacrificed for microscopy during dissection.
3. The choice of induction depends on the characteristics of the cell type being profiled. We typically use 16 h for embryonic neuroblasts, 24 h for larval cells and 48 h for adult neurons. In general, longer induction times increase the signal–noise ratio, but may be inappropriate for profiling specific stages in developmentally important cell types. Profiling for <12 h is not recommended.
4. Some degree of tissue dissection is recommended in most instances, although it is possible to use whole animals with highly specific GAL4 drivers ([13] and J. Newland, C. Delandre and O. Marshall, unpublished data). We do recommend caution and careful scrutiny of both the driver expression and the resulting data if not dissecting tissues, however. Note, in particular, that many GAL4 lines drive expression in the salivary glands and these will contribute significantly to the resulting binding profiles if not removed.

5. Check the top sieve under the microscope to assess how many fly heads actually separated from the rest of the body. If it is lower than 90%, try vortexing longer (we usually vortex three times for 10 s each and cool them on dry ice in between). We get the highest rates of fly head removal when vortexing while moving the conical tube so the sides are well shaken.
6. Check under the microscope that the tissue has been completely digested; only tissues such as cuticle and mouth hooks should be visible.
7. When working with fly heads, try to reduce the number of heads being transferred onto the column to prevent blockage.
8. If using a vacuum manifold, use disposable connectors (e.g., VasConnectors; QIAGEN) to avoid cross-contamination. DamID is an extremely sensitive technique that involves the PCR amplification of material, and we have observed significant signal contamination (especially from bacterial plasmids) when not using disposable connectors. If not using a vacuum manifold, centrifugation steps can be used as per the manufacturer's instructions without issue; however, we strongly recommend the use of a manifold for time-saving and ease-of-use considerations.
9. There should be a single band on the top of the gel and not a smear. Although it is sensible to perform this check the first time TaDa is carried out (or if using difficult tissues such as the gut) in general we rarely observe significant DNA degradation at this step.
10. The digestion can be reduced to 2 h if required (with some potential loss of sensitivity). DpnI can also be optionally heat-inactivated at 80 °C for 20 min, although it is effectively removed in the subsequent DNA cleanup step.
11. We recommend the use of 8-well PCR strips for medium throughput applications, and use fresh strip caps after every opening to avoid contamination. There are generally very little or undetectable amounts of DNA at this point, as uncut genomic DNA (which should be the majority of DNA in most use cases) will not pass through the spin column. The purified, cut DNA should only come from induced cells. If working with large amounts of tissue and/or if concerned about DNA shearing, the sample concentration can be measured and a maximum of 750 ng should be used going forward. However, such a high yield is unusual in practice. Note that undetectable yields at this stage can produce excellent data.
12. If the yield at Subheading 3.6, **step 5** is low (e.g., <1 μg) it is possible to repeat the processing on this stored aliquot and merge the two repeats at Subheading 3.7, **step 1**. We rarely find that this is necessary. Otherwise, we recommend retaining this aliquot for verification/backup purposes.

13. Unless indicated, we typically use a PCR machine for the enzyme incubations.
14. Splitting the reaction into 40 μL aliquots provides optimum reaction efficiency—we find that final yields are lower if the reaction is not split.
15. Both the initial long extension for 10 min and the subsequent long cycles are required. A 3 min extension time, as present in the final 17 cycles, should amplify all fragments below 5 kb; these represent 99.97% of all GATC fragments in the *Drosophila melanogaster* genome.
16. We find that trace carryover amounts of the MyTaq PCR buffer can significantly affect quantitation via a spectrophotometer, and we recommend using a DNA-chelating-dye-based quantitation method (such as a Qubit), especially if yields are expected to be low. Note that the yield of DNA at this point will greatly depend on both the starting number of cells and the extent of DNA binding exhibited by the protein in question. We have seen extremely low amplification values (e.g., <20 ng total yield) at this stage generate excellent quality final binding profiles.

    In terms of total expected yields, broad chromatin binding proteins (e.g., Polycomb, Brm, Lamin) can generate 2× to 10× the amount of DNA amplification observed in the Dam-only control. DNA polymerase components (e.g., RpII215 or RpII18) typically generate 0.5× to 1×, respectively, the amount of DNA amplification observed in the Dam-only control. The yield from transcription factors greatly depends on the specificity of binding.
17. If the yield from Subheading 3.6, **step 5** is <2 μg, use the total amount of sample at this point.
18. All sonicators are different and conditions should be optimized for the specific device. Note that sonicators typically provide custom-made sample tubes, and these must be used for reproducible fragmentation (the variation when using standard 1.5 mL tubes is unusably large in our experience). If possible, we recommend the use of modern sonicators with inbuilt water cooling—if using these, no ice is required, but prechilling of the unit is necessary.
19. Slight variations in fragment size are acceptable. Very large fragments (>600 bp) may impede clustering efficiency and sequencing yields, and we aim to avoid this (although we have never observed a sequencing library to fail in practice). It is acceptable to perform additional cycles of sonication if the fragment size is deemed too large at this step.
20. AlwI is a Type IIS restriction enzyme that cuts four nucleotides downstream from a GGATC site. Using this enzyme avoids the initial low-diversity of most library fragments commencing

with GATC, and is important for optimal cluster detection efficiency on many sequencing platforms. However, the use of ordered flow cells may alleviate this problem—and in these cases the enzyme Sau3A1 may be used instead to retain the full length of the GATC fragment. Other than reduced read numbers when retaining the initial GATCs, we do not observe any significant difference in TaDa profiles when cutting with either enzyme.

21. Mix by pipetting ~20 times, flick the PCR strip, and pulse down. Rapid mixing of beads and sample is important for even DNA precipitation.
22. If multiplexing four or fewer libraries, selecting adaptors with barcodes that are too similar may result in a reduced number of reads passing the filter. In this case, the preferred indexes to use (in order) are 4, 7, 6, and 8 (i.e., if multiplexing two samples, use indices 4 and 7). When using more than four libraries, we simply use library indices in numerical order and have never observed any issues with reduced read numbers on multiple NGS platforms.
23. This second cleanup step is required to ensure complete removal of sequencing adaptor dimers; should these be present in any library they will out-compete genuine library sequences in the multiplex when hybridizing to the flow cell.
24. The peak should have an average size of 600 bp, with no secondary peak present between 100 and 200 bp (that would indicate unremoved adaptors). A secondary peak twice the size of the original peak could mean the PCR reaction (Subheading 3.8.7, **step 2**) has been exhausted; this generally happens when performing eight or ten (or more) cycles. This peak results from concatemers of the amplified product. In such cases, reduce the number of PCR cycles. It is unclear whether these secondary peaks affect sequencing read numbers, and although we try to avoid them we have not observed any issues when they are present. They will not affect the downstream data quality.
25. Dam-only control files can also be manually specified on the command-line with the --dam = [filename] flag, but will be automatically detected if a single file starting with "Dam" is present.
26. If multiple replicates have been multiplexed, we recommend using the --just_align flag to generate BAM files from all samples, and then run the pipeline again on each set of BAM files from each replicate. Ideally, alignment rates should be above 90% (95% is typical).

References

1. Southall TD, Gold KS, Egger B, Davidson CM, Caygill EE, Marshall OJ, Brand AH (2013) Cell-type-specific profiling of gene expression and chromatin binding without cell isolation: assaying RNA Pol II occupancy in neural stem cells. Dev Cell 26:101–112. https://doi.org/10.1016/j.devcel.2013.05.020
2. Marshall OJ, Southall TD, Cheetham SW, Brand AH (2016) Cell-type-specific profiling of protein–DNA interactions without cell isolation using targeted DamID with next-generation sequencing. Nat Protoc 11:1586–1598. https://doi.org/10.1038/nprot.2016.084
3. van Steensel B, Delrow J, Henikoff S (2001) Chromatin profiling using targeted DNA adenine methyltransferase. Nat Genet 27:304–308. https://doi.org/10.1038/85871
4. van Steensel B, Henikoff S (2000) Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. Nat Biotechnol 18:424–428. https://doi.org/10.1038/74487
5. Luo GZ, He C (2017) DNA N 6-methyladenine in metazoans: functional epigenetic mark or bystander? Nat Struct Mol Biol 24:503–506. https://doi.org/10.1038/nsmb.3412
6. Doupé DP, Marshall OJ, Dayton H, Brand AH, Perrimon N (2018) Drosophila intestinal stem and progenitor cells are major sources and regulators of homeostatic niche signals. Proc Natl Acad Sci 115:12218–12223. https://doi.org/10.1073/pnas.1719169115
7. Marshall OJ, Brand AH (2017) Chromatin state changes during neural development revealed by in vivo cell-type specific profiling. Nat Commun 8:2271. https://doi.org/10.1038/s41467-017-02385-4
8. Aughey GN, Estacio Gomez A, Thomson J, Yin H, Southall TD (2018) CATaDa reveals global remodelling of chromatin accessibility during stem cell differentiation in vivo. elife 7:1–22. https://doi.org/10.7554/elife.32341
9. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, van Steensel B (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453:948–951. https://doi.org/10.1038/nature06947
10. Cheetham SW, Brand AH (2018) RNA-DamID reveals cell-type-specific binding of roX RNAs at chromatin-entry sites. Nat Struct Mol Biol 25:109–114. https://doi.org/10.1038/s41594-017-0006-4
11. Luukkonen BG, Tan W, Schwartz S (1995) Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. J Virol 69:4086–4094
12. Brand AH, Perrimon N (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118:401–415
13. Marshall OJ, Brand AH (2015) damidseq_pipeline: an automated pipeline for processing DamID sequencing datasets. Bioinformatics 31:3371–3373. https://doi.org/10.1093/bioinformatics/btv386
14. Delandre C, McMullen JPD, Marshall OJ (2020) Membrane-bound GFP-labelled vectors for Targeted DamID allow simultaneous profiling of expression domains and DNA binding. bioRxiv 1–4. https://doi.org/10.1101/2020.04.17.045948
15. Aughey GN, Delandre C, Southall TD, Marshall OJ (2020) FlyORF-TaDa allows rapid generation of new lines for *in vivo* cell-type specific profiling of protein-DNA interactions in *Drosophila melanogaster*. bioRxiv 2020.08.06.239251. https://doi.org/10.1101/2020.08.06.239251
16. Tosti L, Ashmore J, Tan BSN, Carbone B, Mistri TK, Wilson V, Tomlinson SR, Kaji K (2018) Mapping transcription factor occupancy using minimal numbers of cells in vitro and in vivo. Genome Res 28:592–605. https://doi.org/10.1101/gr.227124.117
17. Cheetham SW, Gruhn WH, van den Ameele J, Krautz R, Southall TD, Kobayashi T, Surani MA, Brand AH (2018) Targeted DamID reveals differential binding of mammalian pluripotency factors. Development 145:dev170209. https://doi.org/10.1242/dev.170209
18. Katsanos D, Barkoulas M (2020) Tissue-specific transcription factor target identification in the Caenorhabditis elegans epidermis using targeted DamID. bioRxiv
19. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. Nat Biotechnol 29:24–26. https://doi.org/10.1038/nbt.1754

20. Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, Habermann B, Akhtar A, Manke T (2018) High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nat Commun 9(1):189. https://doi.org/10.1038/s41467-017-02525-w
21. Tarazona S, Furió-Tarí P, Turrà D, Di Pietro A, Nueda MJ, Ferrer A, Conesa A (2015) Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. Nucleic Acids Res 43(21):e140. https://doi.org/10.1093/nar/gkv711
22. Hatch HAM, Belalcazar HM, Marshall OJ, Secombe JA (2021)KDM5-Prospero transcriptional axis functions during early neurodevelopment to regulate mushroom body formation. Elife. 2021 Mar 17;10:e63886. https://doi.org/10.7554/eLife.63886. PMID: 33729157; PMCID: PMC7997662