# **High-Throughput Genome Scanning in Constant Tension Fluidic Funnels**

# **Electronic Supplemental Information**

Joshua W. Griffis, Ekaterina Protozanova, Douglas B. Cameron, Robert H. Meltzer\*

Pathogenetix, Inc., 12 Gill St., Suite 3150, Woburn, MA 01801

\* To whom correspondence should be addressed. Email: <u>rmeltzer@pathogenetix.com</u>; Tel: (781) 939 6481; Fax: (781) 938 0060

### **Sample Preparation**

Genomic DNA from Escherichia coli K12 (MG1655, GenBank U00096.2)<sup>1, 2</sup> was prepared in an automated mini-reactor as described previously.<sup>3</sup> The purified genomic DNA was digested using the NotI restriction endonuclease (New England Biolabs, Ipswich, MA), and tagged with ATTO-550 conjugated bis-PNA tags (p58) designed to recognize the sequence AAAAGAAG.<sup>3</sup> PNA probes were designed by Pathogenetix and manufactured by Panagene (Daejeon, Korea). The prepared DNA sample was eluted at  $\sim 0.5 - 1.5$  ng/µl in TE buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, purchased as 100x concentrate from Sigma-Aldrich, St. Louis, MO). Immediately prior to stretching on the microfluidic device, a 10 µl aliquot of DNA was gently mixed with POPO-1 intercalator (Life Technologies, Grand Island, NY) in order to uniformly stain the DNA backbone. DNA concentration was quantified by ethidium bromide-stained gel prior to intercalation in order to standardize a ratio of three nucleic acid base pairs per dye. The length to which a given fragment of DNA is stretched under tension depends on the ratio of dye to DNA.<sup>4, 5</sup> For all comparisons presented in this study, paired experiments were therefore performed with a common stock of intercalated DNA. This also assured that the DNA concentration was uniform in each paired experiment set.

## **Chip Operation and Data Acquisition**

Single-molecule data acquisition was performed as previously described.<sup>6-9</sup> A 5  $\mu$ l sample of prepared, intercalated DNA was pipetted into the sample port of the device, and the sheath buffer ports were filled with fresh TE buffer. The sheathing buffer streams centered the DNA stream in the stretching funnel, targeting it to the projected laser spots in the detection channel. Also, all DNA molecules were subjected to a uniform flow stream in the center of the funnel, thus avoiding variations in velocity and fluidic path length near the outer walls of the funnel.

The chip was mounted to a custom confocal fluorescence microscope with accompanying pneumatic controls. Vacuum was applied at the waste port on the chip, thus drawing the DNA sample and sheathing buffer through the stretching and detection funnel. The microfluidic device was alternatively driven by applying positive pressure at the DNA injection port and sheath buffer ports simultaneously. This configuration was required to achieve high fluid velocities in 1  $\mu$ m etch depth devices.

Fluorescence emission from DNA flowing through the detection funnel was collected at three spatially separated spots along the fluidic channel (Fig 2A). Two 455 nm wavelength laser spots, separated by 20  $\mu$ m along the channel, elicited fluorescence from the POPO-1 stained DNA. A 532 nm wavelength laser spot excited ATTO-550 dyes attached to the PNA probes. This detection spot was fixed at 5  $\mu$ m before the first of the 455 nm spots.

The position of the excitation spots in the detection channel affected the observed stretching of DNA as discussed further in this investigation. The position of the lasers was therefore defined for each experimental condition. Precise alignment of the laser detection spots to the detection channel was achieved using a fiducial ruler etched alongside the detection channel in each device (not shown). The position of the laser spots was referenced to the transition from the stretching to the detection portion of the channel.

#### **DNA stretching in GSS fluidics**

GSS leverages the relative inextensibility of linearized DNA between 5 and 65 pN to provide a broad range of analyzable molecule lengths using a single funnel and flow velocity. Figure S1A and B show two-dimensional histograms (heat maps), plotting the distribution of fragment lengths versus average POPO-1 fluorescence intensity. The *E. coli* NotI digest was selected to provide DNA fragments ranging from 40 to 361 kb. In Fig. S1A, the sample was run on the CS 30 at 30  $\mu$ m/ms to demonstrate uniform DNA extension across the entire fragment length range. In Fig. S1B, the same sample was run on CV 7.5 at 30  $\mu$ m/ms to accentuate overstretching of longer DNA fragments. Representative individual molecules from positions noted in Fig. S1B are presented in Fig S1C. As noted in Eq. 13 in the text, the maximum tension on a molecule is proportional to the square of its length. Very short molecules do not fully extend in the GSS funnel, resulting in a population of fragments that exhibit increased fluorescence intensity and shorter observed length than would be expected (Fig. S1C-1). The extent of stretching of such short fragments is highly stochastic, resulting in a hyperbolic distribution of fragments in this plot, with a constant product of length and intensity.

In Fig. S1B, molecules ranging in length from 25 to 70  $\mu$ m achieved complete stretching. This was evidenced by their uniform average fluorescence intensity. Individual molecules also typically displayed uniform intercalator fluorescence along the length of each molecule (Fig S1C-2). Molecules longer than 75um however, showed characteristics of overstretching. In particular, the average molecule fluorescence dropped below the level of the well-stretched shorter fragments. Each molecule also had a characteristic dip in fluorescence intensity at its center (Fig S1C-3). This is consistent with the parabolic tension distribution predicted along the stretched molecule. The goal in design of the GSS funnel is to maximize the range of DNA lengths that reside between the bounds of the under-stretching and over-stretching thresholds. Fig. S1A demonstrates ideal stretching across the entire range of *E. coli* restriction fragments.



**Figure S1:** *E. coli* NotI digest stretching in GSS. A, B) Heat map of molecule length *vs.* average POPO-1 fluorescence intensity. Data was acquired in CS 30 at 30  $\mu$ m/ms (A) and CV 7.5 at 30  $\mu$ m/ms (B). Intensity of heat map corresponds to number of observed molecules. C) Single representative molecules from positions 1, 2, and 3 in B. Blue and red traces are from duplicate records of POPO-1 fluorescence intensity. The green trace indicates fluorescence intensity of site-specific PNA probes.

### Tension distribution along a molecule in GSS

Tension varies along the length of an extended DNA molecule in GSS. The tension at any point along the molecule is given by Eq. 12 in the text, and is dependant on the difference in velocity of the molecule and the fluid velocity at the point of interrogation. The tension distribution is therefore dependent on the surrounding channel geometry, as demonstrated in Fig. S2. A 100 µm DNA molecule was simulated at multiple positions within the CV30 and CS30 funnels, assuming a fluid velocity of 30 µm/ms. In the CV30 device, the molecule tension distribution was calculated for a molecule positioned at the origin of the constant velocity channel, as well as in 25 µm increments along the channel (Fig. S2 A, B). At the start of the detection channel, the tension distribution on the molecule is nearly parabolic, with a peak tension near 65 pN. As the molecule proceeds along the channel, the peak tension decreases and shifts towards the rear of the molecule as it becomes surrounded by the constant velocity flow. Once the entire length of the molecule has traversed into the constant-velocity channel, the tension drops to zero. In contrast, the same 100 µm molecule was calculated at three positions separated by 25  $\mu$ m, but fully enclosed within the constant strain-rate detection channel of CS30. In this geometry, the tension distribution is parabolic along the length of the molecule and identical at all three positions within the funnel.



**Figure S2:** Tension distribution along a DNA molecule in GSS. A) Cartoon of molecule within CV funnel. B) Tension distribution along length of the molecule, calculated by Eq. 12 from the text. Numbered curves correspond to molecules positioned with the leading end at the corresponding locations in A. C) Cartoon of molecule within CS funnel. B) Tension distribution of molecule fully enclosed within CS channel.

#### Acceleration correction in constant strain rate detection channels

The quantitative, analytical expression of acceleration can be derived as follows. Within the constant strain rate portion of the funnel, fluid velocity increases linearly with position x from the origin of the funnel, as defined in the text (Eq. 8). The molecule velocity is equal to the average fluid velocity over its length, as shown in Eq. 11. The molecule velocity is identical to the fluid velocity at the center of the molecule ( $x_m$ ):

$$v_{mol} = v(x_m) = \delta x_m$$
 Eq. S1

The time required for a molecule to travel a distance *dx* is thus:

$$dt = \frac{dx_m}{v(x_m)} = \frac{dx_m}{\mathscr{A}_m}$$
 Eq. S2

where  $x_m$  is the position of the middle of the molecule.

Integrating this starting from an arbitrary molecule position  $x_{0m}$  yields:

$$\Delta t = \frac{1}{\mathscr{K}} \log(1 + \frac{\Delta x}{x_{0m}})$$
 Eq. S3

where  $\Delta t$  is the time required for the molecule to travel a distance  $\Delta x$  starting at  $x_{0m}$ .

In an interval of time  $\Delta t$  the molecule travels a distance given by inverting Eq. S3:

$$\Delta x = x_{0m}(e^{\Delta t} - 1)$$
 Eq. S4

This differs from constant velocity in which the distance traveled would be proportional to the time interval.

If  $x_{0m}$  is the position of the middle of the middle of a completely linearized as the head of the molecule enters the tag spot at  $x_{tag}$ , then  $x_{0m} = x_{tag} - L/2$ . The distance a molecule travels over a time  $\Delta t$  starting when the molecule enters the tag spot is:

$$\Delta x = (x_{tag} - L/2)(e^{\varepsilon \Delta t} - 1)$$
 Eq. S5

A molecule transiting the tag spot will travel a distance:

$$\Delta x = L$$
 Eq. S6

requiring a time (using Eq. S3):

$$\Delta t_{tag} = \frac{1}{\mathscr{K}} \log \left( 1 + \frac{L}{x_{tag} - L/2} \right)$$
 Eq. S7

The average velocity during this time is therefore:

$$\overline{v} = \frac{\Delta x}{\Delta t_{tag}} = \frac{L \&}{\log \left(1 + \frac{L}{x_{tag} - L/2}\right)}$$
Eq. S8

Defining  $\tau$  as the relative time required for the molecule to transit the tag spot:

$$\tau \equiv \frac{\Delta t}{\Delta t_{tag}}$$
 Eq. S9

The difference,  $\Delta x_c$ , between the distance traveled assuming this constant velocity and that actually traveled,  $\Delta x$ , is:

$$\Delta x_c = \Delta x - L\tau \qquad \text{Eq. S10}$$

Using Eq. S1 for  $\Delta x$ , the trace acceleration correction for a tapered funnel is given by:

$$\Delta x_c(\tau) = (x_{tag} - L/2)(e^{\varepsilon \tau \Delta t_{tag}} - 1) - L\tau \qquad \text{Eq. S11}$$

Using the time required for the molecule to transit the tag spot,  $\Delta t_{tag}$ , from Eq. S7, then:

$$\Delta x_{c}(\tau) = (x_{tag} - L/2) \left( \left( 1 + \frac{L}{x_{tag} - L/2} \right)^{\tau} - 1 \right) - L\tau \quad \text{Eq. S12}$$

This is defined as the error in position across the molecule that results from assuming a constant velocity. This is an analytically exact expression of the trace acceleration correction. The correction is zero at the start  $\tau = 0$  and end  $\tau = 1$  of the molecule.

The exact expression for the acceleration correction,  $\Delta x_c(\tau)$ , can be simplified and approximated by expressing the equation as a series expansion for the length of the molecule, *L*. When the molecule length *L* is much less than the distance to the funnel origin, i.e.  $x_{tag} > L$ , the first term dominates so

$$\Delta x_c \simeq -\frac{L^2}{2x_{tag}} \tau (1-\tau)$$
 Eq. S13

This approximation for acceleration correction is applied in the main text (Eq. 15).

## References

- M. Riley, T. Abe, M. B. Arnaud, M. K. Berlyn, F. R. Blattner, R. R. Chaudhuri, J. D. Glasner, T. Horiuchi, I. M. Keseler, T. Kosuge, H. Mori, N. T. Perna, G. Plunkett, 3rd, K. E. Rudd, M. H. Serres, G. H. Thomas, N. R. Thomson, D. Wishart, and B. L. Wanner, *Nucleic Acids Res.*, 2006. 34(1): p. 1-9.
- F. R. Blattner, G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao, *Science*, 1997. 277(5331): p. 1453-62.
- E. Protozanova, M. Zhang, E. J. White, E. T. Mollova, D. Ten Broeck, S. V.
  Fridrikh, D. B. Cameron, and R. Gilmanshin, *Anal. Biochem.*, 2010. 402: p. 83-90.
- R. Eckel, R. Ros, A. Ros, S. D. Wilking, N. Sewald, and D. Anselmetti, *Biophys. J.*, 2003. 85(3): p. 1968-1973.
- 5 K. Gunther, M. Mertig, and R. Seidel, *Nucleic Acids Res.*, 2010. **38**(19): p. 6526-32.
- R. E. Burton, E. J. White, T. R. Foss, K. M. Phillips, R. H. Meltzer, N. Kojanian, L. W. Kwok, A. Lim, N. L. Pellerin, N. V. Mamaeva, and R. Gilmanshin, *Lab Chip*, 2010. 10(7): p. 843-51.
- J. W. Larson, G. R. Yantz, Q. Zhong, R. Charnas, C. M. D'Antoni, M. V. Gallo, K. A. Gillis, L. A. Neely, K. M. Phillips, G. G. Wong, S. R. Gullans, and R. Gilmanshin, *Lab Chip*, 2006. 6: p. 1187-1199.
- 8 E. J. White, S. V. Fridrikh, N. Chennagiri, D. B. Cameron, G. P. Gauvin, and R. Gilmanshin, *Clin. Chem.*, 2009. **55**(12): p. 2121-2129.
- 9 R. H. Meltzer, J. R. Krogmeier, L. W. Kwok, R. Allen, B. Crane, J. W. Griffis, L. Knaian, N. Kojanian, G. Malkin, M. K. Nahas, V. Papkov, S. Shaikh, K. Vyavahare, Q. Zhong, Y. Zhou, J. W. Larson, and R. Gilmanshin, *Lab Chip*, 2011. **11**(5): p. 863-73.