INgen: Intracellular Genomic DNA Amplification for Downstream Applications in Sequencing and Sorting

Crossland[1,2*], P., Schmidlin[1,2*], K., Valli-Doherty[1,3], I., Brettner,[1,2,+], L., Geiler-Samerotte[1,2+], K.

[1] Biodesign Center for Mechanisms of Evolution, Arizona State University, Tempe, AZ, 85287
[2] School of Life Sciences, Arizona State University, Tempe AZ, 85287
[3] School of Biological and Health Systems Engineering, Arizona State University, Tempe AZ, 85257
[*]co-first authors
[+]co-senior authors

Abstract

Here, we introduce intracellular genomic amplification (INgen), a method that harnesses the cell membrane as a natural reaction chamber for DNA amplification, enabling downstream sequencing and cell sorting. Unlike traditional single-cell techniques, INgen utilizes a strand-displacing, isothermal polymerase to amplify DNA *within* fixed, permeabilized cells while maintaining the cell's structural integrity. This approach overcomes challenges associated with both typical single-cell DNA sequencing and hindrances encountered when previously attempting to sequence genetic material from fixed microbial cells, allowing amplification of genomic regions up to 100 kb and sequencing of whole genomes from diverse cell types, including *S. cerevisiae*, *B. subtilis*, and *E. coli*. Additionally, INgen can be adapted for targeted DNA enrichment using biotinylated primers and for fluorescence-based cell sorting.

Introduction

Single-cell DNA sequencing has the potential to transform our ability to study genetic heterogeneity in mixed microbial communities (e.g., microbiomes) and in any population where not all cells are genetically identical (e.g. tumors, infections, laboratory experiments). However, a major bottleneck that limits many single-cell 'omics technologies is the reliance on physical cell isolation which constrains scalability. When performing single-cell genomics, this often restricts studies to hundreds (Venkataram et al. 2016; Gasch et al. 2017; Schmidlin et al. 2024) or even dozens of cells (Behringer et al. 2024; Zion et al. 2024). Additionally, such techniques can require time-consuming labor, specialized expertise, and/or costly instrumentation which limits accessibility. While most single-cell DNA methods sort cells into vessels to perform amplification, ironically, each cell's genetic material is naturally compartmentalized by its membrane, providing a built-in reaction chamber. Using the cell as its own reaction container has been extensively harnessed for single-cell transcriptomics (Cao et al. 2017; Rosenberg et al. 2018; Blattman et al. 2020; Kuchina et al. 2021; Martin et al. 2023; Brettner et al. 2024). However, intracellular DNA-based applications have not been widely implemented due to several technical challenges (Hully 1999; Speel et al. 1999), including the hurdle of amplifying DNA in formaldehyde-fixed cells in such a way that allows it to be sequenced. A robust method for in situ DNA amplification that allows for downstream DNA sequencing thus stands to enable the optimization of higher-throughput single-cell DNA sequencing protocols.

A significant challenge in intracellular DNA amplification is the need to fix cells with formaldehyde before the *in situ* reactions. Fixing with formaldehyde stabilizes the cell and the genetic material inside, but the crosslinks that are generated are frequently cited as obstacles to efficient DNA amplification and can hinder sequencing (O'Leary et al. 1994; Clingenpeel et al. 2014; Pereira et al. 2022; Kirsch et al. 2024). While some studies suggest that formaldehyde can be reversed such that they do not severely inhibit amplification (Bagasra 2007; Gavrilov and Razin 2009), concerns about formaldehyde have driven efforts to perform FISH and FACS in live cells instead (Pereira et al. 2022). While short, engineered DNA sequences (15 bp) (Payne et al. 2021) and RNA transcripts have been successfully amplified in fixed cells (Rosenberg et al. 2018; Kuchina et al. 2021; Brettner et al. 2024; Salatino et al. 2024), we found no reports of sequencing genomic DNA that was amplified inside of formaldehyde-fixed cells. Instead, recent intracellular methods randomly insert transposons that drive transcription of DNA to RNA, cleverly circumventing the challenge of intracellularly amplifying and sequencing DNA (Vitak et al. 2017; Yin et al. 2019).

Here we introduce intracellular genomics (INgen), a novel method that addresses existing challenges and enables intracellular DNA amplification in cell suspension in a way that unlocks downstream applications such as cell sorting and/or DNA sequencing. A key innovation of INgen is the use of phi29 polymerase, which can overcome formaldehyde crosslinks through its strand-displacing activity (Lee et al. 2006; Aviel-Ronen et al. 2006; Tang et al. 2020). Another important feature of phi29 is that it is an isothermal polymerase, which does not require temperature cycling, and thus avoids the cell membrane damage, extracellular amplification, and false positives or negatives that derailed earlier efforts to amplify DNA intracellularly (Speel et al. 1999). Another key innovation is the use of biotinylated primers, which allows us to streptavidin bind intracellularly amplified DNA and purify it for downstream reactions such as DNA sequencing. Using INgen, we successfully sequenced target genes up to 3kb and whole genomes in diverse organisms, including *S. cerevisiae*, *B. subtilis*, and *E. coli*. Further, we demonstrated that up to 14% of the yeast genome can be intracellularly amplified within a single yeast cell. Finally, we demonstrate that the presence of additional DNA within a cell allows us to sort and sequence cells containing intracellularly amplified DNA based on the strength of the fluorescent signal of a DNA-binding dye. In sum, INgen is a robust and versatile method for intracellular DNA amplification in diverse cell types which paves the way for a wide array of applications, including high-throughput single-cell sequencing, rare cell genomics, and genotype-based sorting. This approach has the potential to advance the capabilities of single-cell 'omics, with broad potential applications in fields such as microbial ecology, cancer genomics, and

evolutionary biology.

Results

A flexible method for intracellular DNA Amplification (INgen)

In this manuscript, we tackle the challenge of amplifying DNA within a single cell and demonstrating this DNA can be sequenced. The INgen workflow consists of four key steps: (1) fixation and permeabilization of cells to stabilize the cellular structure and allow reagent access (**Fig 1a–c**), (2) intracellular DNA amplification using an isothermal polymerase (**Fig 1d**) and biotinylated primers, (3) isolation of newly synthesized DNA via streptavidin-coated beads, and (4) downstream sequencing of amplified genomic material. We will show that intracellular amplification and subsequent sequencing of amplified products is not only possible but can be successfully applied in multiple contexts which highlights this method's utility for downstream single-cell sequencing applications.

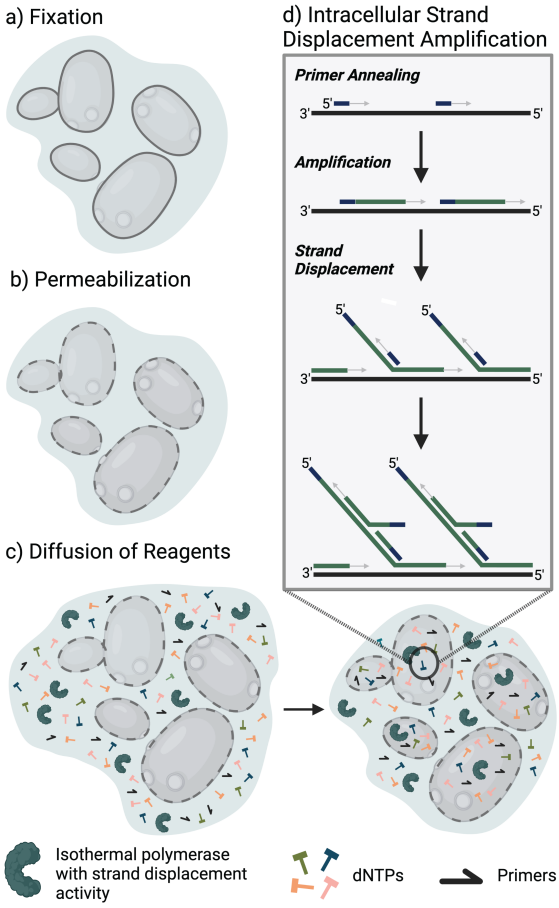


Figure 1. Schematic of the intracellular DNA amplification (INgen) workflow. This robust and flexible protocol can be altered at each step based on the cell type being amplified. **a)** Fixation stabilizes the cell, which needs to remain intact during the entire amplification reaction. Optimal fixation conditions can vary by organism: 4% formaldehyde for yeast and 70% ethanol for bacteria. **b)** Permeabilization allows reagents to diffuse inside cells and optimal methods again differ across taxa. Yeast are treated with zymolyase and triton-X, while bacterial membranes become permeable during ethanol fixation. **c)** Reaction reagents including DNA polymerase (phi29), dNTPs, and biotinylated primers are diffused into the cell by the crowding agent sorbitol. See methods for a detailed master mix recipe. **d)** Amplification occurs via strand displacement which previous work suggests may help overcome formaldehyde crosslinks.

To assess the broad applicability of our method across taxa, we tested intracellular amplification

in the eukaryotic model organism *Saccharomyces cerevisiae* (yeast) and two prokaryotic species, *Bacillus subtilis* (a gram-positive bacteria) and *Escherichia coli* (a gram-negative bacteria). Using species-specific fixation and permeabilization protocols, we successfully amplified genomic material in all three organisms. For yeast, fixation with formaldehyde followed by treatment with zymolyase and triton-X enabled effective permeabilization, while ethanol fixation proved most effective for *E. coli* and *B. subtilis*. The results demonstrate that our platform is robust to the characteristic differences in the cell wall and membrane composition of gram-negative and gram-positive bacteria and fungi, making it broadly applicable across phylogenetically distant organisms.

In-situ amplification and sequencing of targeted genomic regions

To evaluate the ability of INgen to selectively amplify specific genomic regions, we designed biotinylated primers targeting genes of varying lengths, including three genes from yeast (SOD1, 450bp; PDR3, 3,000bp; and an engineered barcode BC, 300bp) and one gene from bacteria (16S rRNA, 1,000bp). Cells were incubated with phi29 master mix for 16 hours at 30C, lysed to release DNA, and newly synthesized DNA amplicons were isolated by binding their covalently-bound biotin molecules to streptavidin-coated beads. Once bound to the streptavidin-coated beads and washed three times, a secondary extracellular amplification reaction with the same targeted primers (modified for extracellular amplification) was performed to generate sufficient DNA that is not bound by streptavidin beads for library preparation, which prepares DNA for next generation sequencing. Gel electrophoresis analysis confirmed successful targeted amplification off of the beads, with distinct bands of the expected sizes observed for all tested loci (**Fig 2A**). In contrast, amplification reactions performed using the wash (non-bead bound DNA) failed to produce detectable products on an agarose gel, even when large amounts of bacteria DNA were spiked into the tube prior to washing. This confirms that amplification is specific to intracellularly amplified, biotinylated DNA and not native genomic DNA (**Fig S1**; also see **Fig 4**).

We amplified a wide range of gene lengths with the smallest being 300bp and the largest being 3000bp. While we didn't attempt to amplify any regions above 3kb, the highly processive nature of phi29 suggests that longer targets could also be amplified (Blanco et al. 1989). Following size confirmation on an agarose gel, DNA fragments were prepared for sequencing using the Illumina DNA Flex kit. Targeted genomic regions were successfully sequenced to a minimum depth of 15X which is sufficient for variant calling (**Fig 2B**). The ability to target multiple genes across multiple species highlights the flexibility and broad applicability of this method.

Whole genome amplification with random hexamers in diverse cell types

A potential concern with INgen is whether certain genomic regions might be inaccessible to the isothermal polymerase, Our experiments suggest this is not a major limitation. To assess whether specific genomic regions are resistant to intracellular amplification, we performed INgen whole genome amplification in bacterial (*E. coli*) cells using biotinylated random hexamers. We then isolated newly synthesized (i.e. biotinylated) DNA from single cells, performed additional rounds of PCR using random hexamers to copy these DNA sequences off of the beads, and submitted the resulting pool of DNA for sequencing.

Gel electrophoresis of the off-the-beads PCR product revealed smears spanning a broad range of fragment sizes, consistent with successful whole-genome amplification. Fragment sizes ranged from 200 bp to 50 kb (**Fig 3**). TapeStation analysis further confirmed that intracellular amplification generated long DNA fragments, with an average size of 20kB, and some spanning up to several hundred kilobases (**Fig S2, S3 & S4**), consistent with the expected performance of phi29 polymerase. These results suggest that formaldehyde crosslinking does not significantly inhibit this polymerase's activity in our reaction conditions. Here we demonstrate that it is possible to amplify the majority of the *E. coli* genome by performing in-situ amplification with biotinylated random hexamers (**Fig 3**). From a single in situ PCR reaction which contained 1 million cells, 98.59% of the *E. coli* genome was covered to an average depth of 19.81 (**Fig 3**). This result suggests that the INgen method can

be used to capture complete genomes.

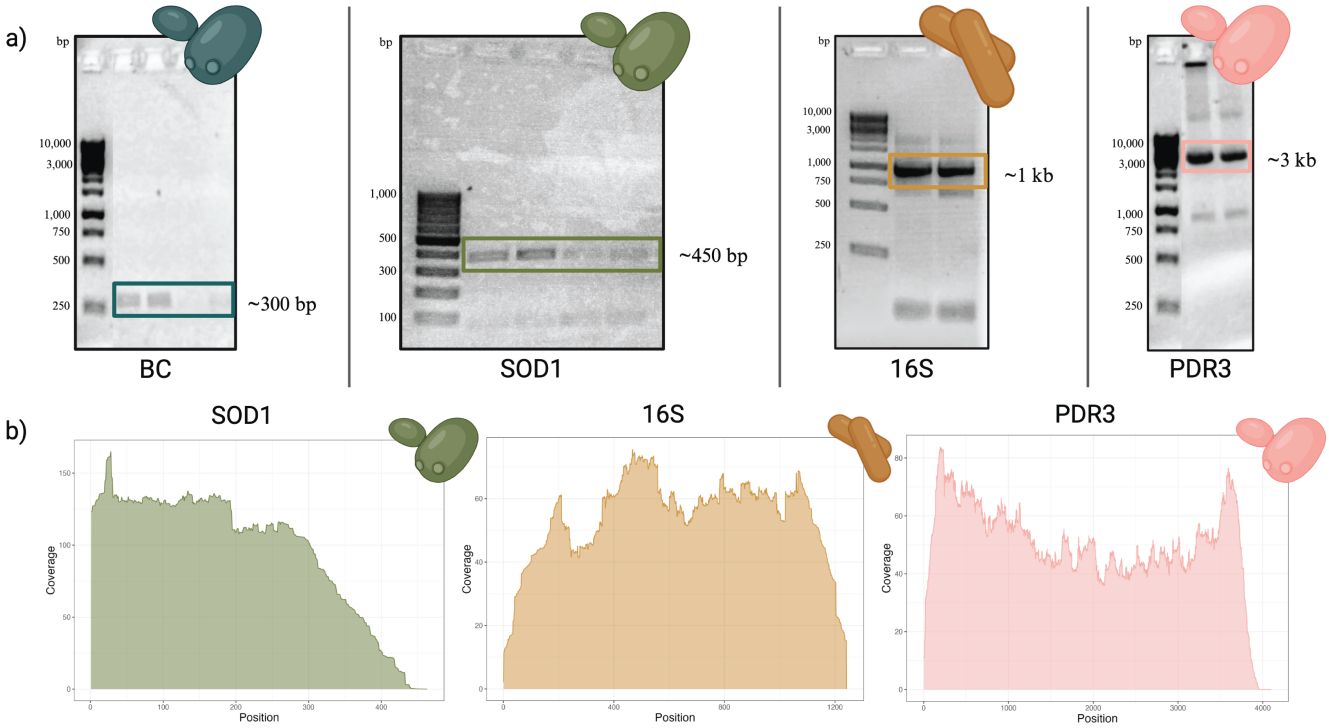


Figure 2. Targeted amplification and sequencing of specific genomic regions. a) Gel electrophoresis of amplified DNA fragments targeting different genomic loci. BC: barcode region (300 bp, teal), SOD1 (450 bp, green), 16S rRNA (1 kb, orange), and PDR3 (3 kb, pink). Bands corresponding to the expected fragment sizes are highlighted with colored boxes. **b)** Sequencing coverage plots for SOD1 (green), 16S rRNA (orange), and PDR3 (pink), show the read depth across each amplified region. Coverage varies across each fragment, but most bases are covered over 15X.

The use of biotinylated primers is particularly helpful in whole genome sequencing with random hexamers. Our approach mitigates contamination issues commonly encountered in traditional single-cell genome sequencing, where cells are lysed before DNA amplification. Post-lysis amplification can result in exogenous DNA (i.e. contamination) being amplified along with the genome of interest. This problem is exacerbated in single-cell sequencing by the limited amount of starting genomic DNA within a single cell (Nawy 2014; Lusk 2014; Lux et al. 2015). By using biotinylated primers, INgen enables selective enrichment of intracellularly amplified DNA, while allowing non-biotinylated DNA and DNA that was biotinylated while outside of a cell to be effectively removed by wash steps. This is particularly important when amplifying DNA with random hexamers that are not specific to a sequence of interest. By selectively retaining only genomic DNA, our method improves sequencing accuracy and reliability.

Successful sorting and sequencing of individual cells with intracellularly amplified DNA

To further validate that DNA amplification is occurring intracellularly, we designed control experiments in yeast cells where we omitted either polymerase or primers from our intracellular reactions. We then stained control and test yeast cells with EvaGreen, a fluorescent dye that binds to all double-stranded DNA, and compared fluorescence distributions. Indeed, test populations exhibited significantly greater average fluorescence compared to controls, likely because the test reactions possessed all necessary reagents to intracellularly amplify DNA (**Fig 4A**). Since EvaGreen stains both the native genome and newly made DNA, we used flow cytometry to identify only those test cells exhibiting brighter fluorescence than measured in any control cell. Of the test cells that were brighter than any control, we then gated the brightest cells assuming that higher fluorescence indicated more robust intracellular genomic amplification. The fluorescent cells that met our criteria were sorted

into individual wells of a 96-well plate for downstream single-cell sequencing. Our primary objective was to determine if the brightest cells indeed contained biotinylated DNA. Our second objective was to learn what fraction of the genome amplified within any given single cell.

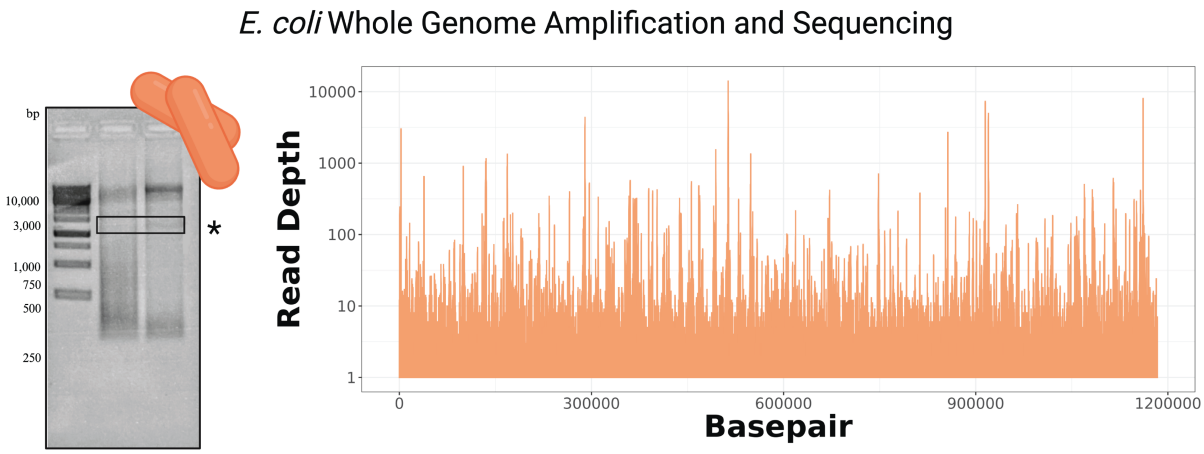
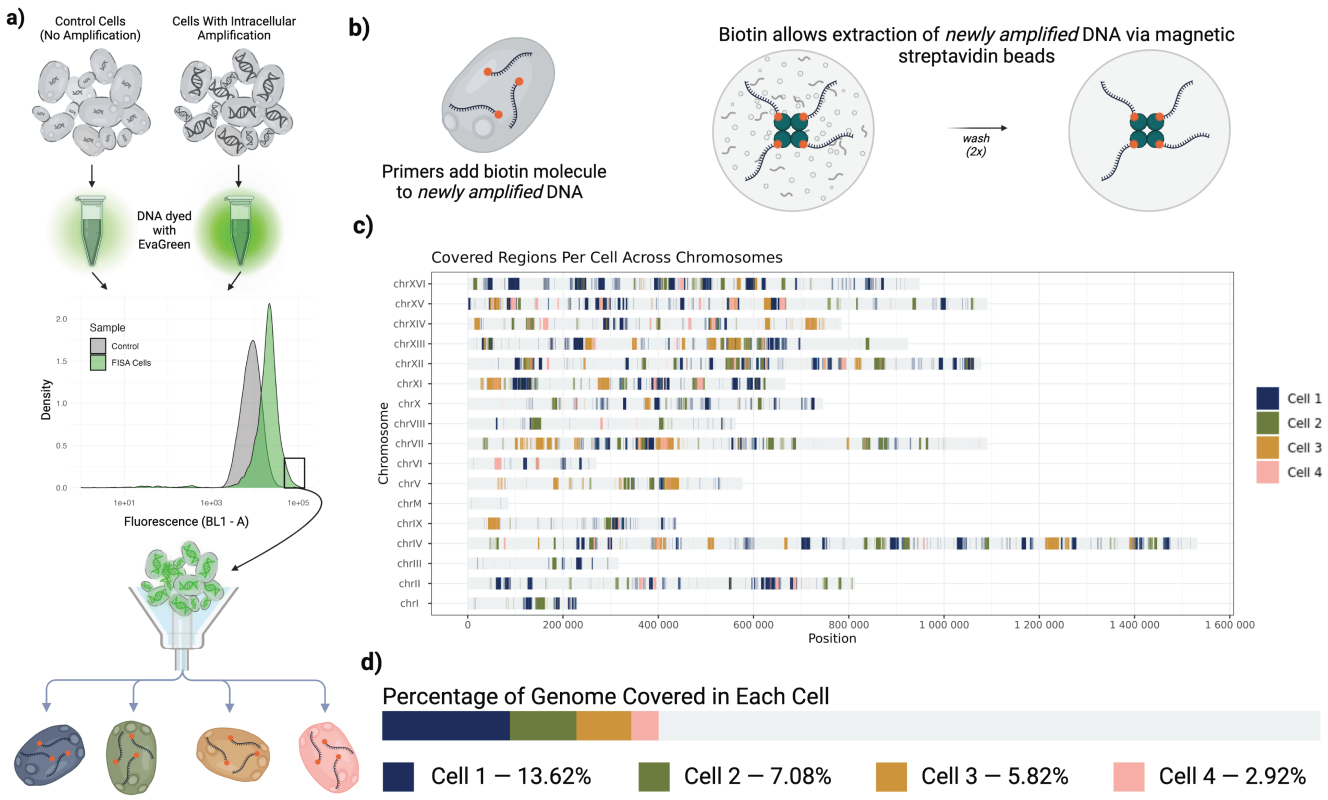


Figure 3. Whole Genome Amplification and Sequencing in *E. coli*. *E. coli* whole genome amplification and sequencing results. Left: Gel electrophoresis image of amplified *E. coli* DNA, with a prominent high-molecular-weight smear and a boxed region highlighting a specific fragment size. An asterisk (*) indicates an additional band corresponding to the size of a 16S amplicon, as 16S primers were added alongside random hexamers in this reaction. This suggests that multiple primer types can be used within the same reaction. Right: Read depth across the *E. coli* genome, showing sequencing coverage in orange.

We performed single-cell sequencing on 4 of the 96 sorted single cells. Each cell was lysed using Proteinase K, and the newly synthesized, biotinylated DNA was captured with streptavidin-coated beads and washed several times to remove native genomic DNA (**Fig 4B**). For all four cells, the bead-bound DNA was then re-amplified using random hexamers to generate enough starting material ($>$5 ng/μL) to construct sequencing libraries. This enabled the specific sequencing of only newly synthesized DNA from individual cells specifically, which allows us to then report the percent of the genome this method can amplify per cell. Sequencing data revealed that between 3 and 13% of the yeast genome was recovered from each cell, a recovery rate comparable to current single-cell sequencing technologies (**Fig 4C & D**) (Navin and Hicks 2011; Laks et al. 2019). Importantly, there was minimal overlap between the genomic regions recovered from different cells (**Fig 4C**), suggesting that our method does not introduce significant bias in amplification and provides a more representative sampling of the genome when analyzing multiple single cells.

The ability to sequence only the newly synthesized DNA from a single cell is a significant advantage of our method. This capability could be particularly valuable for studies investigating transposable element activity (Chen et al. 2017), de novo mutation detection (Muyas et al. 2024), or DNA replication dynamics (Müller et al. 2019; Georgieva et al. 2020), where distinguishing newly amplified sequences from unwanted native DNA is essential. Furthermore, by selectively enriching amplified regions, our method prioritizes high-quality genomic data, which may be especially beneficial in resource-limited sequencing applications. However, by washing away non-biotinylated DNA before sequencing, our method significantly reduces contamination, ensuring that only amplified DNA from the target cell is sequenced.



Intracellular Genome Amplification and Sequencing Coverage Analysis. **a)** Schematic representation of intracellular genome amplification. Control cells (left) do not undergo amplification, while treated cells (right) contain newly synthesized DNA. DNA was stained with EvaGreen, and fluorescence intensity measured by flow cytometry, revealing increased fluorescence in cells that underwent intracellular amplification compared to controls. Amplified cells are then sorted based on high fluorescent signals for sequencing. **b)** Biotinylated primers enable selective extraction of newly amplified DNA using magnetic streptavidin beads, followed by multiple washes to purify the sample. **c)** Genome coverage analysis of four individual cells mapped to *S. cerevisiae* chromosomes. Each colored segment represents a covered region for a specific cell. (d) Summary of genome coverage for each sequenced cell. The percentage of the genome covered is highest in Cell 1 (13.62%) and lowest in Cell 4 (2.92%).

Discussion

Our intracellular amplification platform demonstrates robust performance across diverse experimental conditions, including different species, varied fixation and permeabilization protocols, and amplification targets. By incorporating biotinylated primers into the amplification process, our method can selectively isolate newly synthesized DNA confirming that amplification is indeed occurring inside of cells. We validate this method by amplifying and sequencing three distinct genomic loci in yeast and one in bacteria (**Fig 2**). Beyond targeted amplification, we demonstrate whole genome in situ amplification is possible. A single, pooled in situ PCR reaction resulted in the recovery of over 98% of the *E. coli* genome. This indicates that nearly all regions of the bacterial genome are accessible for in situ amplification (**Fig 3**). We extended and adapted this approach to investigate single-cell whole-genome amplification in *S. cerevisiae*. After amplification, yeast cells were sorted based on DNA concentration calculated by EvaGreen staining. We show up to 14% of the yeast genome can be amplified in a single cell (**Fig 4**). Notably, there was minimal overlap in the regions sequenced between individual cells which indicates, like bacteria, a substantial portion of the yeast genome is accessible for amplification. By validating the method in both eukaryotic and prokaryotic species, as well as through targeted and genome-wide amplification, we highlight its adaptability for a wide range of biological applications, including its potential for use in single-cell sequencing

applications.

Comparison with Existing Amplification and Sequencing Techniques

When comparing INgen to other amplification and intracellular techniques, several key advantages become apparent. Most directly related to this work is In Situ PCR (IS-PCR) which enhances signal for probe hybridization by utilizing intracellular DNA amplification. (Murray 1993; Nuovo 1997; Muro-Cacho 1997; Hully 1999; Speel et al. 1999; Bagasra 2007). This technique has traditionally been used to localize cells containing genetic elements within tissues or tumors (Murray 1993; Hully 1999; Bagasra 2007; Janiszewska et al. 2015). Specific applications of IS-PCR have included detecting HIV (Bagasra et al. 1992; Gibellini et al. 1998), and malaria in human cells (Hashimoto et al. 2018), analyzing oncogenes in tumors (Janiszewska et al. 2015; Salatino et al. 2024), and detecting pathogens in environmental samples (Tani et al. 1998). However, the widespread adoption of this technique has been stalled by technical issues including, compromised cell membranes, extracellular amplification, non-uniform amplification, and false signals (O'Leary et al. 1994; Speel et al. 1999; Yin et al. 2019). Further, IS-PCR is usually performed on paraffin-embedded samples which harshly restricts downstream applications such as sorting labeled cells for 'omics analyses of genotypes of interest (Bagasra 2007; Janiszewska et al. 2015; Hashimoto et al. 2018). INgen can overcome many of the technical challenges of IS-PCR. We can work with cells in solution, our method's low temperatures do not compromise membranes, and we can isolate newly synthesized DNA from native genomic DNA and extracellular contaminants.

Another advantage of our approach over traditional amplification techniques is its ability to work with fixed cells. Conventional multiple displacement amplification (MDA) and other whole-genome amplification methods typically require the sorting of live cells and subsequent lysis. This can present challenges in applications where cell viability is difficult to maintain or harsh lysis agents are required as they can interfere with amplification reactions (Spits et al. 2006; Picher et al. 2016). By enabling whole-genome amplification from fixed cells, our method overcomes these constraints and expands the scope of potential applications, particularly in clinical and environmental settings where sample integrity and preservation are essential (Dietmaier et al. 1999; Huang et al. 2015).

Unlike the more common single-cell RNA-sequencing approaches, which are restricted to transcriptomic analysis, our method enables the amplification and subsequent sequencing of genomic DNA from individual cells, which allows for the detection of rare mutations, structural variations, and large genomic regions that may be missed in transcriptomic studies (de Bourcy et al. 2014; Borgström et al. 2017).

The INgen method presented in the paper serves as the first step towards improved single-cell DNA sequencing that does not require cell isolation and is flexible enough to work with a wide range of organisms. Intracellular whole-genome amplification with barcoded primers could enable high-throughput single-cell DNA sequencing without the need to physically separate cells, as it has for single-cell RNA sequencing (Rosenberg et al. 2018; Kuchina et al. 2021). Finally, INgen also allows combining both targeted and random primers within a single cell (**Fig 3**), which can help provide an anchor point —such as a 16S rRNA gene or an engineered barcode— to cluster multiple single-cell genomes into a consensus genotype. Thus, by selecting the primers and reagents diffused into cells, INgen can be adapted for applications spanning single-cell genomics, population-level studies, and targeted enrichment of specific genomic regions.

Methods

Fixation and Permeabilization

Fixation and permeabilization are crucial steps in our method, as they maintain cellular integrity while enabling reagent diffusion for intracellular genome amplification. This balance is necessary to preserve cell structures while ensuring effective reagent penetration. To achieve this, we evaluated multiple fixation strategies to identify optimal conditions.

Formaldehyde-based fixation provided a robust approach by crosslinking proteins and stabilizing cellular architecture while preserving nucleic acids. We tested protocols ranging from short incubations at room temperature (e.g., 10 minutes, as described in Payne et al., 2021) to extended fixation at 4°C, a commonly used method for yeast preservation (Brettner et al. 2024). Both protocols maintained cell morphology and supported efficient downstream amplification.

Ethanol fixation/permeabilization offered an alternative strategy. Ethanol rapidly precipitates proteins and disrupts membranes, enhancing permeability (Wagner et al. 1998). While harsher than formaldehyde, this method effectively permeabilized bacterial cells such as *E. coli* DH5. Its simplicity and rapid processing time (under 15 minutes) made it an attractive option.

Interestingly, the fixation method had minimal impact on genomic amplification efficiency or fluorescent labeling, underscoring the robustness of our approach. However, differences in cell durability and fluorescence intensity were observed, suggesting that formaldehyde fixation may better preserve intracellular targets for flow cytometry.

Following fixation, permeabilization was required to facilitate reagent entry. We tested enzymatic treatments, such as zymolyase digestion for yeast cell walls, and detergent-based methods using Tween-20 and Triton X-100. Zymolyase treatment efficiently converted yeast cells into spheroplasts while maintaining structural integrity. Detergent-based approaches provided a rapid, scalable alternative applicable to both yeast and bacterial cells.

These optimizations of fixation and permeabilization enabled intracellular amplification across diverse cell types. Our approach refines existing protocols to meet the specific demands of fluorescently labeled intracellular genome amplification.

Intracellular Amplification

The central innovation of our method is the ability to amplify genomic DNA directly within fixed and permeabilized cells. By performing amplification intracellularly, we mitigate biases inherent in traditional single-cell sequencing, such as uneven amplification or loss of fragile DNA regions. Instead, our method leverages the cellular environment as a natural reaction vessel. The selection of an appropriate DNA polymerase was critical for the success of this approach. We employed a thermostable polymerase optimized for isothermal amplification, capable of functioning despite residual fixative or membrane debris. This polymerase facilitated consistent amplification across various cell types, including yeast and bacteria. Another key feature of our approach is the intracellular reaction setup. By allowing reagents to diffuse into cells, we ensured uniform genome-wide amplification, avoiding the biases of traditional workflows that rely on DNA extraction, purification, and in vitro amplification. Intracellular amplification preserved genomic integrity and improved the detection of low-abundance sequences, such as rare mutations or transposable elements.

Identifying Highly Amplified Cells: EvaGreen Staining for Genome Quantification and Sorting

To quantify intracellular genome amplification, cells were stained with EvaGreen, a DNA-binding dye that fluoresces upon binding to double-stranded DNA. Cells were incubated with EvaGreen for 30 minutes to ensure sufficient dye uptake while minimizing background fluorescence. Following

incubation, excess dye was removed by centrifugation, and the cells were washed to prevent non-specific staining. Fluorescence intensity was measured using an Attune flow cytometer, allowing for precise quantification of amplified DNA within individual cells. Cells exhibiting high fluorescence—indicative of robust genome amplification—were identified and selected for sorting. Using fluorescence-activated cell sorting (FACS), highly amplified cells were isolated and deposited into individual wells of a 96-well plate for downstream processing. This approach enabled the targeted recovery of cells with successful intracellular amplification, improving the efficiency and accuracy of genomic analysis.

DNA Recovery and Sequencing Preparation: Capturing and Amplifying Biotinylated Genomes

Following intracellular amplification, cells were lysed using proteinase K to release genomic material. To inactivate the proteinase K, 5 μL of 100 μM PMSF (resuspended in isopropanol) was added to each tube immediately after heat treatment, followed by a 10-minute incubation at room temperature.

For selective capture of biotinylated DNA, 100 μL of resuspended streptavidin-coated C1 magnetic beads were added to each sample and incubated at room temperature for 60 minutes with agitation. The samples were then placed on a magnetic rack, and the supernatant was removed. Beads were resuspended in 250 μL of 1X wash solution, agitated for 5 minutes, and subjected to two consecutive wash steps. After the final wash, beads were resuspended in 250 μL of 10 mM Tris-HCl (pH 8.0) with 0.1% Tween-20. At this stage, beads could either be rinsed with molecular-grade water and processed immediately or stored overnight at 4°C in a Tris-Tween buffer.

Captured DNA was re-amplified using phi29 polymerase with random hexamers to generate sufficient material for sequencing. The amplified DNA was then prepared for Illumina sequencing, ensuring high-quality genomic data for downstream analysis.

Example Fixation and Permeabilization Permutations

a) Overnight Formaldehyde Fixation with Zymolyase and Tween-20 Permeabilization. At the time of sampling, 3 mL of yeast cultures, grown to saturation in YPD medium, were immediately spun down in a room-temperature centrifuge at 5000 g for 3 minutes. The cell pellet was fixed with 4% formaldehyde and incubated overnight at +4°C on a shaker. Fixed cells were centrifuged at 5000 g for 3 minutes, the supernatant was discarded, and the pellet was resuspended in ice-cold Buffer B (1.2 M sorbitol, 0.1 M potassium phosphate dibasic, pH 7.5). Zymolyase (1 mg/mL) was added, and cells were incubated at 30°C for 15 minutes. Permeabilization was performed with 0.1% Tween-20 at room temperature for 10 minutes before proceeding to downstream steps.

b) Overnight Formaldehyde Fixation with Zymolyase and TritonX-100 Permeabilization. At the time of sampling, 3 mL of yeast cultures, grown to saturation in YPD medium, were immediately spun down in a room-temperature centrifuge at 5000 g for 3 minutes. The cell pellet was fixed with 4% formaldehyde and incubated overnight at +4°C on a shaker. Fixed cells were centrifuged at 5000 g for 3 minutes, the supernatant was discarded, and the pellet was resuspended in ice-cold Buffer B (1.2 M sorbitol, 0.1 M potassium phosphate dibasic, pH 7.5). Zymolyase (1 mg/mL) was added, and cells were incubated at 30°C for 15 minutes. Permeabilization was performed with 0.1% TritonX-100 at room temperature for 10 minutes before proceeding to downstream steps.

c) 10 Minute Formaldehyde Fixation with Zymolyase and TritonX-100[PC1] Permeabilization ("INgen fix"). At the time of sampling, 3 mL of yeast cultures, grown to saturation in YPD medium, were immediately spun down in a room-temperature centrifuge at 5000 g for 3 minutes. The cell pellet was fixed with 4% formaldehyde for 10 minutes at room temperature (+20°C). Fixed cells were centrifuged at 5000 g for 3 minutes, the supernatant was discarded, and the pellet was resuspended in ice-cold Buffer B (1.2 M sorbitol, 0.1 M potassium phosphate dibasic, pH 7.5). Zymolyase (1

mg/mL) was added, and cells were incubated at 30°C for 15 minutes. Permeabilization was performed with 0.1% TritonX-100 at room temperature for 10 minutes before proceeding to downstream steps.

d) 30 Minute Formaldehyde Fixation with Zymolyase and Overnight Ethanol Permeabilization ("FISH fix"). At the time of sampling, 3 mL of yeast cultures, grown to saturation in YPD medium, were immediately spun down in a room-temperature centrifuge at 5000 g for 3 minutes. The cell pellet was fixed with 4% formaldehyde for 30 minutes at room temperature (+20°C) and then transferred to +4°C for overnight fixation on a shaker. Fixed cells were centrifuged at 5000 g for 3 minutes, the supernatant was discarded, and the pellet was resuspended in ice-cold Buffer B (1.2 M sorbitol, 0.1 M potassium phosphate dibasic, pH 7.5). Zymolyase (1 mg/mL) was added, and cells were incubated at 30°C for 15 minutes. The permeabilization step used ethanol treatment overnight at room temperature before downstream steps.

e) Ethanol Fixation and Permeabilization. At the time of sampling, 3 mL of Escherichia coli cultures, grown to the desired density, were immediately spun down in a room-temperature centrifuge at 5000 g for 3 minutes. The cell pellet was resuspended in 70% ethanol and incubated at room temperature (+20°C) for 10 minutes. Fixed cells were centrifuged again at 5000 g for 3 minutes, the supernatant was discarded, and the pellet was resuspended in ice-cold PBS. This washing step was repeated twice to ensure the removal of excess ethanol. Fixed cells were then processed directly for intracellular amplification or other downstream applications.

Additional Methods Considerations

Yeast cell culture. S. cerevisiae C5W4 WTC-SOD1-D102S was used. Cells were streaked on YP plus 2% dextrose agar plates from frozen -80°C glycerol stocks and grown at 30°C for 48 hours. Single colonies of (yeast strain) were inoculated into YP plus 2% dextrose liquid media (YPD) and grown at 30°C with shaking [200 rotations per minute (rpm)] for 24 hours.

Bacterial cell culture. E. coli DH5 was used. Cells were streaked on LB agar plates from frozen -80°C glycerol stocks and grown at 37°C for 24 hours. Single colonies of DH5 were inoculated into fresh LB medium and grown at 37°C with shaking [200 rotations per minute (rpm)] for 24 hours.

Cell Lysis. After amplification, the entire reaction volume was transferred into a microcentrifuge tube and 1 mL of cold PBS and 5 μL of 10% Triton X-100 was added. Samples were spun down at 4°C, 5000 g for 3 min. The supernatant was carefully aspirated off, leaving 30 μL to avoid removing the pellet. Cells were then resuspended in 1 mL cold PBS, spun down, and resuspended in 1 mL cold PBS for a total of two washes. After washing cells, the supernatant was aspirated off and cells were resuspended in 50 μL of cold PBS, 50 μL of 2X lysis buffer [20 mM Tris (pH 8.0), 400 mM NaCl, 100 mM EDTA (pH 8.0), and 4.4% SDS], and 10 μL of proteinase K solution (20 mg/mL). Cells were incubated at 55°C for 2 hours with periodic vortexing to lyse the cells and reverse the formaldehyde cross-links.

Preparing Streptavidin Beads for Sample Binding. To prepare beads for sample binding, they must first be washed. For each lysate, 44 μL of Dynabeads MyOne Streptavidin C1 (Invitrogen) were washed three times with 800 μL of a 1X wash solution of 5 mM Tris-HCl pH 8.0, 1 M NaCl, 500 μM EDTA, and 0.05% Tween-20 using a magnetic 1.5 mL tube rack. The beads were then resuspended in 100 μL per sample of a 2X wash solution containing 10 mM Tris-HCl pH 8.0, 2 M NaCl, and 1 mM EDTA.

Sample Binding to Streptavidin Beads. Directly after lysates were removed from heat, 5 μL of 100 μM PMSF (resuspended in isopropanol) was added to each tube and incubated at room temperature for 10 minutes to inactivate the proteinase K. To bind DNA to C1 beads, 100 μL of resuspended C1 beads were added to each sample tube and agitated at room temperature for 60 minutes. The

samples were then placed on the magnetic rack and the supernatant was removed. Sampled were removed from the magnetic rack and resuspended in 250 μL of the 1X wash solution and agitated at room temperature for 5 min. Samples were replaced on the magnetic rack and the subsequent steps were repeated for a total of two wash steps. After the two 1X wash steps, the supernatant was removed, and the samples were resuspended in 250 μL of 10 mM Tris-HCl pH 8.0 and 0.1% Tween-20. At this point, the beads could be rinsed with 250 μL of molecular-grade water while the beads were still bound to the magnetic rack and moved on to subsequent steps or resuspended in 250 μL of the Tris-HCL Tween-20 buffer and stored at 4°C overnight.

Off-the-beads Amplification. If the samples were stored overnight in the Tris-Tween buffer, the tubes were placed in a magnetic rack and the beads were rinsed with 250 μL of molecular-grade water. The bead-bound DNA was then amplified in 220 μL reactions with 2X high fidelity polymerase (KAPA HiFi) and 0.4 μM BC_0062 and BC_0108 primers[p5] (Supporting Information S2: Table2[p6]) for 3 min at 95°C, and then five cycles of 98°C for 20 s, 65°C for 45 s, and 72 for 3 min.

Quantitative PCR. The off-the-beads PCR product was then placed against a magnetic rack and the supernatant was transferred to new optical-grade PCR tubes with qPCR dye (EvaGreen 20X). The samples were then amplified on a qPCR machine for a further 10–20 cycles until the amplification curves exited the log-linear phase.

Size selection bead clean up. The PCR products were then cleaned using a 0.8X SPRI size selection and eluted in 20 μL of molecular-grade water.

Gel electrophoresis. 5 μL of the product which was eluted during the bead clean-up was then run on a 2% agarose gel at 120 V for 15-20 min. A single properly amplified gene should appear as a dark band on the gel at the correct size of the gene. The genome that has been properly amplified with random hexamers should appear as a smear starting at approximately 5–7 kB and ending at approximately 300 bp on a gel.

Funding

This work was supported by a National Institutes of Health grant R35GM133674 (to KGS), an Alfred P Sloan Research Fellowship in Computational and Molecular Evolutionary Biology grant FG-2021-15705 (to KGS), and a National Science Foundation Biological Integration Institution grant 2119963 (to KGS).

References

Aviel-Ronen S, Qi Zhu C, Coe BP, et al (2006) Large fragment Bst DNA polymerase for whole genome amplification of DNA from formalin-fixed paraffin-embedded tissues. BMC Genomics 7:312

Bagasra O (2007) Protocols for the in situ PCR-amplification and detection of mRNA and DNA sequences. Nat Protoc 2:2782–2795

Bagasra O, Hauptman SP, Lischner HW, et al (1992) Detection of human immunodeficiency virus type 1 provirus in mononuclear cells by in situ polymerase chain reaction. N Engl J Med 326:1385–1391

Behringer MG, Ho W-C, Miller SF, et al (2024) Trade-offs, trade-ups, and high mutational parallelism underlie microbial adaptation during extreme cycles of feast and famine. Curr Biol 34:1403–1413.e5

Blanco L, Bernad A, Lázaro JM, et al (1989) Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. J Biol Chem 264:8935–8940

Blattman SB, Jiang W, Oikonomou P, Tavazoie S (2020) Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. Nat Microbiol 5:1192–1201

Brettner L, Eder R, Schmidlin K, Geiler-Samerotte K (2024) An ultra high-throughput, massively multiplexable, single-cell RNA-seq platform in yeasts. Yeast 41:242–255

Cao J, Packer JS, Ramani V, et al (2017) Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 357:661–667

Chen C, Xing D, Tan L, et al (2017) Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI). Science 356:189–194

Clingenpeel S, Schwientek P, Hugenholtz P, Woyke T (2014) Effects of sample treatments on genome recovery via single-cell genomics. ISME J 8:2546–2549

Gasch AP, Yu FB, Hose J, et al (2017) Single-cell RNA sequencing reveals intrinsic and extrinsic regulatory heterogeneity in yeast responding to stress. PLoS Biol 15:e2004050

Gavrilov A, Razin SV (2009) Formaldehyde fixation of cells does not greatly reduce the ability to amplify cellular DNA. Anal Biochem 390:94–96

Georgieva D, Liu Q, Wang K, Egli D (2020) Detection of base analogs incorporated during DNA replication by nanopore sequencing. Nucleic Acids Res 48:e88

Gibellini D, Zauli G, Re MC, et al (1998) Flow cytometry in situ polymerase chain reaction as an innovative approach to HIV-1 proviral DNA detection. Eur J Histochem 42:137–141

Hashimoto M, Sakamoto H, Ido Y, et al (2018) In situ loop-mediated isothermal amplification (LAMP) for identification of Plasmodium species in wide-range thin blood smears. Malar J 17:235
Hully JR (1999) In situ PCR. In: PCR Applications. Elsevier, pp 169–194

Janiszewska M, Liu L, Almendro V, et al (2015) In situ single-cell analysis identifies heterogeneity for PIK3CA mutation and HER2 amplification in HER2-positive breast cancer. Nat Genet 47:1212–1219

Kirsch JM, Hryckowian AJ, Duerkop BA (2024) A metagenomics pipeline reveals insertion sequence-driven evolution of the microbiota. Cell Host Microbe 32:739–754.e4

Kuchina A, Brettner LM, Paleologu L, et al (2021) Microbial single-cell RNA sequencing by split-pool barcoding. Science 371.: https://doi.org/10.1126/science.aba5257

Laks E, McPherson A, Zahn H, et al (2019) Clonal decomposition and DNA replication states defined by scaled single-cell genome sequencing. Cell 179:1207–1221.e22

Lee CIP, Leong SH, Png AEH, et al (2006) An isothermal method for whole genome amplification of fresh and degraded DNA for comparative genomic hybridization, genotyping and mutation detection. DNA Res 13:77–88

Lusk RW (2014) Diverse and widespread contamination evident in the unmapped depths of high throughput sequencing data. PLoS One 9:e110808

Lux M, Hammer B, Sczyrba A (2015) Automated contamination detection in single-cell sequencing. bioRxiv

Martin BK, Qiu C, Nichols E, et al (2023) Optimized single-nucleus transcriptional profiling by combinatorial indexing. Nat Protoc 18:188–207

Müller CA, Boemo MA, Spingardi P, et al (2019) Capturing the dynamics of genome replication on individual ultra-long nanopore sequence reads. Nat Methods 16:429–436

Muro-Cacho CA (1997) In situ PCR. Overview of procedures and applications. Front Biosci 2:c15–29

Murray GI (1993) In situ PCR. J Pathol 169:187–188 Muyas F, Sauer CM, Valle-Inclán JE, et al (2024) De novo detection of somatic mutations in high-throughput single-cell profiling data sets. Nat Biotechnol 42:758–767

Navin N, Hicks J (2011) Future medical applications of single-cell sequencing in cancer. Genome Med 3:31

Nawy T (2014) Single-cell sequencing. Nat Methods 11:18

Nuovo GJ (1997) PCR in Situ Hybridization: Protocols and Applications. Lippincott Williams & Wilkins

O'Leary JJ, Browne G, Landers RJ, et al (1994) The importance of fixation procedures on DNA template and its suitability for solution-phase polymerase chain reaction and PCR in situ hybridization. Histochem J 26:337–346

Payne AC, Chiang ZD, Reginato PL, et al (2021) In situ genome sequencing resolves DNA sequence and structure in intact biological samples. Science 371.: https://doi.org/10.1126/science.aay3446

Pereira AC, Tenreiro A, Cunha MV (2022) When FLOW-FISH met FACS: Combining multiparametric, dynamic approaches for microbial single-cell research in the total environment. Sci Total Environ 806:150682

Rosenberg AB, Roco CM, Muscat RA, et al (2018) Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science 360:176–182

Salatino R, Franco M, Romero-Toledo, et al (2024) Cell sorting based on single nucleotide variation enables characterization of mutation-dependent transcriptome and chromatin states. bioRxiv.

https://doi.org/10.1101/2024.07.05.602247

Schmidlin K, Apodaca S, Newell D, et al (2024) Distinguishing mutants that resist drugs via different mechanisms by examining fitness tradeoffs across hundreds of fluconazole-resistant yeast strains. Elife. https://doi.org/10.7554/elife.94144.1

Speel EJ, Hopman AH, Komminoth P (1999) Amplification methods to increase the sensitivity of in situ hybridization: play card(s). J Histochem Cytochem 47:281–288

Tang Y-W, Lozano L, Chen X, et al (2020) An Isothermal, Multiplex Amplification Assay for Detection and Genotyping of Human Papillomaviruses in Formalin-Fixed, Paraffin-Embedded Tissues. J Mol Diagn 22:419–428

Tani K, Kurokawa K, Nasu M (1998) Development of a direct in situ PCR method for detection of specific bacteria in natural environments. Appl Environ Microbiol 64:1536–1540

Venkataram S, Dunn B, Li Y, et al (2016) Development of a Comprehensive Genotype-to-Fitness Map of Adaptation-Driving Mutations in Yeast. Cell 166:1585–1596.e22

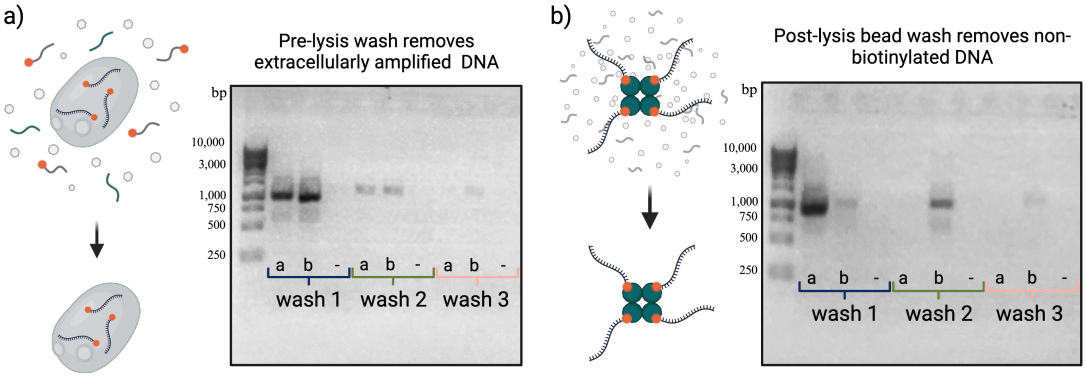
Vitak SA, Torkenczy KA, Rosenkrantz JL, et al (2017) Sequencing thousands of single-cell genomes with combinatorial indexing. Nat Methods 14:302–308

Wagner M, Schmid M, Juretschko S, et al (1998) In situ detection of a virulence factor mRNA and 16S rRNA inListeria monocytogenes. FEMS Microbiol Lett 160:159–168

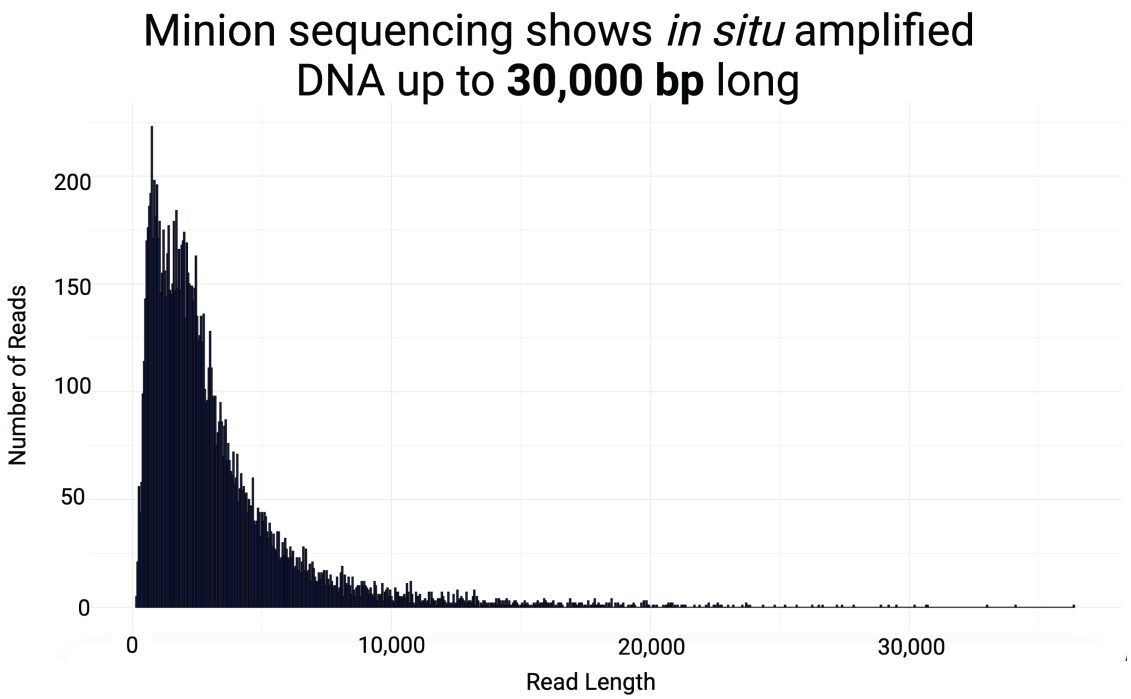
Yin Y, Jiang Y, Lam K-WG, et al (2019) High-Throughput Single-Cell Sequencing with Linear Amplification. Mol Cell 76:676–690.e10

Zion S, Katz S, Hershberg R (2024) Escherichia coli adaptation under prolonged resource exhaustion is characterized by extreme parallelism and frequent historical contingency. PLoS Genet 20:e1011333

SUPPLEMENT



Supplemental Figure 1: Intracellular Amplification and Recovery of Amplified DNA. a) Pre-lysis wash removes bacterial 'contaminating' DNA that we spiked in at very high levels. **b)** Post-lysis wash removes non-biotinylated DNA representing the native genome.



Supplemental Figure 2: Sequencing reads recovered can exceed 30KB, even after fragmentation, suggesting INgen amplifies long pieces of DNA intracellularly.

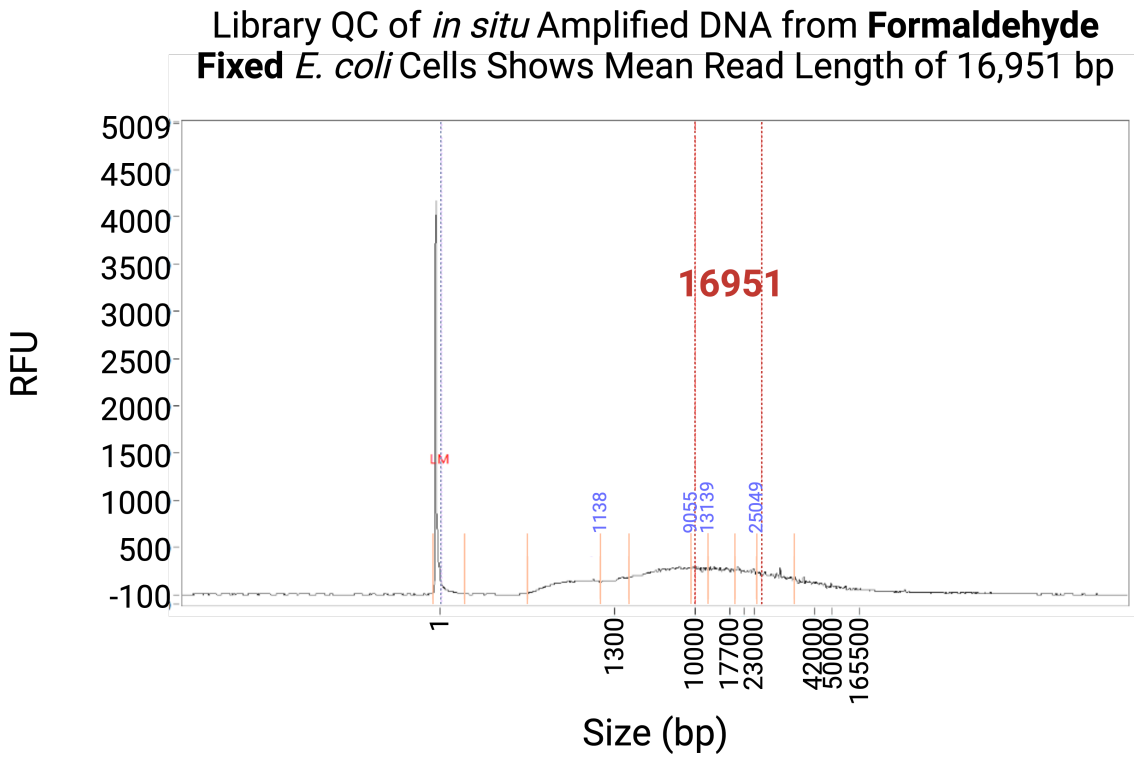


Figure S3: Sequencing reads recovered are on average 20KB, spanning over 100KB, suggesting INgen amplifies long pieces of DNA intracellularly.

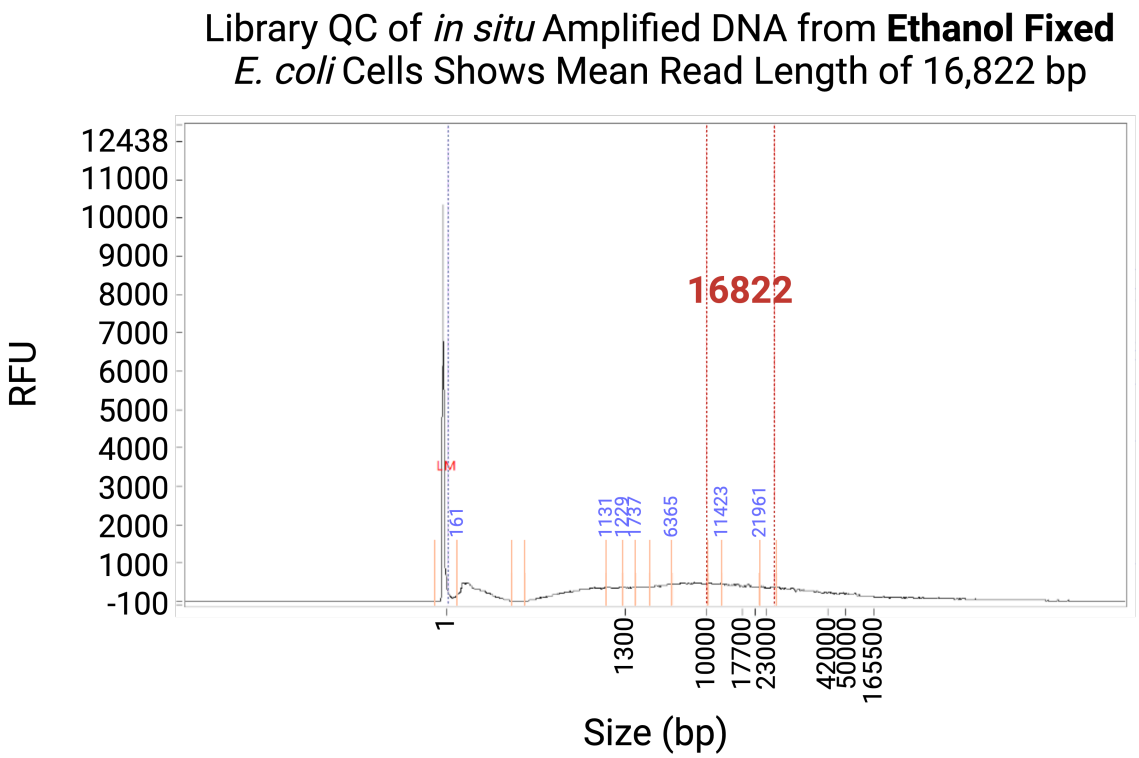


Figure S4: Sequencing reads recovered are on average 20KB, spanning over 100KB, suggesting INgen amplifies long pieces of DNA intracellularly.