Generation of gene-level resolution chromosome contact maps in bacteria and archaea

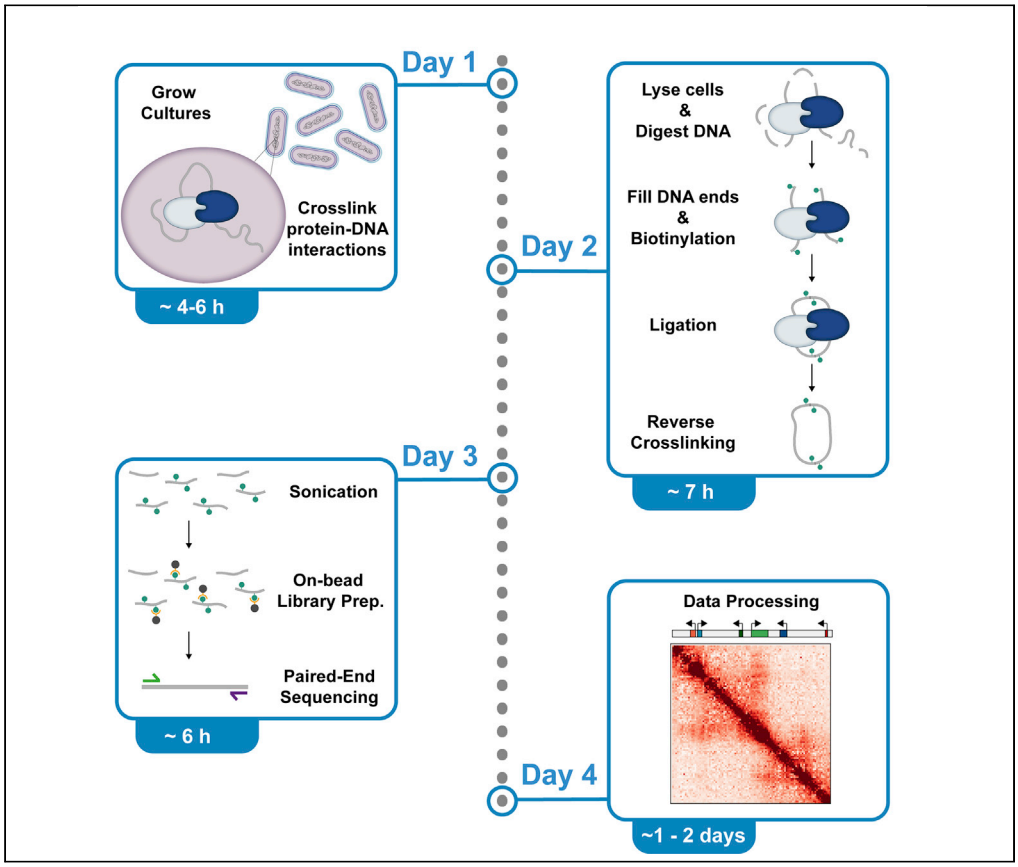
Charlotte Cockram, Agnès Thierry, Romain Koszul

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Generation of gene-level resolution chromosome contact maps in bacteria and archaea



Chromosome conformation capture (Hi-C) has become a routine method for probing the 3D organization of genomes. However, when applied to bacteria and archaea, current protocols are expensive and limited in their resolution. By dissecting the different steps of published eukaryotic and prokaryotic Hi-C protocols, we have developed a cost- and time-effective approach to generate high-resolution (down to 500 bp – 1 kb) contact matrices of both bacteria and archaea genomes.

Charlotte Cockram, Agnès Thierry, Romain Koszul

romain.koszul@pasteur.fr (R.K.)
agnes.thierry@pasteur.fr (A.T.)

Highlights

Optimized Hi-C protocol for archaeal and bacterial genomes

Generation of genome-wide contact maps up to 1 kb resolution

Detailed description of steps from cell fixation to sequencing library preparation

A cost- and time-effective approach offering gene-level resolution contact maps

Cockram et al., STAR Protocols *2*, 100512
June 18, 2021
https://doi.org/10.1016/j.xpro.2021.100512

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Generation of gene-level resolution chromosome contact maps in bacteria and archaea

Charlotte Cockram,[1] Agnès Thierry,[1,2,*] and Romain Koszul[1,3,*]

[1]Institut Pasteur, Unité Régulation Spatiale des Génomes, CNRS UMR 3525, 5015 Paris, France

[2]Technical contact

[3]Lead contact

*Correspondence: romain.koszul@pasteur.fr (R.K.), agnes.thierry@pasteur.fr (A.T.)
https://doi.org/10.1016/j.xpro.2021.100512

SUMMARY

Chromosome conformation capture (Hi-C) has become a routine method for probing the 3D organization of genomes. However, when applied to bacteria and archaea, current protocols are expensive and limited in their resolution. By dissecting the different steps of published eukaryotic and prokaryotic Hi-C protocols, we have developed a cost- and time-effective approach to generate high-resolution (down to 500 bp – 1 kb) contact matrices of both bacteria and archaea genomes.

For complete details on the use and execution of this protocol, please refer to Cockram et al. (2020).

BEFORE YOU BEGIN

Timing: approximately 30 min

Considerations for designing the Hi-C experiment

1. Choosing the correct restriction enzyme:

The frequency of restriction sites correlates with the GC-content of a genome and thus the choice of restriction enzymes can introduce significant distortions to 3C/Hi-C contact maps (Cournac et al., 2012; Marbouty et al., 2017; Yaffe and Tanay, 2011). For this reason, eukaryotic Hi-C protocols now typically use the 4-cutter DpnII (^GATC), that offers a relatively uniform distribution of restriction sites along the genome (Rao et al., 2014). This is also what we use and recommend for prokaryotes (Marbouty and Koszul, 2017; Marbouty et al., 2015). However, DpnII is not an adequate solution for most bacterial species, whose genomes are subject to the Dam methylation of GATC sites. It is therefore important to consider other enzymes that will provide sufficient restriction positions within a particular genome, to reach at a proper resolution.

We recommend – as of today – using HpaII (C^CGG) with NEBuffer 1 (10× NEBuffer 1: (200 mM Tris-HCl pH 7.5, 100 mM MgCl2, 10 mM DTT, 1 mg/mL BSA) for most applications. In case the investigated genome is particularly AT-rich, or if a specific AT-rich region is under investigation, then we suggest using MluCI (^AATT) with the same buffer.

The enzyme used will determine if either Biotin-14-dATP or Biotin-14-dCTP is used for the subsequent biotinylation step of the protocol.

Preparation of solutions and buffers

Timing: approximately 1–2 h

2. Prepare cell culture medium and all solutions listed below in the "materials and equipment" section. Ensure there are enough for the number of samples being processed.
3. All experiments should be performed with fresh formaldehyde solution (<1 week). We recommend purchasing 25 mL bottles to limit waste.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
2 ml VK05 microorganism Precellys tubes	Bertin Instruments	Cat# P000913-LYSK0-A
microTUBE AFA Fiber Pre-Slit Snap-Cap	Covaris	Cat# 520045
100 mM dNTPs (set of 4)	Dutscher	Cat# 755086
Proteinase K (20 mg/mL)	Eurobio Scientific	Cat# GEXPRK01-B5
Ethidium Bromide Solution	Eurobio Scientific	Cat# GEPBET02-AF
RNase A	Euromedex	Cat# RB0473
36.5%–38% Formaldehyde solution	Sigma-Aldrich	Cat# F8775
Glycine	Sigma-Aldrich	Cat# G8898
1 M $MgCl_2$	Sigma-Aldrich	Cat# M1028
DL-Dithiothreitol solution	Sigma-Aldrich	Cat# 43816
Bovine Serum Albumin	Sigma-Aldrich	Cat# A7906
ATP	Sigma-Aldrich	Cat# 10519987001
Sodium acetate	Sigma-Aldrich	Cat# S2889
Phenol:Chloroform:Isoamyl Alcohol 25:24:1 (pH 8.0)	Sigma-Aldrich	Cat# P2069
cOmplete Mini EDTA-free protease inhibitor cocktail	Sigma-Aldrich	Cat# 11836170001
T4 DNA ligase (30 U/ul)	Thermo Fisher Scientific	Cat# EL0013
20% SDS solution	Thermo Fisher Scientific	Cat# 10607633
Triton-X-100	Thermo Fisher Scientific	Cat# 10671652
1 M Tris-Hcl pH7.5	Thermo Fisher Scientific	Cat# 10573145
0.5 M EDTA (pH 8.0)	Thermo Fisher Scientific	Cat# 10182903
Absolute EtOH	Thermo Fisher Scientific	Cat# 10680993
10× PBS solution	Thermo Fisher Scientific	Cat# 10649743
UltraPure Agarose	Thermo Fisher Scientific	Cat# 16500500
50× TAE Solution	Thermo Fisher Scientific	Cat# 10490264
Biotin-14-dCTP	Thermo Fisher Scientific	Cat# 19518018
Biotin-14-dATP	Thermo Fisher Scientific	Cat# 19524016
NaCl	Thermo Fisher Scientific	Cat# 10616082
Tween 20	Thermo Fisher Scientific	Cat# 10113103
Dynabeads MyOne Streptavidin C1	Thermo Fisher Scientific	Cat# 10202333
GeneRuler 1 Kb Plus DNA ladder	Thermo Fisher Scientific	Cat# SM1331
Orange DNA Loading Dye	Thermo Fisher Scientific	Cat# R0631
Critical commercial assays		
Agencourt AMPure XP beads	Beckman Coulter	Cat# A63881
Qubit dsDNA HS Assay Kit	Thermo Fisher Scientific	Cat# 10606433
NextSeq 500/550 v2.5 High Output Kit (75 Cycles)	Illumina	Cat# 20024906
TruSeq DNA CD Indexes (96 Indexes)	Illumina	Cat# 20015949
HpaII	New England Biolabs	Cat# R0171M
MluCI	New England Biolabs	Cat# R0538L

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
DNA Polymerase I, Large (Klenow) Fragment	New England Biolabs	Cat# M0210L
10× NEBuffer 2	New England Biolabs	Cat# B7002S
Quick Ligation Kit	New England Biolabs	Cat# M2200L
10× T4 DNA Ligase Reaction Buffer	New England Biolabs	Cat# B0202S
Phusion High-Fidelity PCR Master Mix with HF Buffer	New England Biolabs	Cat# M0531L
T4 Polynucleotide Kinase	New England Biolabs	Cat# M0201L
T4 DNA Polymerase	New England Biolabs	Cat# M0203L
Klenow Fragment (3'–5'exo-)	New England Biolabs	Cat# M0212L
Software and algorithms		
Bowtie 2	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
SCN normalization procedure	(Cournac et al., 2012)	https://github.com/koszullab/hicstuff
R	(RStudio Team, 2020)	https://rstudio.com
Chromosight	(Matthey-Doret et al., 2020)	https://github.com/koszullab/chromosight
HTseq - count	(Anders et al., 2015)	https://github.com/htseq/htseq
DESeq2	(Love et al., 2014)	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
SAMtools	(Li et al., 2009)	http://www.htslib.org
bamCoverage (deepTools)	(Ramírez et al., 2016)	https://github.com/deeptools/deepTools/blob/master/docs/content/tools/bamCoverage.rst
Deposited data		
Codes and functions for Hi-C analysis	(Matthey-Doret et al., 2020)	https://github.com/koszullab/hicstuff
Raw data and processed data	(Cockram et al., 2020)	PRJNA587586
Other		
0.22 μm Filter, Nalgene filtration	Thermo Scientific	596-3320
Precellys Evolution tissue homogenizer	Bertin Instruments	P000062-PEVO0-A
Covaris S220	Covaris	500217
NEXTFLEX® DNA Barcodes Kits 48 (384 rxns)	PerkinElmer	NOVA 5141-13
Experimental models: organisms/strains		
E. coli K12 MG1655 (*F- lambda- ilvG- rfb-50 rph-1*)	(Blattner et al., 1997)	N/A
H. volcanii H26 ΔpyrE2	(Allers et al., 2004)	N/A

MATERIALS AND EQUIPMENT

Solutions

Glycine Solution		
Reagent	Final concentration	Amount
Glycine	2.5 M	187.7 *g*
ddH_2O	n/a	Up to 1 L

Note: Filter sterilize through a 0.22 μm filter, store (we use glycine solution stored for a maximum of 6 months; this indication and the others below are based on our experience, and doesn't exclude longer conservation time: we typically consume all of these buffers before the indicated time) at 19°C–23°C.

BSA solution		
Reagent	Final concentration	Amount
BSA	10 mg/mL	0.2 *g*
ddH_2O	n/a	Up to 20 mL

Note: Filter sterilize through a 0.22 μm filter, aliquot into 1 mL tubes, store at −20°C (we use BSA solution for a maximum of 6 months).

ATP solution		
Reagent	Final concentration	Amount
ATP	100 mM	1 g
NaOH (1 M)	n/a	~ 1.6 mL
ddH_2O	n/a	Up to 16.5 mL

Note: Volume of NaOH indicated is a guide, verify pH is 7.0 before adjusting to final volume with ddH_2O. Filter sterilize through a 0.22 μm filter, aliquot into 1 mL tubes, store at −20°C (we use ATP solution storedfor a maximum of 6 months).

10× digestion buffer		
Reagent	Final concentration	Amount
Tris-HCl (1 M, pH7.5)	200 mM	10 mL
$MgCl_2$(1 M)	100 mM	5 mL
DTT (1 M)	10 mM	0.5 mL
BSA solution (10 mg/mL)	1 mg/mL	5 mL
ddH_2O	n/a	Up to 50 mL

Note: Filter sterilize through a 0.22 μm filter, store at 19°C–23°C (we use TE stored up to 6 months).

1× TE solution		
Reagent	Final concentration	Amount
Tris-HCl (1 M, pH7.5)	10 mM	5 mL
EDTA (0.5 M)	1 mM	1 mL
ddH_2O	n/a	Up to 500 mL

Note: Filter sterilize through a 0.22 μm filter, aliquot into 5 mL tubes, store at −20°C (we use digestion buffer stored up to 6 months).

10× ligation buffer		
Reagent	Final concentration	Amount
Tris-HCl (1 M, pH7.5)	500 mM	25 mL
$MgCl_2$(1 M)	100 mM	5 mL
DTT (1 M)	100 mM	5 mL
ddH_2O	n/a	Up to 50 mL

Note: Filter sterilize through a 0.22 μm filter, aliquot into 5 mL tubes, store at −20°C (we use ligation buffer stored up to 6 months).

1× Tween Washing Buffer (TWB)		
Reagent	Final concentration	Amount
Tris-HCl (1 M, pH7.5)	5 mM	0.5 mL
EDTA (0.5 M)	0.5 mM	0.1 mL
NaCl (5 M)	1 M	20 mL
Tween-20	0.05%	50 μL
ddH_2O	n/a	Up to 100 mL

Note: Filter sterilize through a 0.22 μm filter, store at 19°C–23°C (we use TWB solution stored up to a year).

2× Binding Buffer (BB)		
Reagent	Final concentration	Amount
Tris-HCl (1 M, pH7.5)	10 mM	1 mL
EDTA (0.5 M)	1 mM	0.2 mL
NaCl (5 M)	2 M	40 mL
ddH_2O	n/a	Up to 100 mL

Note: Filter sterilize through a 0.22 μm filter, store at 19°C–23°C (we use binding buffer stored up to 6 months).

10 mM dNTP mix		
Reagent	Final concentration	Amount
dATP (100 mM)	10 mM	20 μL
dTTP (100 mM)	10 mM	20 μL
dGTP (100 mM)	10 mM	20 μL
dCTP (100 mM)	10 mM	20 μL
ddH_2O	n/a	Up to 200 μL

Note: Store at −20°C (we use dNTP solution stored up to a year).

10 mM dAGTTP mix		
Reagent	Final concentration	Amount
dATP (100 mM)	10 mM	20 μL
dGTP (100 mM)	10 mM	20 μL
dTTP (100 mM)	10 mM	20 μL
ddH_2O	n/a	Up to 200 μL

Note: Store at −20°C (we use dAGTTP solution stored up to a year).

10 mM dCGTTP mix		
Reagent	Final concentration	Amount
dCTP (100 mM)	10 mM	20 μL
dGTP (100 mM)	10 mM	20 μL
dTTP (100 mM)	10 mM	20 μL
ddH_2O	n/a	Up to 200 μL

Note: Store at −20°C (we use dCGTTP solution stored up to a year).

10 mM dATP mix		
Reagent	Final concentration	Amount
dATP (100 mM)	10 mM	20 μL
ddH_2O	n/a	Up to 200 μL

Note: Store at −20°C (we use a dATP solution stored up to a year).

STEP-BY-STEP METHOD DETAILS

Day 1: Cell fixation

Timing: approximately 4–6 h

The subsequent steps describe a general overview of the growth and fixation of bacterial and archaeal cultures.

1. Grow bacterial or archaeal culture in an appropriate medium and at the required temperature until you have a total of ~1 × 10^8 cells growing in the exponential growth phase.

   Note: For *E. coli*, *Hbt. salinarum*, *H. volcanii*, and *T. kodakarensis* we usean OD_{600nm} of ~0.2 – 0.25, but this can be adjusted for other growth phases, whilst maintaining of ~1 × 10^8 cells. Note that trial and error will be needed for species requiring specific growth conditions. For instance, for hyperthermophile archaea crosslinking cannot be performed at high temperatures, and those cells will have to be cooled down immediately prior to fixation (see Cockram et al., 2020; Takemata et al., 2019). We recommend paying attention to this step, as it is a crucial one with respect to the quality of the resulting data.

2. Crosslink protein-DNA interactions by adding fresh formaldehyde (3% final concentration)
3. Incubate for 30 min at 19°C–23°C on an orbital shaker (50–100 rpm).

   ⚠ CRITICAL: Always use fresh formaldehyde for crosslinking.

   Optional: Crosslinking can also be performed using a magnetic stirrer.

   ⚠ CRITICAL: When crosslinking thermophilic organisms, such as the Euryarchaeota *T. kodakarensis*, we recommend mixing 20 mL of the cell culture with 80 mL of room temperature formaldehyde (19°C–23°C) diluted in 1 X PBS (4% final concentration, Cockram et al., 2020) see also (Takemata et al., 2019).

4. Quench formaldehyde by adding glycine (0.5 M final concentration).
5. Incubate for 20 min at 19°C–23°C on an orbital shaker (50–100 rpm).

   ⚠ CRITICAL: When working with halophilic organisms, such as the Euryarchaeota *H. volcanii*, which are very sensitive to salt concentrations, we recommend preparing 2.5M glycine in 18% (w/v) saltwater solution (Cockram et al., 2020).

6. Transfer culture to a sterile centrifuge tube and collect cells by centrifugation (4000. × *g*, 10 min, at 19°C–23°C).
7. Decant supernatant.

⚠ CRITICAL: Depending on the bacterial or archaeal species, pellets can be unstable and the supernatant may need to be removed by careful pipetting.

8. Carefully resuspend the cell pellet in 25 mL of 1× PBS by gentle pipetting and transfer to a 50 mL conical tube.
9. Collect cells by centrifugation (4000×*g*, 10 min at 19°C–23°C).
10. Resuspend the pellet in 1 mL of 1× PBS by gentle pipetting. Transfer to a 1.5 mL tube.
11. Centrifuge again (4000 × *g*, 5 min at 19°C–23°C).
12. Carefully remove the supernatant using a pipette, so that the pellet is as dry as possible.
13. Store the cell pellet at −80°C.

Pause point: Pellets can be stored for at least 2 years at −80°C with no impact on the quality of the subsequent Hi-C library.

Day 2: Hi-C library construction

Timing: approximately 7 h

14. Remove the cell pellet from the −80°C freezer and place on ice until completely thawed.
15. Gently resuspend the pellet by pipetting in 1.2 mL of 1× TE containing complete protease inhibitor cocktail.
16. Transfer to a 2 mL VK05 Precellys tube.
17. Lyse cells using the Precellys Evolution tissue homogenizer (V7500: 5 × 30 s, 20 s pause).

Note: For Precellys machines without the Cryolys cooling attachment, the program is: V6700: 9 × 20s, (30s pause). Tubes should be removed every 3 cycles and placed on ice for 5 min to prevent sample degradation from overheating.

⚠ CRITICAL: For Gram-negative species, such as *E. coli*, lysis can also be performed by incubating the cells with lysozyme (0.4 µg/mL final concentration) for 30 min at 37°C. This will produce similar results to mechanical disruption. However, formaldehyde-crosslinked Gram-positive bacteria, such as *B. subtilis*, can be difficult to lyse with lysozyme and the enzyme is inactive on archaea, since they do not contain peptidoglycan in their cell walls. In such cases, we recommend either disrupting cells with glass beads or a French Press (Marbouty et al., 2014).

18. Transfer the lysate (~1 mL in volume) to a 5 mL microcentrifuge tube. Take care not to transfer any of the glass beads.
19. Add 50 µL of 10% SDS solution (0.5% final concentration) and incubate for 10 min at 19°C–23°C.
20. Prepare the Digestion Mix:

Reagent	Volume
10× Digestion Buffer	500 µL
10% Triton-X-100	500 µL
ddH_2O	3 mL

Note: If processing multiple samples, we recommend preparing a master mix.

21. Add the Digestion Mix to the tube containing the lysate and mix well by pipetting.
22. Remove 400 µL of the sample and transfer it to a 1.5 mL microcentrifuge tube.

Note: This is the non-digested (**ND**) control to check for sample degradation and should be stored on ice until required (troubleshooting 1 Figure 1).

23. Add 1000U of restriction enzyme to the remaining sample and incubate at 37°C for 3 h in a shaking incubator (~180 rpm).
24. Following digestion, remove 400 µL of the sample and transfer it to a 1.5 mL microcentrifuge tube.

Note: This is the digested (**D**) control to check for efficient cutting of the genomic DNA. Store on ice until required (troubleshooting 2 Figure 1).

25. Pellet the insoluble fraction by centrifugation (16,000 × *g*, 20 min at 19°C–23°C).

Note: This pellet contains protein-DNA complexes of interest

26. Decant the supernatant.
27. Carefully resuspend the pellet in 400 µL of dH2O by pipetting.

Note: Depending on the species, this pellet can be quite difficult to resuspend. Ensure it is fully resuspended before continuing.

28. Prepare the Biotinylation mix:

Reagent	Volume
10× Ligation Buffer	50 µL
10 mM dAGTTP mix	4.5 µL
Biotin-14-dCTP	37.5 µL
DNA Polymerase I – Large (Klenow) Fragment (5U/µL)	8 µL

Note: If processing multiple samples, we recommend preparing a master mix.

Note: If using MluCI for the digestion, exchange Biotin-14-dCTP for Biotin-14-dATP and replace the 10 mM dAGTTP mix with 10 mM dCGTTP mix.

29. Add the Biotinylation mix to the resuspended pellet and mix by pipetting
30. Incubate at 37°C for 45 min in a shaking incubator (~180 rpm).
31. Prepare the ligation mix:

Reagent	Volume
10× Ligation Buffer	120 µL
10 mg/mL BSA	12 µL
100 mM ATP	12 µL
ddH_2O	540 µL
T4 DNA ligase (30U/µL)	16 µL

Note: If processing multiple samples, we recommend preparing a master mix.

32. Add ligation mix to the reaction and mix by pipetting.
33. Incubate with gentle agitation for 3 h at 19°C–23°C.

34. Following ligation, add 20 μL 500 mM EDTA, 80 μL 10% SDS, and 100 μL 20 mg/mL proteinase K to the Hi-C library.
35. Remove the ND and D controls from ice and add 20 μL 500 mM EDTA, 20 μL 10% SDS, and 10 μL 20 mg/mL proteinase K.
36. Incubate the Hi-C library, ND and D controls overnight (10–12 h) at 65°C to reverse formaldehyde-mediated protein-DNA crosslinks.

Note: we processed indiscriminately bacteria and archaea cells at this step, regardless of their normal growth condition.

Day 3: DNA purification and sequencing library preparation

Timing: approximately 6 h

37. Purification of DNA by phenol extraction

Timing: approximately 1 h

a. Add an equal volume (~1.2 mL) of Phenol:Chloroform:Isoamyl alcohol to the Hi-C sample and vortex for 30 s.

⚠ CRITICAL: Always handle phenol under the hood, as it causes severe burns and eye damage and is toxic if inhaled or ingested.

b. Centrifuge (12,000 × *g*, 5 min at 19°C–23°C).
c. Carefully remove the upper aqueous phase (~1.2 mL) and transfer to a new 5 ml microcentrifuge tube.
d. Precipitate DNA by adding 2.5× volume (~3 mL) of ice-cold 100% EtOH and 1/10 volume (~120 μL) of 3 M NaOAc (pH 5.0).
e. Precipitate DNA of ND and D controls by adding 2.5× volume (~1 mL) of ice-cold 100% EtOH and 1/10 volume (~ 40 μL) of 3 M NaOAc (pH 5.0).

Note: to reduce the amount of phenol chloroform used, we proceed straight to ethanol precipitation for the ND and D controls.

f. Incubate all samples at −80°C for 30 min.
g. Pellet DNA by centrifugation (12,000 × *g*, 20 min, 4°C).
h. Carefully remove supernatant using a pipette
i. Wash pellets with 500 μL of ice-cold 70% EtOH.
j. Pellet DNA by centrifugation (12,000 × *g*, 5 min, 4°C).
k. Remove EtOH and dry pellets on a 37°C heat block for 5–10 min.

Note: Briefly centrifuge (~10 sec) after removing the ethanol to collect any residual EtOH on the side of the tube.

l. Add 140 μL 1× TE buffer containing 1 mg/mL RNase to pellets and incubate for 15–30 min at 37°C in a shaking incubator (~180 rpm).
m. Sample 10 μL for the ligation (L) control (troubleshooting 3 Figure 1).
n. Run 10 μL of the Hi-C library and 10 μL of the ND, D and L controls on a 1% agarose gel (Figure 1).

⚠ CRITICAL: These controls are important for determining that the Hi-C library construction was successful.

o. Discard ND and D controls.

Pause point: Hi-C library can be stored at −20°C.

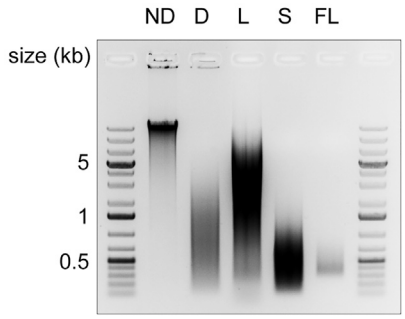


Figure 1. Expected results from each step of the Hi-C and sequencing library preparation

Representative image from the preparation of an *E. coli* Hi-C bank and subsequent sequencing library. The non-digested (ND) sample is a control for degradation during the cell lysis and restriction digest steps of the Hi-C library preparation. Following digestion, an aliquot is taken (D) to check for efficient cutting of genomic DNA by the restriction enzyme (D). The DNA is then subject to end-filling with a biotinylated nucleotide and blunt-end ligation, the efficiency of these steps can be determined by the checking that the DNA smear seen in the digested control has shifted higher molecular weight in the Ligation (L) control. Following reverse crosslinking and DNA purification, the DNA is sonicated to yield an average fragment size of 300 bp (S). The Hi-C library is then prepared for Illumina sequencing, with the size and quality of final library (FL) then checked before paired-end sequencing.

38. Sonication and size-selection of DNA

Timing: approximately 45 min

a. Transfer 130 μL of DNA to a Covaris microtube.

Note: We recommend a maximum of 5 μg of DNA is used for sonication (as determined by Qubit analysis). Samples that have less DNA are OK, but if the Hi-C library exceeds 5 μg, then an aliquot should be taken and the remaining DNA stored at −20°C, as a backup. Once the correct quantity of DNA is obtained, the sample volume should be adjusted to 130 μL using 1× TE buffer.

b. Sonicate DNA using a Covaris S220 machine and the following parameters:

Parameter	Value
Peak Incidence Power (W)	140
Duty Factor	10%
Cycles per Burst	200
Temperature	7°C
Treatment Time	80 sec

Note: These parameters will generate a fragment size of ~300 bp.

c. Transfer sheared DNA to a 1.5 mL microcentrifuge tube.
d. Add an equal volume of AMPure XP beads.
e. Mix sample 10× by gentle pipetting.
f. Incubate for 5 min at 19°C–23°C to allow DNA fragments to bind to the magnetic beads.
g. Transfer tube to a magnetic rack to allow the beads to separate to the side of the tube (~ 1 min).
h. Carefully remove the supernatant and wash with 70% EtOH, keeping the tube on the magnet and ensuring that the beads are not disturbed.
i. Repeat step h.
j. Remove the EtOH, and airdry the pellet for 2 min to remove residual EtOH.

⚠ CRITICAL: do not allow the pellet to over-dry as this will reduce the elution efficiency.

k. Remove beads from the magnet and resuspend in 320 μL of 10 mM Tris-HCl (pH 7.5).
l. Mix sample 10× by gentle pipetting.
m. Incubate for 5 min at 19°C–23°C and then place back onto the magnetic rack for 1 min.

n. Carefully remove supernatant and transfer to a new 1.5 mL microcentrifuge tube.
o. Check sonication and size-selection by running 18 μL of the DNA on a 1% agarose gel (sonication control (S), troubleshooting 4 Figure 1).
p. The remaining 300 μL is used to prepare the sequencing library.

Pause point: DNA can be stored at −20°C.

39. Biotin pull-down

Timing: approximately 30 min

a. Thoroughly mix the Streptavidin C1 Dynabeads before use.
b. Transfer 30 μL of beads to a 1.5 ml microcentrifuge tube.
c. Place the tube on a magnet and wait ~1 min to clear the supernatant.
d. Remove the supernatant and then resuspend the beads in 500 μL of 1× TWB.
e. Place the tube back on the magnet and wait ~1 min to clear the supernatant
f. Remove the supernatant and resuspend the beads in 300 μL of 2× BB.
g. Add 300 μL of the Hi-C sample to the beads.
h. Incubate on a tube rotator (~10 rpm) for 15 min at 19°C–23°C.
i. Place the tube back on the magnet and wait ~1 min to clear the supernatant.
j. Remove the supernatant and then resuspend the beads in 500 μL of 1× TWB.
k. Incubate in a thermomixer (2 min, 55°C, 1,000 rpm).
l. Place the tube back on the magnet and wait ~1 min for the supernatant to clear.
m. Repeat steps j–l.
n. Discard the supernatant, resuspend the beads in 100 μL of 1× T4 ligase buffer and transfer to a new 1.5 mL microcentrifuge tube.

40. End repair

Timing: approximately 45 min

a. Prepare the end-repair mix by combining the following:

Reagent	Volume
1× T4 ligase buffer	85 μL
10 mM dNTPs	5 μL
T4 Polynucleotide Kinase (10U/μL)	5 μL
T4 DNA polymerase (3U/μL)	4 μL
DNA Polymerase I, Large (Klenow) Fragment (5U/μL)	1 μL

Note: If processing multiple samples, we recommend preparing a master mix.
b. Place the Hi-C library on the magnet and wait ~1 min to clear the supernatant.
c. Discard the supernatant and resuspend the beads in the end-repair mix.
d. Incubate for 30 min at 19°C–23°C, without agitation.
e. Place the tube on the magnet and wait ~1 min to clear the supernatant.
f. Remove the supernatant and then resuspend the beads in 500 μL of 1× TWB.
g. Incubate in a thermomixer (2 min, 55°C, 1,000 rpm).
h. Place the tube back on the magnet and wait 1 min for the supernatant to clear.
i. Repeat steps f–h.
j. Discard the supernatant, resuspend the beads in 100 μL of 1× NEB Buffer 2, and transfer to a new 1.5 mL microcentrifuge tube.

Pause point: DNA can be stored at −20°C.

41. A-tailing

Timing: approximately 45 min

a. Prepare the A-tailing mix by combining the following:

Reagent	Volume
1× NEBuffer 2	90 μL
10 mM dATP solution	5 μL
Klenow Fragment (3′–5′ exo-, 5U/μL)	5 μL

Note: If processing multiple samples, we recommend preparing a master mix.

b. Place the Hi-C library on the magnet and wait 1 min to clear the supernatant.
c. Discard the supernatant and resuspend the beads in the A-tailing mix.
d. Incubate for 30 min at 37°C, without agitation.
e. Place the tube on a magnet and wait ~1 min to clear the supernatant.
f. Remove the supernatant and then resuspend the beads in 500 μL of 1× TWB.
g. Incubate in a thermomixer (2 min, 55°C, 1,000 rpm).
h. Place the tube back on the magnet and wait ~1 min for the supernatant to clear.
i. Repeat steps f–h.
j. Discard the supernatant, resuspend the beads in 50 μL of 1× Quick Ligase Buffer, and transfer to a new 1.5 mL microcentrifuge tube.

Optional: adapter ligation can also be performed using T4 DNA ligase for 2 h at 19°C–23°C or overnight (10 – 12 h) at 16°C.

If doing this, resuspend in 50 μL 1× T4 DNA ligase buffer instead and follow the manufacturer's instructions (https://www.thermofisher.com/order/catalog/product/EL0013?fr&en#/EL0013?fr&en) for setting up the reaction.

42. Adapter ligation

Timing: approximately 30 min

a. Prepare the Ligation mix by combining the following:

Reagent	Volume
1× Quick Ligase Buffer	48 μL
Quick Ligase	2 μL

Note: If processing multiple samples, we recommend preparing a master mix.

b. Place the Hi-C library on the magnet and wait ~1 min to clear the supernatant.
c. Discard the supernatant and resuspend the beads in the Ligation mix.
d. Add 2 μL of sequencing adapter and mix by pipetting.
e. Centrifuge a few seconds to collect sample at bottom of tube.
f. Incubate at 19°C–23°C for 10 min.
g. Place the tube on a magnet and wait ~1 min to clear the supernatant.
h. Remove the supernatant and then resuspend the beads in 500 μL of 1× TWB.
i. Incubate in a thermomixer (2 min, 55°C, 1,000 rpm).
j. Place the tube back on the magnet and wait 1 min for the supernatant to clear.
k. Repeat steps h–j.
l. Discard the supernatant, resuspend the beads in 50 μL of 10 mM Tris-HCl (pH 8) and transfer to a new 1.5 mL microcentrifuge tube.

Pause point: beads containing the Hi-C library can be stored at −20°C for at least 2 years

43. PCR amplification

Timing: approximately 1 h

a. Prepare the PCR mix by combining the following:

Reagent	Amount
Phusion 2× High Fidelity Master Mix	40 µL
2 µM Primer Mix (NEXTflex, PerkinElmer)	5 µL
Streptavidin beads containing Hi-C library	3 µL
ddH_2O	32 µL

Note: If processing multiple samples, we recommend preparing a master mix.

b. Amplify the library using the following conditions:

PCR cycling conditions			
Steps	Temperature	Time	Cycles
Initial denaturation	98°C	30 sec	1
Denaturation	98°C	10 sec	12
Annealing	61°C	15 sec	
Extension	72°C	15 sec	
Final extension	72°C	10 min	1
Hold	4°C	forever	

c. Place the PCR reaction on a magnet and wait ~1 min to clear the supernatant.
d. Transfer the supernatant to a new 1.5 mL microcentrifuge tube. Discard the Dynabeads.
e. Purify the PCR reaction by adding an equal volume (~80 µL) of AMPure XP beads.

Note: if multiple PCR reactions were performed (i.e., for low concentration samples), the reactions can be combined at this step and the volume of AMPure XP beads should be adjusted accordingly.

f. Mix sample 10× by gentle pipetting
g. Incubate for 5 min at 19°C–23°C to allow DNA fragments to bind to the magnetic beads.
h. Transfer tube to a magnetic rack to allow the beads to separate to the side of the tube (~ 1 min).
i. Carefully remove the supernatant and wash with 70% EtOH, keeping the tube on the magnet and ensuring that the beads are not disturbed.
j. Repeat step i.
k. Remove the EtOH, and airdry the pellet for 2 min to remove residual EtOH.

⚠ CRITICAL: do not allow the pellet to over-dry as this will reduce the elution efficiency.

l. Remove beads from the magnet and resuspend in 50 µL of 10 mM Tris-HCl pH 8.
m. Mix sample 10× by gentle pipetting.
n. Incubate for 5 min at 19°C–23°C and then place back onto the magnetic rack for ~1 min.
o. Carefully remove the supernatant and transfer it to a new 1.5 mL microcentrifuge tube.
p. Remove a 5 µL aliquot and transfer to a new tube. This is the final library (FL) control.
q. Run the FL sample on a 1% agarose gel to determine the fragment size of the sequencing library and to check for the presence of primer-dimers (troubleshooting 5, Figure 1).
r. Determine DNA concentration using the Qubit dsDNA HS kit on a Qubit Fluorometer.
s. Store the final sequencing library at −20°C.

Note: Samples can be stored at −20°C for more than 2 years with no impact on the quality of the Hi-C data.

t. Prepare libraries for paired-end sequencing using the 75 cycle High-Output Kit v2.5. according to Illumina's instructions (https://support.illumina.com/content/dam/illumina-support/

documents/documentation/system_documentation/nextseq-550dx/nextseq-550dx-high-output-reagent-kit-v2-5-75-cycles-package-insert-can-1000000081984-01.pdf).

Day 4: Data processing

Timing: approximately 1–2 days

44. All data were processed into contact maps accordingly to standard Hi-C analysis that include iterative alignment, filtering, and normalization (Cournac et al., 2012; Imakaev et al., 2012; Le et al., 2013; Lioy et al., 2018; Servant et al., 2015). Details relating to the datasets shown in Figure 2 are described in (Cockram et al., 2020).

EXPECTED OUTCOMES

The typical results obtained for the different steps of the bacterial/archaeal Hi-C protocol are shown in Figure 1. Following cell lysis and the preparation of the digestion reaction, we remove an aliquot and designate it as the non-digested (ND) control. This control monitors the amount of DNA degradation during these first steps. Next, we use the digested (D) control to check for efficient cutting of the genomic DNA by the restriction enzyme, this is shown by a DNA smear that is absent in the ND control. Following end-filling with a biotinylated nucleotide and blunt-end ligation, the efficiency of the ligation can be determined by checking that the DNA smear (L) has shifted to a higher molecular weight. Following reverse crosslinking, DNA purification and sonication, the average size of the DNA fragments is checked (S) before continuing with the preparation of the sequencing library. The size and quality of the final library (FL) are then checked on gel before pair-end sequencing. DNA concentration is determined by Qubit analysis. When starting with ~1×10^8 cells, we typically obtain a concentration of 5–15 ng/ul for the final Hi-C library. Figure 2 shows the typical Hi-C maps obtained from wild-type, exponentially growing cultures of the bacterium *E. coli* (Figure 2A) and the Euryarchaeota *H. volcanii* (Figure 2B (Cockram et al., 2020; see also Le et al., 2013 for an earlier example of bacteria contact map). This protocol is capable of generating 0.5 – 1 kb Hi-C matrices for species of both domains of life, offering gene-level resolution contact maps (Figures 2C and 2D).

LIMITATIONS

The Hi-C method relies on a population of cells and as a result the contact maps reflect an average of the population which may mask the presence of cell-cell heterogeneity in different conditions and mutant backgrounds. In addition, archaeal species are generally cultured in extreme conditions, such as high-salt or high-temperature. These species-specific growth conditions may require specific optimization. Finally, the Euryarchaeota and some bacterial species, such as *D. radiodurans*, are polyploid, so inter-chromosomal contacts between replicated chromosomes cannot be determined using the Hi-C technique.

TROUBLESHOOTING

Problem 1

Degradation of DNA in the non-digested (ND) control (step 22).

Potential solution

If some DNA degradation is observed, then the final Hi-C library will likely be fine, but if there is significant degradation, we recommend starting the protocol again with new buffers.

Problem 2

Incomplete digestion of genomic DNA (step 23).

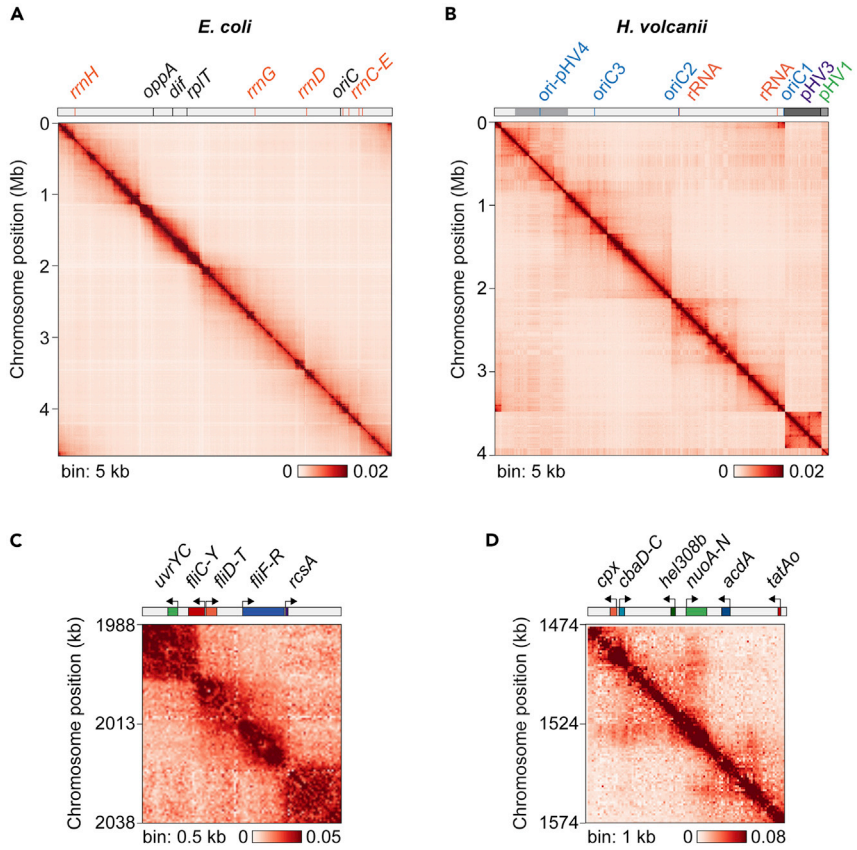


Figure 2. Representative examples of the quality of Hi-C matrices obtained using the optimized protocol for bacteria and archaea. b

Normalized Hi-C contact maps of asynchronously growing populations of WT *E. coli* and *H. volcanii* cells. The X and Y axes represent the coordinates of the chromosome and the color scale reflects the frequency of contacts between two regions of the genome (arbitrary units), from white (rare contacts) to dark red (frequent contacts). Features of interest are indicated along the top axis.

(A) Hi-C contact matrix of WT *E. coli* binned at 5 kb.

(B) Hi-C contact matrix of WT *H. volcanii* binned at 5 kb.

(C) Magnification of a 50 kb region surrounding the *E. coli* fliF-R operon (coordinates: 1988 – 2038 kb, bin = 0.5 kb).

(D) Magnification of a 100 kb region of the main *H. volcanii* chromosome (coordinates 1474 kb – 1574 kb; bin = 1 kb).

Potential solution

In our experience, some restriction enzymes work better than others and there isn't always a clear reason for this. We have tested HpaII and MluCI in a range of bacteria and archaea and both seem to work very well. However, if your organism requires a specific enzyme, we suggest performing steps 1–35 with a test sample first to ensure efficient digestion. A second reason for incomplete digestion could be that there is an issue with the digestion buffer, so we would suggest re-making this and trying the digestion again. Finally, it may be that you have too much DNA, so perhaps try diluting the lysate before digestion.

Problem 3

No shift in higher molecular weight following ligation (Figure 1; control **L**; steps 31–34).

Potential solution

This step may fail for one of two reasons. Either the end-fill step didn't work, or the blunt-end ligation was unsuccessful. Unfortunately, the only solution is to repeat the Hi-C library preparation from the beginning. We would recommend re-making all solutions for both the end-fill and ligation steps. If

the digestion (D) smear looks correct, then the buffers for this step are fine and do not need to be remade.

Problem 4

Poor sonication efficiency (step 38).

Potential solution

If the average fragment size is too high following sonication, we suggest re-checking the DNA concentration by Qubit analysis. The Covaris parameters detailed in this protocol are suitable for a maximum of 5 μg of DNA. If this is exceeded, the efficiency of sonication will be reduced and the fragment size will be too large. If you're using a different sonication method, we suggest checking the manufacturer's instructions and performing a time course experiment to determine the optimal conditions.

Problem 5

Low DNA yield in the final Hi-C library (step 43, r)

Potential solution

This is why we strongly encourage all users to take aliquots at the steps suggested and run the controls on an agarose gel (Figure 1). This will help them to identify which step went wrong. For example, if the gel shows a high concentration of DNA following sonication and size-selection, but the final library has a poor yield, then it suggests a problem occurred during the preparation of the sequencing library. We would first suggest repeating the final PCR amplification step, as it's the easiest and least-expensive problem to fix. If this doesn't solve the issue, then it suggests the DNA was lost in one of the previous steps, such as binding of the Hi-C library to the streptavidin beads and means the sequencing library needs to be prepared again. Alternatively, if there was already a low amount of material following sonication and size-selection, then it may mean that you need to scale up and perform two PCR reactions instead of one, we do not recommend increasing the number of PCR cycles. In our experience, as long as the final Hi-C library has a DNA concentration >1 ng/ul by Qubit analysis, then it should be ok to sequence.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Romain Koszul (romain.koszul@pasteur.fr).

Materials availability

This study did not generate new unique reagents

Data and code availability

The datasets generated during using protocol are detailed in (Cockram et al., 2020) and are available at PRJNA587586. Open-access versions of the programs and pipelines used for downstream data analysis (*E. coli* analysis pipeline and Chromosight; Matthey-Doret et al., 2020)) are available online via the lab Github account (https://github.com/koszullab/hicstuff).

ACKNOWLEDGMENTS

We thank Roxane Lestini, Aurore Gorlas, and Pierrick Moreau and the members of the RSG laboratory for insightful discussions regarding the development of this protocol. This research was supported by funding to Romain Koszul from the European Research Council under the Horizon 2020 program (ERC grant agreement: 771813).

AUTHOR CONTRIBUTIONS

C.C. and A.T. adapted the Hi-C protocol. C.C. and R.K. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Allers, T., Ngo, H.-P., Mevarech, M., and Lloyd, R.G. (2004). Development of additional selectable markers for the halophilic archaeon haloferax volcanii based on the leuB and trpA genes. Appl. Environ. Microbiol. *70*, 943–953.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., et al. (1997). The complete genome sequence of Escherichia coli K-12. Science *277*, 1453–1462.

Cockram, C., Thierry, A., Gorlas, A., Lestini, R., and Koszul, R. (2020). Euryarchaeal genomes are folded into SMC-dependent loops and domains, but lack transcription-mediated compartmentalization. Mol. Cell. *81*, 459–472.e10.

Cournac, A., Marie-Nelly, H., Marbouty, M., Koszul, R., and Mozziconacci, J. (2012). Normalization of a chromosomal contact map. BMC Genomics *13*, 436.

Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat. Methods *9*, 999–1003.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Meth *9*, 357–359.

Le, T.B.K., Imakaev, M.V., Mirny, L.A., and Laub, M.T. (2013). High-resolution mapping of the spatial organization of a bacterial chromosome. Science *342*, 731–734.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.; 1000 Genome Project Data Processing Subgroup. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Lioy, V.S., Cournac, A., Marbouty, M., Duigou, S., Mozziconacci, J., Espéli, O., Boccard, F., and Koszul, R. (2018). Multiscale structuring of the E. coli chromosome by nucleoid-associated and condensin proteins. Cell *172*, 771–783.e18.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*. https://doi.org/10.1186/s13059-014-0550-8.

Marbouty, M., and Koszul, R. (2017). Generation and analysis of chromosomal contact maps of bacteria. Methods Mol. Biol. *1624*, 75–84.

Marbouty, M., Cournac, A., Flot, J.-F., Marie-Nelly, H., Mozziconacci, J., and Koszul, R. (2014). Metagenomic chromosome conformation capture (meta3C) unveils the diversity of chromosome organization in microorganisms. ELife *3*, e03318.

Marbouty, M., Le Gall, A., Cattoni, D.I., Cournac, A., Koh, A., Fiche, J.-B., Mozziconacci, J., Murray, H., Koszul, R., and Nollmann, M. (2015). Condensin- and replication-mediated bacterial chromosome folding and origin condensation revealed by Hi-C and super-resolution imaging. Mol. Cell *59*, 588–602.

Marbouty, M., Baudry, L., Cournac, A., and Koszul, R. (2017). Scaffolding bacterial genomes and probing host-virus interactions in gut microbiome by proximity ligation (chromosome capture) assay. Sci. Adv. *3*, e1602105.

Matthey-Doret, C., Baudry, L., Breuer, A., Montagne, R., Guiglielmoni, N., Scolari, V., Jean, E., Campeas, A., Chanut, P.H., Oriol, E., et al. (2020). Computer vision for pattern detection in chromosome contact maps. Nat. Commun. *11*, 5795.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic acids research *44*, W160–W165.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

RStudio Team (2020): RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. *16*, 259.

Takemata, N., Samson, R.Y., and Bell, S.D. (2019). Physical and functional compartmentalization of archaeal chromosomes. Cell *179*, 165–179.e18.

Yaffe, E., and Tanay, A. (2011). Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. Nat. Genet. *43*, 1059–1065.