cDNA library construction for next-generation sequencing to determine the transcriptional landscape of Legionella pneumophila

Tobias Sahr, Carmen Buchrieser

cDNA library construction for Next Generation Sequencing (NGS) to determine the transcriptional landscape of *Legionella pneumophila*

Tobias Sahr and Carmen Buchrieser*

Institut Pasteur, Biologie des Bactéries Intracellulaires, 75724 Paris, France and [2]CNRS UMR 3525, 75724 Paris, France

Running title: RNAseq of *Legionella*

Key words: *Legionella pneumophila*, next generation sequencing, transcriptional start site mapping, small ncRNA

Summary

The adaptation of *Legionella pneumophila* to the different conditions it encounters in the environmental and in the host is governed by a complex regulatory system. Current knowledge of these regulatory networks and the transcriptome responses of *L. pneumophila* is mainly based on microarray analysis and limited to transcriptional products of annotated protein-coding genes. The application of the Next Generation Sequencing (NGS) technology allows now genome wide strand-specific sequencing and accurate determination of all expressed regions of the genome to reveal the complete transcriptional network and the dynamic interplay of specific regulators on a genome wide level. NGS based techniques promote deeper understanding of the global transcriptional organization of *L. pneumophila* by identifying transcription start sites (TSS), alternative TSS and operon organization, noncoding RNAs (ncRNA), antisense RNAs and 5'-/3'-untranslated regions. In this chapter we describe the construction of cDNA libraries for (i) RNA deep sequencing (RNA-seq) and (ii) transcription start site (TSS) mapping using the Illumina technology.

1. Introduction

In recent years, the field of microbial genomics has changed considerable due to the development of new generation sequencing techniques (*1*). Thus, deep RNA sequencing is now revolutionizing our understanding of the complexity, plasticity and regulation of microbial transcriptomes (*2*). Recent studies using RNA deep sequencing indicate that the function of untranslated transcriptional regions as posttranscriptional regulator was underestimated in bacteria until now (*3-6*). In particular the unexpected high amount of ncRNAs found so far strongly suggests that they play a more dominant and widespread role in regulation of gene expression including a) transcription interference/termination b) translational interference c) effects on the stability of target RNA and d) interaction with RNA-binding proteins (for review see (*7*)). To date the *L. pneumophila* transcriptome was analyzed mainly using microarrays (*8-12*). Recently the first study using NGS techniques to analyze the intracellular transcriptome was applied to *L. pneumophila* allowing to discover 70 novel small RNAs (*13*). However, the application of NGS for an in-depth transcriptional analysis of *L. pneumophila* may revolutionize our understanding of adaptive and developmental processes during infection, as it will add new levels of control to our regulatory network. In the following, we present several strategies to construct strand-specific cDNA libraries for Illumina sequencing. A protocol for strand-specific RNA deep

sequencing (RNAseq) and two different methods to enrich and define transcriptional start sites (TSS mapping).

Parallel deep sequencing of the whole transcriptome (RNAseq) reveals detailed prediction of gene expression by determination of length and abundance of transcripts including 5'- and 3'-untranslated region or antisense transcription at different conditions. Anyhow, RNAseq has its limitation as it is not possible to distinguish primary transcripts from processed 5' ends, so, it cannot be used e.g. for identification of alternative transcription starts or operon organization. To achieve a more precise view on the level of primary transcripts it is necessary to construct specific TSS mapping libraries in which the 5'ends of the RNA are highly enriched in comparison to an untreated cDNA control library. We can distinguish between two different approaches to analyze the TSS region:

1) Terminator Exonuclease (TEX) method: differential cDNA library pairs treated (+) or not treated (-) with TEX. Primary transcripts in bacteria like mRNA and ncRNA (but not rRNA) are known to be protected by a triphosphate cap at the 5'end. TEX digests specifically RNA having a 5'-monophosphate, but is not able to degrade RNA with a 5'-PPP protection. Therefore, TEX eliminates processed transcripts like partially degraded or sheared RNA leading to an enrichment of primary bacterial transcripts with intact 5'-triphosphate terminus (Fig 1).
2) Tobacco Acid Pyrophosphatase (TAP) method: differential cDNA library pairs treated (+) or untreated (-) with TAP. TAP hydrolyzes pyrophosphate of 5'-PPP RNA resulting in 5'monophosporylated end. This step is essential during the library construction as the 5' RNA adapter will ligated only to a 5'-monophosphorylated RNA but not to RNA with 5'-PPP terminus. For that reason, the untreated TAP (-) control library include all transcripts with 5'-P terminus while the TAP (+) library additionally comprise the start sites of the transcript formally protected by the 5'-triphosphate cap (Fig 2).

2. Materials

2.1. Strains and growth media

1. *Legionella pneumophila* e.g. strains Paris, JR32 or Philadelphia-1

2. AYE medium. For 1 liter dissolve 12 g yeast extract and 10 g ACES, adjust pH to 6.9 with 1 M KOH. Add 10 mL of cysteine 40 g/L and 10 mL of iron pyrophosphate 30 g/L. Fill volume to 1 L with distilled water and filter sterilize.

3. CYE plates. For 1 liter dissolve 10 g yeast extract and 10 g ACES, adjust pH to 6.9 with 1M KOH, add 15 g of agar, 2 g of activated charcoal and autoclave. Add 10mL of filter sterilized cysteine 40 g/L and 10 mL of filter sterilized ferric nitrate 25 g/L. Poor plates.

2.2. RNA extraction

1. Resuspension buffer: ½ volume of Glucose 20% + ½ volume of Tris 25 mM pH 7.6

2. EDTA 10 mM

3. EDTA 0.5 M

4. Glass beads, Sigma (200-300 microns Sigma G1277)

5. Water treated with DEPC

6. Phenol acid pH 4.5 (Interchim)

7. Total RNA extraction, TRIzol reagent (Invitrogen,

8. Chloroforme/alcohol isoamyl 24/1 (v/v)

9. Isopropanol (RT temperature)

10. Ethanol 70% (ice cold)

11. Tris-HCl 10mM pH 7.6 EDTA 1mM

12. FastPrep Instrument to lyse bacteria

13. 1.5 mL microcentrifuge tubes

14. Shaking platform at 37°C

15. Disposable 50 mL polypropylene tubes

2.3. cDNA library construction for Illumina sequencing

1. 10µg RNA

2. DNase I recombinant, RNase free

3. Microbe Express, Bacterial mRNA Enrichment Kit from

4. Terminator™ 5'-Phosphate-Dependent Exonuclease, TEX

5. Phenol solution, pH4.3 (Sigma, P4682-100ML); Chloroform:Isoamyl alcohol 24:1

6. RNase inhibitor (e.g. RNaseOUT™ Ribonuclease Inhibitor)

7. 3M Sodium Acetate, pH5.2

8. Ethanol p.a.; Ethanol 75%

9. Glycogen (20mg/ml)

10. Tobacco acid pyrophosphatase 10U/µl, TAP (Epicentre, T19500)

11. 5'-RNA adapter (25µM) GUUCAGAGUUCUACAGUCCGACGAU

12. T4 RNA Ligase 5U/µl

13. Superscript II Reverse Transcriptase Kit

14. RT Random primer (100µM) CAAGCAGAAGACGGCATACGANNNNNN

15. RT Specific primer (100µM) CAAGCAGAAGACGGCATACGA

16. RNA fragmentation kit (Ambion)

17. Alkaline Phosphatase

18. 3'-RNA adapter (100µM) P-UCGUAUGCCGUCUUCYGCUUGUidT

19. T4 Polynucleotide Kinase

20. 10mM dNTP mix

21. Certified Low Range Ultra Agarose (Biorad, 161-3106)

22. Agarose gel electrophoresis equipment

23. 0.5M EDTA pH8.0

24. Distilled water DNase/RNase free

25. Thermocycler

26. Phusion HF DNA Polymerase

27. Primer 1 (25µM) CAAGCAGAAGACGGCATACGA

28. Primer 2 (25µM) AATGATACGGCGACCACCGACAGGTTCAGA GTTCTACAGTCCGA

29. NucleoSpin® Gel and PCR Clean-up

30. Bionalyzer 2100 (Agilent)

31. RNA Nano Chips (Agilent, 5067-1511)

32. DNA 1000 Chips (Agilent, 5067-1504)

3. Methods

3.1. RNA isolation

1. From a glycerol stock maintained at -80°C, streak the strain on a CYE plate and incubate it at 37°C for 72 h until you obtain colonies in stationary phase.

2. Pick colonies from the plate and grow a pre-culture in 50ml polypropylene tubes overnight, 37°C, shaking

3. Dilute the pre-culture. Grow cultures to a chosen OD and take a sample (10 ml).

4. Centrifuge in a 15ml falcon tube, 5000xg for 5 min at 4°C in a pre-cooled centrifuge.

5. Withdraw the supernatant and flash freeze the pellet on dry ice +EtOH or proceed with the next step immediately.

6. Prepare a sarstedt tube, add 500 µl phenol acid and 0.4 g glass beads.

7. Resuspend the bacterial pellet in 400 µl resuspension buffer and 60 µl EDTA 0.5 M.

8. Transfer suspension to the tubes containing phenol acid and glass beads.

9. Lyse the cells with the FastPrep apparatus with the following settings:
Speed: 6.0; Time: 30s

10. Let stand for 1 min at 4°C and repeat step 4 once more with the same settings.

11. Centrifuge for 5 minutes at 13000 rpm at 4°C. Transfer the top liquid phase to a sterile eppendorf tube.

12. Add 1 ml room-tempered Trizol. Mix very gently with a pipette till it is 'foamy' and let stand for 5 min on bench.

13. Add 100 µl chloroforme/IAA. Mix vigorously by shaking. Let stand for 1 minute on bench.

15. Centrifuge 5 min at 13000 rpm at 4°C. Transfer the aqueous phase to a new eppendorf tube.

16. Add 200 µl chloroforme/IAA. Mix vigorously and let stand for 1 minute at room temperature.

17. Centrifuge for 5 minutes, 13000 rpm at 4°C. Transfer the aqueous phase to a sterile eppendorf tube.

18. Add 500 µl isopropanol and agitate by inversing the tube. Let stand for 30 minutes on ice.

19. Centrifuge 15 minutes, 13000 rpm at 4°C.

20. Rince the pellet with 1 ml ice cold ethanol 70%.

21. Centrifuge 5 minutes, 13000 rpm at 4°C.

22. Withdraw the supernatant and dry the pellet (SpeedVac or airdry on bench).

23. Resuspend the pellet in 50 µl H_2O.

24. Incubate for 15 minutes at 37°C, measure the concentration aliquot and freeze to –80°C or use directly for library construction

3.2. Transcription start sites (TSS) mapping library construction

1. RNA extraction using the TRIzol reagent (see also **Note 1**)
2. Depletion of rRNA using Microbe Express (**Note 2**)
3. TEX treatment: divide the RNA in two similar aliquots and incubate one with or without TEX (TEX-/+ library, for construction of TAP-/+ library see **Note 3**):
   Xµl depleted RNA (recovered from MicrobeExpress)
   Xµl RNase-free water
   0.5µl RNase Inhibitor
   2µl TEX buffer A
   1U TEX
   (final volume of the reaction 20µl)

4. Incubate for 60min at 30°C

5. Terminate reaction by adding 1µl 100mM EDTA + 180ulH2O

6. Purify RNA by organic extraction (PhenolChloroform/Isoamylalcohol, see **Note 4**). After NaAc precipitation (Note 5), resuspend the dry pellet in 44µl distilled water

7. TAP treatment: incubate both samples (with and without TEX) independently with TAP:

8. Denature the RNA from step 3 for 10min at 65°C

9. Incubate 1min on ice

10. Add 5µl TAP buffer and 1µl TAP (10U) and incubate for 1h at 37°C

11. Extract with Phenol/Chloroform/IAA (**see Note 4**)

12. Precipitate with NaAc (**see Note 5**)

13. Resuspend dry pellet in 5.3µl distilled H_2O

14. 5'-RNA adapter ligation:

5.3µl RNA

1.2µl 5'-RNA adapter (25µM)

Incubate mix for 10min at 65°C, put on ice for 1min and add

1µl T4 RNA ligase buffer

1µl ATP solution

0.5µl RNase inhibitor

1µl RNA ligase

And incubate for 6h at 20°C, following 4°C over night

15. Reverse Transcription using RT Random primer

10µl RNA +1.5µl random primer (100µM)

Denature at 65°C for 10min, put on ice 1min and add

6µl first strand buffer

1.5µl DTT (100mM)

1.5µl dNTP

1µl RNase inhibitor

7µl RNase-free H2O

1.5µl Superscript II RT

16. Incubate 10min at 25°C, following 1h at 42°C

17. Size fractionation on a low range ultra 2% agarose gel (**Note 6**)

Cut the zones between 100-170nt and 170-250nt and purify independently with columns (e.g. Nucleospin)

Combine the samples of the two "zones", precipitate and resuspend pellet in 30µl distilled water

18. PCR amplification (**Note 7**)

30µl template cDNA

10µl Phusion HF buffer

8µl H_2O

0.5µl Primer 1 (25µM)

0.5µl Primer 2 (25µM)
0.5µl dNTP (25mM)
0.5µl Phusion DNA Polymerase

98°C 1min

98°C 10sec
60°C 30sec
72°C 30sec
15 cycles

72°C 10min
4°C

19. Purify PCR reaction with columns (e.g. Nucleospin)

20. Precipitate and resuspend in 10µl

21. Check quality and quantity with the Bioanalyzer (DNA 1000)

22. Ready for Sequencing with Illumina

3.3. RNAseq library construction

1. RNA extraction (Trizol method, see also Note 1)
2. Depletion of rRNA using Microbe Express
   After NaAc precipitation, resuspend the dry pellet in 10µl distilled H_2O
3. Fragmentation: the RNA was metal-catalyzed heat fragmented to sizes of around 100-200nt using the RNA fragmentation kit

   1.1µl 10x fragmentation reagent (buffered zinc solution
   10µl RNA
   Incubation at 70°C for 5min
   Terminate reaction by putting on ice and add 1.1µl stop solution

After NaAc precipitation (**Note 5**), resuspend the dry pellet in 26µl RNase-free distilled H_2O

4. RNA dephosphorylation:

a) TAP treatment

Denature the RNA from step 3 for 10min at 65°C

After 1min on ice, add 3µl TAP buffer and 1µl TAP (10U) and incubate for 1h at 37°C

b) Alkaline Phosphatase treatment

Add 14µl distilled H2O, 5µl 10x Dephosphorylation buffer and 1µl Alkaline Phosphatase (10U) and incubate for 30min at 37°C

Phenol/Chloroform/IAA extraction (**Note 4**), NaAc precipitation; resuspend the dry pellet in 5.9 µl distilled water

5. 3'-RNA adapter ligation

5.9µl RNA

0.6µl 3'adapter (100µM)

Incubate mix for 10min at 65°C, put on ice for 1min and add

1µl T4 RNA ligase buffer

1µl ATP solution

0.5µl RNase inhibitor

1µl RNA ligase

Incubate for 6h at 20°C, following 4°C over night, purify with Phenol/Chloroform/IAA extraction and resuspend RNA after NaAc precipitation in 10µl distilled water

6. Re-phosphorylation of 5' end

10µl RNA

2µl PK buffer

1µl ATP solution

0.5µl RNase inhibitor

1µl Polynucleotide Kinase
5.5µl RNase free H2O

7. Size fractionation on a low range ultra 2% agarose gel
Cut the zones between 100-170nt and 170-250nt and purify with columns (e.g. NucleoSpin). Precipitate with NaAc and resuspend in 5.3µl distilled H_2O

8. 5'-RNA adapter ligation:

5.3µl RNA
1.2µl 5'-RNAadapter (25µM)
Incubate 10min at 65°C, on ice for 1min and add
1µl T4 RNA ligase buffer (Epicentre)
1µl ATP solution
0.5µl RNase inhibitor
1µl RNA ligase
And incubate for 6h at 20°C, following 4°C over night

9. Reverse Transcription using RT Specific primer
10µl RNA +0.5µl RT Specific primer (100µM)
Denature at 65°C for 10min, on ice 1min and add
6µl first strand buffer
1.5µl DTT (100mM)
1.5µl dNTP
1µl RNase inhibitor
8µl RNase-free H2O
1.5µl Superscript II RT

Incubate 1h at 42°C

10. Size fractionation on a low range ultra 2% agarose gel (**Note 6**)
Cut the zones between 100-170nt and 170-250nt and purify independently with columns (e.g. Nucleospin)

Combine the samples of the two “zones”, precipitate and resuspend pellet in 30µl distilled water

11. PCR amplification (see also **Note 7**)

30µl template cDNA
10µl Phusion HF buffer
8µl H_2O
0.5µl Primer 1 (25µM)
0.5µl Primer 2 (25µM)
0.5µl dNTP (25mM)
0.5µl Phusion DNA Polymerase

98°C 1min

98°C 10sec
60°C 30sec
72°C 30sec
15 cycles

72°C 10min
4°C

Purify PCR reaction with columns (e.g. Nucleospin), precipitate and resuspend in 10µl

12. Check quality and quantity with the Bioanalyzer (DNA 1000)
13. Ready for Sequencing with Illumina

4. Notes

1. Extracted RNA must be treated with DNase I, purity and concentration can be determined by measuring the absorbance at 260nm an 280nm. Nevertheless, to

guarantee best quality a Bioanalyzer analysis should be performed with total and rRNA-depleted RNA.

2. Microbe Express is performed according to the manufacture's instruction with 2x10µg of total RNA for each condition as starting material.
3. Alternatively to the comparison of libraries treated with and without TEX, a TAP-/+ library can be constructed. For this purpose no TEX treatment is necessary! Instead, divide the depleted RNA after step 2 (MicrobeExpress treatment) in two similar aliquots and incubate one fraction with, the other without TAP as described in step 7-10 and continue with step 11.
4. Phenol/Chloroform/IAA extraction: Add 50% of Phenol to the sample and mix vigorously by vortexing. Add the same amount of Chloroform/Isoamyl alcohol and mix well. Centrifuge at maximal speed for 5 min, transfer the supernatant to a new tube and add 1 vol of Chloroform/Isoamyl alcohol, mix vigorously, centrifuge and transfer supernatant into a new tube
5. NaAc precipitation: Add 10% 3M sodium acetate (pH5.2), 2% glycogen and 2.5vol of ethanol p.a. to the sample and store it min 2h (better over night) at-20°C. Centrifuge at 4°C for 15min at max speed and discard supernatant. Wash RNA pellet with 500µl ice-cold 75% ethanol centrifuge again for 5min, discard supernatant and let the pellet dry for 10min at RT. Resuspend the dry pellet RNase-free distilled H_2O
6. Final library with a median insert size of around 200 bp is ideal for Illumina NGS. Anyhow, if you choose longer fragments you need to reduce your loading concentration otherwise the clusters begin to overlap due to the length. Also a too wide range of fragment size will have a negative effect on the quality and distribution of clustering.
7. Number of cycles during PCR amplification can differ between experiments depending on the amount of cDNA obtained in the previous step. Typically are 12 to maximal 17 cycles. More cycles are not recommended as additional PCR steps may introduce a significant amplification bias in cDNA representation. DNA amounts due to Bioanalyzer DNA1000 analysis are typically in a range of 10-50 nM. If much less material is obtained, adapt number of cycles in the PCR or increase the amount of starting material. For the hybridization on the Cluster Station concentrations of 1–10 pM are recommended.

Acknowledgements

This work received support from the Institut Pasteur, the Centre national de la recherché scientifique (CNRS) and the Institut Carnot-Pasteur MI and from the ANR-10-PATH-004 project, in the frame of ERA-Net PathoGenoMics

References

1. Metzker, M. L. (2010) Sequencing technologies - the next generation, *Nat Rev Genet*, 1.
2. Sorek, R., and Cossart, P. (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity, *Nat Rev Genet 11*, 9-16.
3. Dornenburg, J. E., Devita, A. M., Palumbo, M. J., and Wade, J. T. (2010) Widespread antisense transcription in *Escherichia coli*, *MBio 1*, e00024-00010.
4. Sharma, C. M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R., Stadler, P. F., and Vogel, J. (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*., *Nature 464*, 250-255.
5. Thomason, M. K., and Storz, G. (2010) Bacterial antisense RNAs: how many are there, and what are they doing?, *Annu Rev Genet 44*, 167-168.
6. Vivancos, A. P., Güell, M., Dohm, J. C., Serrano, L., and Himmelbauer, H. (2010) Strand-specific deep sequencing of the transcriptome., *Genome Res 20*, 989-999.
7. Storz, G., Vogel, J., and Wassarman, K. M. (2011) Regulation by small RNAs in bacteria: expanding frontiers., *Mol Cell. 43*, 880-891.
8. Bruggemann, H., Hagman, A., Jules, M., Sismeiro, O., Dillies, M. A., Gouyette, C., Kunst, F., Steinert, M., Heuner, K., Coppee, J. Y., and Buchrieser, C. (2006) Virulence strategies for infecting phagocytes deduced from the in vivo transcriptional program of Legionella pneumophila, *Cell Microbiol 8*, 1228-1240.
9. Dalebroux, Z. D., Yagi, B. F., Sahr, T., Buchrieser, C., and Swanson, M. S. (2010) Distinct roles of ppGpp and DksA in *Legionella pneumophila* differentiation, *Mol Microbiol 76*, 200-219.
10. Faucher, S. P., Mueller, C. A., and Shuman, H. A. (2011) *Legionella pneumophila* Transcriptome during Intracellular Multiplication in Human Macrophages, *Front Microbiol 2*, 60.
11. Sahr, T., Bruggemann, H., Jules, M., Lomma, M., Albert-Weissenberger, C., Cazalet, C., and Buchrieser, C. (2009) Two small ncRNAs jointly govern virulence and transmission in *Legionella pneumophila*, *Mol Microbiol 72*, 741-762.
12. Hovel-Miner, G., Pampou, S., Faucher, S. P., Clarke, M., Morozova, I., Morozov, P., Russo, J. J., Shuman, H. A., and Kalachikov, S. (2009) SigmaS controls multiple pathways associated with intracellular multiplication of *Legionella pneumophila*, *J Bacteriol 191*, 2461-2473.

13. Weissenmayer, B. A., Prendergast, J. G., Lohan, A. J., and Loftus, B. J. (2011) Sequencing illustrates the transcriptional response of *Legionella pneumophila* during infection and identifies seventy novel small non-coding RNAs, *PLoS One 6*, e17570.

Fig 1 Example for TEX- (black) and TEX + (green) library comparison: Artemis software image of the *lpp0001-0003* region. Peaks are representing the relative coverage of strand-specific reads obtained from sequencing cDNA libraries with Illumina HiSeq and mapped to the *L. pneumophila* Paris genome.

Fig 2 Comparison of TAP- (black) and TAP+ (green) library: Artemis software image of lpp0001-0003 region. Peaks are representing the relative coverage of strand-specific reads obtained from sequencing cDNA libraries with Illumina HiSeq and mapped to the *L. pneumophila* Paris genome.

Figure 1.

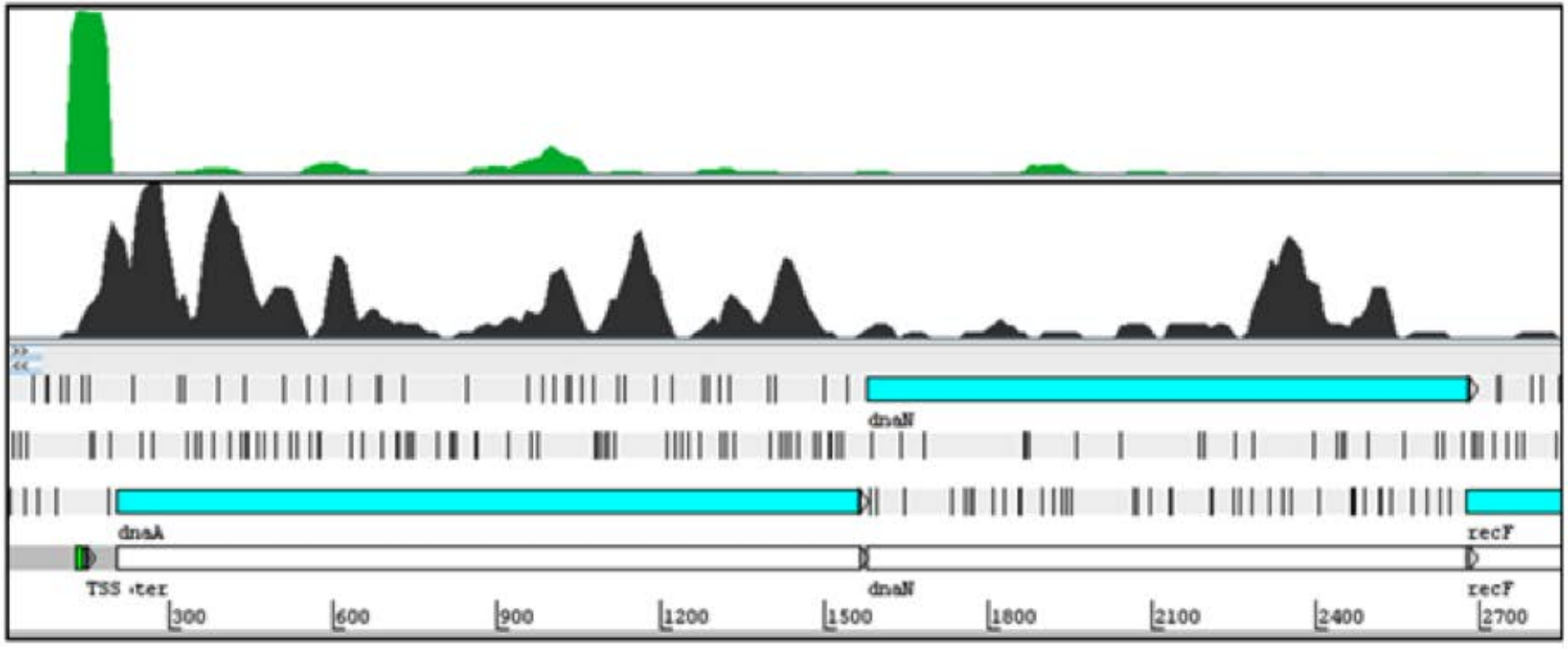
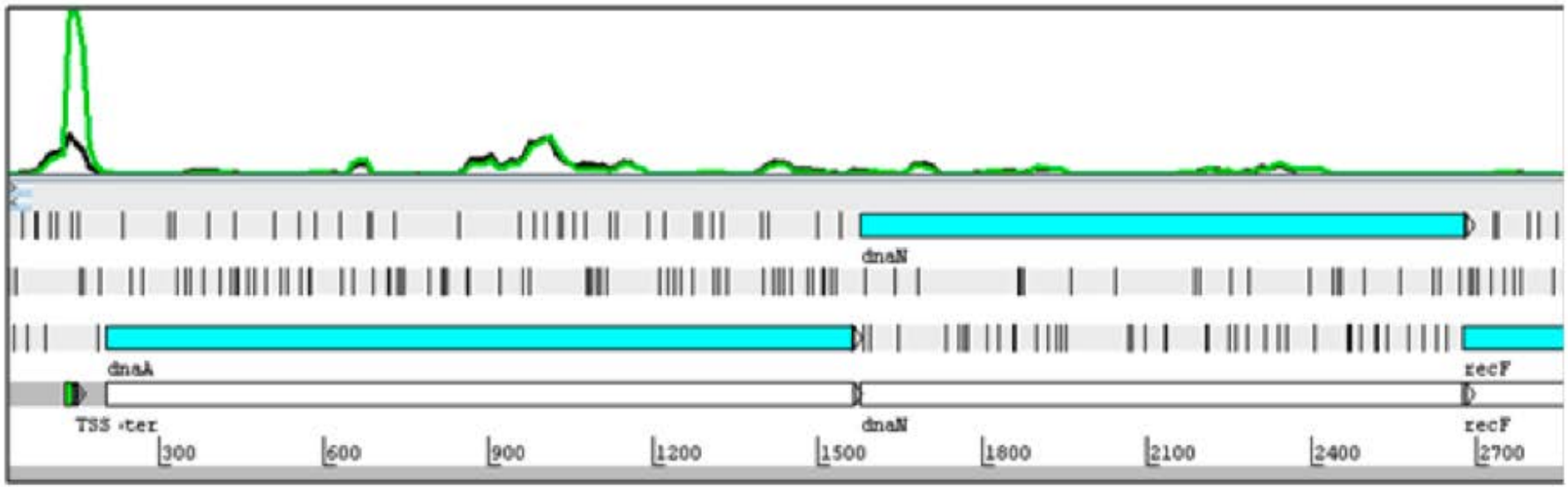

dnaN
dnaA
recF
TSS
ter
dnaN
recF
300
600
900
1200
1500
1800
2100
2400
2700

Figure 2.


dnaN
dnaA
recF
TSS
ter
dnaN
recF
300
600
900
1200
1500
1800
2100
2400
2700