High-Throughput CRISPR Typing of Mycobacterium tuberculosis Complex and Salmonella enterica Serotype Typhimurium

Christophe Sola, Edgar Abadia, Simon Le Hello, François-Xavier Weill

High-throughput CRISPR typing of *Mycobacterium tuberculosis* complex and *Salmonella enterica* serotype Typhimurium

Running title : "CRISPR Typing for public health"

Christophe Sola[1], Edgar Abadia[2], Simon Le Hello[3], François-Xavier Weill[3]

Summary

Spoligotyping was developed almost 16 years ago and still remains a popular first-lane genotyping technique to identify and subtype *Mycobacterium tuberculosis* complex (MTC) clinical isolates at a phylogeographic level. For other pathogens, such as *Salmonella enterica,* recent studies suggests that specifically designed spoligotyping techniques could be interesting for public health purposes. Spoligotyping, was in its original format a reverse line-blot hybridization method using capture probes designed on "spacers" and attached to a membrane's surface and a PCR product obtained from Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs). Cowan *et al*., (2004) and Fabre *et al*., (2012) were the first to propose a high-throughput Spoligotyping method based on microbeads for MTC and *S. enterica* serotype Typhimurium, respectively [1, 2]. The main advantages of the high throughput Spoligotyping techniques we describe here are their low cost, their robustness and the existence (at least for MTC) of very large databases that allow comparisons between spoligotypes from anywhere.

Key words: spoligotyping, CRISPR locus, Microbeads, High-throughput, Molecular

1 Institut de Génétique et Microbiologie, CNRS-University of Paris-Sud 11, Orsay, France
2 Centre Muraz, Bobo-Dioulasso, Burkina Faso
3 Institut Pasteur, Unité des Bactéries Pathogènes Entériques, Paris

epidemiology.

1. Introduction

The discovery of a region within a *Mycobacterium bovis* BCG strain characterized by the presence of short repeats, each interspaced by unique sequences that were highly polymorphic allowed the invention of the Spoligotyping technique [3] [4]. The name of this technique stands for *sp*acer *oligo*nucleotide *typing*, an acronym that was created by a research team in the National Institut of Health and Environment in Bilthoven, The Netherlands, who patented and standardized the technique for *Mycobacterium tuberculosis* complex (MTC). In 2002, the unique and peculiar genetic structure of this region was designated as CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) [5]. CRISPR loci were found to be present in nearly all archaea and in almost 50% of bacteria [6, 7]. These structures represent at least in some species such as *Streptococcus thermophilus*, an adaptative immune system that allows the bacteria to defend against invader DNA or RNA [8]. The discovery of other physiological roles of this complex RNA-based interference mechanism of regulation is expanding [9]. The extreme molecular diversity of these CRISPR loci make them ideal to target bacterial strain diversity and perform subtyping, indirectly allowing clues on the natural history and evolutionary genetics of the underlying disease in the case of bacterial pathogens. Subtyping methods based on analyses of the spacers of CRISPR loci have since been developed for other bacteria of medical interest, such as *Yersinia pestis* and *Y. pseudotuberculosis* [10], *Corynebacterium diphtheriae* [11], *Salmonella enterica* [2, 12-14], *Legionella pneumophila* 1 [15], and *Streptococcus agalactiae* [16]. Hermans *et al.* revealed the presence of the DR region in MTC strains through sequencing [17]. The DR region consists of direct variable repeats (DVR), each made up of a constant and a variable part [3]; in MTC, the constant is represented by identical repeated sequences of 36 bp length (DR, direct repeats) interspaced by unique variable sequences (spacers) of 35-41 bp length that generate the polymorphism.

The absence of some spacers may be characteristic of a given sub-species or sublineages (e.g. the absence of spacer 3, 9, 16 and 39-43 in *M. bovis* BCG). Another example is the rare *M. canetti* subspecies which harbors specific spacers (69-104) or the signature of absence of the spacers 1-33 and presence of spacers 34-43 for the « Beijing » lineage. In 2000 van Embden *et al* provided more knowledge on the genetic diversity of this locus on MTC strains and also some hypothesis about how the region may evolve [18]. Filliol *et al* showed a good correlation between spoligotypes signatures and geographic regions which in turn could be the result of MTC strains genomic changes and adaptation to their host [19]. It seems that the region evolves mainly losing spacers so the way particular spacers are being lost may represent phylogenetic signatures during their evolution. Gagneux *et al.* proposed that MTC lineages are adapted to particular human populations [20]. Indeed, strains from different lineages of MTC are indeed strongly associated with specific geographical regions and with patient country of origin [21, 22].

The locus's schematic view and the technique's principle are shown in Figure 1 and a raw experiment and deduced pattern are shown in Figure 2. Briefly, the power of the technique relies on the amplification of all the spacers which are present in the CRISPR region at once using one pair of primers that are complementary to each DR sequence. One of the primers needs to be biotin-labelled; thus biotin-labelled single-strand DNA of heterogeneous size will be produced. PCR products will be hybridized over a membrane to which a set of predefined complementary oligonucleotides (capture probes) were previously chemically attached. The membrane is supported by a matricial device (miniblotter) in which the hybridization procedure takes place. After the hybridization, washing steps will be done and will allow to get rid of non-specific hybridization events. The biotinylated-hybridized PCR fragments will be revealed after exposure to a streptavidin-peroxydase conjugate through a classical electrochemiluminescence autoradiogram. The result is a matrix of hybridized (black spots) or

non-hybridized spots (no spots) depending on the presence or absence of the corresponding spacers in the original DNA sample (Figure 2). Each DNA produces a unique pattern that will provide a first raw identity of a patient-specific clinical isolate. Some patterns are highly patient and strain-specific whereas others are highly common and poorly significant, requiring further typing. Patterns have to be compared to databases to reveal their informativeness [23]. The hybridizing/non hybridizing spacers patterns transcription is done from the membrane to a spreadsheet of OpenOffice® or Excel® software. The order is strictly conserved from 1 to 43. For better diplay of results, an ultrametric font such as « Monotype Sort » or « Zapf Dingbats » size 10, using characters « n » for each spacer present and « o » for each spacer absent should be used. With such a display, a 43-character black/white pattern is easily recognized by human beings and even more easily by computers using machine-learning algorithms. Other spoligotype pattern display methods were developed according to the need of handling a code with fewer characters, like the octal code (15 digit) or the Hexadecimal code (12 characters) [24]. Just a single script (Excel's macro) can translate one code to another one (directly available from the authors upon request).

In *Salmonella*, there are two CRISPR loci, CRISPR1 and CRISPR2, separated by less than 20 kb. The CRISPR1 locus is located downstream from the *iap* gene, whereas CRISPR2 is located upstream from the *ygcF* gene. The ordered CRISPR-associated (*cas*) genes belonging to the Ecoli subtype are located between the CRISPR loci. The DRs of both CRISPR loci were conserved. They were 29 bp long and had the consensus sequence 5′-CGGTTTATCCCCGCTGGCGCGGGGAACAC-3′. The CRISPR analysis by PCR and sequencing of 783 strains belonging to 130 serotypes revealed the presence of 3800 spacers (mean size 32 bp) [2]. The spacer contact was found correlated with both serotype and multilocus sequence type. Furthermore, spacer microevolution (duplication, triplication, loss or gain of spacers, presence of SNP variant spacers or VNTR variant spacers) discriminated

between subtypes within prevalent serotypes such as Typhimurium (STM), the most prevalent serotype worldwide. In eight genomes and 150 strains of serotype Typhimurium and its monophasic 1,4,[5],12:i:- variant, it was found 57 CRISPR1, 62 CRISPR2 alleles and 83 CRISPR1-CRISPR2 combined alleles. Forty unique spacers (including four with variants, such as SNP or VNTR variants) were identified in CRISPR1. Thirty-nine unique spacers (including two with a SNP variant) were identified in CRISPR2. Particular well-characterized populations, such as multidrug-resistant DT104 isolates, African MDR ST313 isolates, and DT2 isolates from pigeons, each had typical CRISPR alleles. Based on this high polymorphism of the spacer contents, a microbead-based liquid hybridization assay, CRISPOL (for CRISPR polymorphism) has been developped for the serotype Typhimurium and its monophasic variant. This assay targets 72 of the 79 spacers identified previously as it is not possible, for the time-being, to distinguish between some of the remaining seven spacers by a Luminex approach. For example, STMB8var1 has a single SNP located in position 1 of the spacer compared with STMB8 or the four VNTR variants of STM18 only differ from each other by the number of an hexanucleotide repeat.

2. Materials

2.1 Consumables

-1.5 mL co-polymer microcentrifuge tubes (USA Scientific 415-2500 or VWR International or equivalent)

-1.5 mL Eppendorf Protein loBind microcentrifuge tubes (Fisher M54943 or VWR international 525-0133 or equivalent)

-2.0 ml screw-cap microcentrifuge tubes (Fisher 05669-8) or ambered safe-lock for storage (Fisher 51645) or equivalent

-0.2 ml PCR tubes-sterile (VWR International 732-0546) or (Dominique Dutscher, 116380) or equivalent

-10 µL, 250 µL, 1000 µL pipette tip refills (VWR or equivalent)

-96 well microplate aluminium sealing tape (Costar 6570 or Fisher or equivalent)

Thermowell 96-well P polycarbonate clear PCR Plates (Costar 6509 or Fisher or Equivalent) (recommended for XYP heater block)

-Sealing Mat for 96-well Thermowell P Plates (Costar of Fisher or equivalent)

-Microseal 'A' film (MJ Research)

-1.2 µM PVDF filter microliter plates (Millipore)

-1µM PTFE filter microtiter plates (Millipore)

-1.2 µM Supor filter microtiter plates (Pall Life Sciences or VWR)

-96-well black half-area flat bottom plates (Costar)

-96-well half-area flat bottom plates (Costar, 3693)

(white, non treated, recommended for no-wash assays)

-96-well round bottom polystyrene solid plates (Costar)

(recommended for no-wash assays) (Costar 3694)

2.2. Chemicals

-MES (2[N-Morpholino] ethanesulfonic acid) (Euromedex, EU0033-B) or (Sigma M2933) or equivalent

-NaOH (Fisher SS256-500) or equivalent

-TWEEN 20 (Polyoxyethylenesorbitan monolaurate) (VWR international 28829-296) or equivalent

-SDS (Sodium lauryl sulfate) powder (MP Biomedicals, 811030) or 10% solution (Sigma L4522)

-EDTA powder (Euromedex EU0007) or Tris-EDTA Buffer, pH 8.0, 100X (Sigma T9285) or equivalent

-TMAC powder (MP Biomedicals, 152113or Fisher M0704L) or 5M TMAC (Sigma T3411)--

N-Lauroylsarcosine (sarkozyl) sodium salt 20% solution (Sigma T7414) or equivalent

-1 M Tris-HCl, pH 8.0 and 0.5 M EDTA, pH 8.0 (prepared from powder or -Tris-EDTA Buffer, pH 8.0, 100X (Sigma T9285) or any supplier, molecular biology grade

-Molecular Biology grade water

-SSPE, 20X (Phosphate buffer, pH 7.4, sodium chloride, EDTA) (Sigma S2015) or equivalent

-Triton X-100 (MP Biomedicals, 807426) or equivalent

-Betain Chlorhydrate (VWR International, 8.14633.1000 or Euromedex BK185-A) or equivalent

-DMSO (Euromedex, UD8050-05-A) or equivalent

-EDC (1-ethyl-3-[3dimethylaminopropyl carbodiimide hydrochloride (Perbio 22981 or Pierce 77149 or 22980) or equivalent

2.3 Buffers

-0.1 M MES, pH 4.5 (Coupling buffer)

for 250 ml : weight 4,88 g of MES, complete to <250ml, adjust pH with NaOH (≅5 drops), complete to 250 ml, filter sterilize and store at room temperature

-0.02% Tween 20 (Washing Buffer) pipet 50 µL of Tween 20, complete to 250 mL with water, filter sterilize and store at room temperature

-0.1% SDS (Washing buffer): pipet 2.5 mL of SDS 10%, complete to 250 mL with water, filter sterilize and store at room temperature

-Tris-EDTA (TE) Buffer, pH 8.0, 1X: pipet 2.5 mL of TE 100X, complete to 250 mL with

water; Filter Sterilize and store at Room Temperature

-1.5 X TMAC Hybridization Solution (Microsphere diluent) : pipet 225 mL of TMAC 5 M, add 1.88 mL of 20% Sarkosyl solution, add 18.75 ml of 1 M Tris-HCl, pH 8.0, add 3 mL of 0.5 M EDTA, pH 8.0, complete with water to 250 mL (1.37 mL) Store at room Temperature

-1.0 X TMAC Hybridization Solution (Microsphere diluent) : pipet 150 mL of TMAC 5 M, add 1.25 mL of 20% Sarkosyl solution, add 12.5 ml of 1 M Tris-HCl, pH 8.0, add 2 mL of 0.5 M EDTA, pH 8.0, complete with water to 250 mL (84.25 mL) Store at room emperature

2.4 Equipment

-Centrifuge for 96-well plates (Eppendorf 5804) (VWR 53513-800 or Fisher 05-400-90 or Eppendorf 022622501)

-Sonicator (mini) (Cole Parmer 08849-00) or Ultrasonic cleaner, Branson 200 (VWR international, 142-2595)

-Bench Microcentrifuge (VWR international, 521-2844)

-Vortex Mixer (VWR international)

-DNA Engine PTC200 (Biorad) or Equivalent

-Luminex 200®, BioPlex200® or MagPix® (Luminex Corp, Austin, TX ; Biorad, Hercules, CA)

-Hemacytometer (Sigma or equivalent) or Bio-rad TC20 cell counter (BioRad)

-Refrigerated Eppendorf centrifuge (Sigma, Eppendorf or equivalent)

-Pipetors P10, P20, P100, P100, 8 Ch. (Rainin or equivalent) ; we use Pipetman and Pipetman concept 8 Ch. (0.5 µL-10 µL) (Gilson, Villiers le Bel, France)

2.4 Oligonucleotides, Enzymes, PCR reagents, Microbeads

-5’ amino-C12 linker oligonucleotides (IDT, or Eurogentec) ; capture probes according to

published sequences

The specific list of probes for the TB-SPOL (43-Plex) and STM-CRISPOL (72-Plex) are found in Kamerbeek *et al.* modified by van Embden *et al.* for MTC and in Fabre *et al.* for STM [2, 4, 18].

-standard and biotinylated oligonucleotides (IDT or Eurogentec) ; PCR primers.

Specific Primers for MTC-spoligotyping (TB-SPOL) :

DRa (5′Biot-GGTTTTGGGTCTGACGAC-3′) and DRb (5′-CCGAGAGGGGACGGAAAC-3′).

Specific Primers for STM-CRISPOL :

DRSTMA (5′-CCGCTGGCGCGGGGAACA-3′) and DRSTMB (5′Biot-CGCCAGCGGGGATAAACC-3′)

-MicroPlex® Microspheres Regions 1 through 100, choose 43 regions for TB-SPOL, 72 regions for STM-CRIPSOL among LC10001 to LC10100 references or MagPlex® among MC10012 to MC10100 references (Luminex, Austin, TX).

-or MagPlex xMap microspheres Regions 1 through 1000 (to run only on MagPix®).

-Taq DNA Polymerase, deoxynucleotides, and buffers (Homemade or Promega)

-Streptavidin-R-Phycoerythrin Lumi Grade (Roche, 05065925103) or Invitrogen (Molecular Probes S-866) or (Interchim FP-77776A)

-Separately purchased Oligonucleotides and Luminex Microbeads can advantageously be replaced by directly available, quality controled coupled-microspheres sold by the Institut of Genetics and Microbiology upon request (IGM, Orsay, France) or by purchasing full reagent kits : TB-SPOL (43-Plex) and STM-CRISPOL (72-Plex), that also include Taq Polymerase, buffers, positive controls, and Streptavidin-Phycoerythrin ; please visit www.igmors.u-psud.fr « beads4med » and/or contact us for quotes.

2.5 DNA

The quality of DNA may vary from crude to purified. For MTC, crude thermolysates, purified Cetyl-Trimethylammonium bromide (CTAB) extracted, or any commercial kit-extracted DNA can be used as templates. For STM, the following DNA extraction protocol can be followed. Take a 10 µl loop of bacteria and suspend it in 200 µl of molecular biology-grade water. The suspension is vortexed for 10s, incubated at 95°C for 10 min and then centrifuged for 5 min at 10.000 g in a Jouan A14 centrifuge. The supernatant is transferred to a 1.5 ml microtube and stored at −20°C until use.

3. Methods

3.1 Generalities

Since the advent of multiplexed analyzers, an alternative to membrane-based spoligotyping is high-throughput microbead-based spoligotyping [25]. The transfer from the membrane-based towards the microbead-based format was indeed achieved in USA by Cowan *et al.* at the CDC-Atlanta in 2004 and in 2009 by Zhang *et al.* at University Paris-Sud [1, 26] and by the Institut Pasteur in 2011 on *S. enterica* serotype Typhimurium [2, 12, 13]. However, for cost reasons, in many laboratories worldwide the technique still relies on a membrane-based procedure run on the Immunetics miniblotter. Other alternative techniques, e.g. MALDI-TOF (Matrix-Assisted Laser Desorption/Ionisation time-of-flight) mass spectrometry-based Spoligotyping were also recently developed and will not be described here [27].

Briefly the principle of the Luminex system relies on the use of polystyrene or magnetic colored microbeads of different types (up to 500 types in the latest FlexMap 3D® version, 100 types on Figure 3A), that can be individually recognized by a laser (L1) in a microfluidic system (Figure 3B and 3C). On each set of beads it is possible to link a large variety of sensor-targets (antibodies, antigens, nucleic acids) that can thus be individually assessed. In

our case, these markers are amino-linked oligonucleotide with a C12 linker. The second laser of the system (L2, Figure 3B), combined with a second optical mean (in our case Streptavidin-Phycoerythrin or SA-PE), allows to detect the microbead-fixed ligands thus permitting the quantification of results on each microbead type. Alternatively to the use of the biotin/SA-PE detection principle, 5'labelled oligonucleotides using Cyanine or Alexa Fluor markers can also be used for quantification by L2. As many as 500 analytes can theoretically be individually assayed in a unique sample. In Figure 3C-right, two microbead types are represented, type 1 and type n, each previously coupled with a specific oligonucleotide (DR1 to DRn capture sequence).

Users that have been previously producing their own spoligotyping membranes will easily be able to produce spoligotyping microbeads. Chemical constraints or precautions to link oligonucleotides to membranes are not much different. A simple list of requirements is to be followed : (1) always use low-binding Eppendorf tubes since polystyrene microbeads may adsorb to classical polypropylene tubes ; (2) order 5'-amino oligonucleotides with a C12 amino-link instead of a C6 amino-link arm for membranes to increase gyration radius ; (3) keep EDC powder frozen in aliquots at -20°C and do not reuse freshly prepared solutions.

3.2. Protocols

3.2.1. Coupling of oligonucleotides to microbeads

Step 1 : microbead washing (also cf. section 4, Note 1)

1. Let some fresh aliquot of EDC powder come back from -20°C to laboratory temperature.

2. Dissolve amino-linked oligonucleotides ("probe" or "capture" oligo) to 1 mM (1 nanomol/microliter) in sterile water.

3. Resuspend the microbeads stock by vortexing (20 s) and sonication (20 s).

4. Transfer 5.10^6 (400 µL) of MicroPlex® or MagPlex® microbeads stocks into LowBind

Eppendorf tubes.

5. Centrifuge microbeads at 8000 g during 1-2 min.

6. Discard supernatant and resuspend the microbeads in 50 µL of 0.1M MES (2-(N-morpholino)ethanesulfonic acid) buffer pH=4.5 by vortexing and sonicating during 20 s. This washing step is done to eliminate microbeads conservation buffer including antimicrobial agents (cf. section 4, Note 2).

Step 2 : Chemical coupling of oligonucleotides to microbeads (cf. Section 4, Note 3)

7. Prepare a 1/10 dilution of capture oligonucleotides in sterile water (0.1 nanomole/microliter ; 0.1 mM or 100 µM).

8. Add 2 µl (0.2 nmol) of the 1/10 oligo-solution prepared above to the resuspended and vortexed microbeads.

9. Prepare a fresh EDC solution (10 mg/ml) in sterile water.

10. One by one, for each coupling reaction, add 2.5 µl of freshly prepared EDC solution to the microbeads, mix by vortexing.

11. Incubate during 30 min at room temperature in a dark room.

12. Prepare a new fresh EDC solution (10 mg/ml) in sterile water (cf. Section 4, Note 4)

13. Again, one by one, for each coupling reaction, add 2.5 µl of freshly prepared EDC solution to the microbeads, mix by vortexing.

14. Again, Incubate during 30 min at room temperature in a dark room.

Step 3 : Washing of coupled microbeads.

The microbeads are washed successively with Tween-20 and SDS to prevent microbeads aggregation and adsorption to Eppendorf tubes walls, as well as to block hydrophobic sites on the microbeads surfaces.

15. Add 1.0ml of 0.02% Tween-20 to the coupled microbeads (cf. Section 4, Note 5).

16. Centrifuge the coupled microbeads at 8000 g during 1-2 min.

17. discard the supernatant and resuspend the coupled microbeads within1 ml of 0.1% SDS by vortexing (cf. section 4, Note 6).

18. Centrifuge the coupled microbeads by centrifuging at 8000 g during 1-2 min.

19. Discard the supernatant and resuspend the coupled microbeads in 100 µl TE 1X pH=8 by vortexing and sonicating during approximately 20 s (cf. Section 4, Note 7).

20. Store coupled microbeads between 2 to 8°C protected from light. Coupled beads can still be used after 6 months.

Step 4 ; Counting Microbeads on an hemacytometer.

1. Dilute the coupled microbeads 1/100 in sterile water.

2. Load at the proper place 10µL of microbeads dilution in an hemacytometer (cell counter).

3. Count all microbeads seen in one of the corner within a 4x4 grid as shown below (the model of classical hemacytometer may vary from country to country).

4. compute the result by using the following formulae

total microbeads number =

(Number read on 4x4 grids corner) x (1x 10^4) x (dilution factor) x (volume of microbeads suspension in mL).

The beads mix might alternatively be prepared and controled on a TC20 Cell Counter (Biorad, Hercules, CA) which provides the easiest way to check bead counts.

Step 5. Control of oligonucleotide-coupling. The Objective of this step is to verify the right fabrication of oligonucleotides coupled MicroPlex or MagPlex microbeads reagents using complementary probes

1. Defrost 1µL of biotinylated target stock oligonucleotides (antisense nucleotides).

2. Dilute the target oligonucleotides at 10fmoles/µL in TE 1X, pH=8.

3. Select the appropriate coupled microbeads and biotinylated target oligonucleotides sets to control.

4. Resuspend the microbeads by vortexing and sonication during approximately 20 s.

5. Prepare a « Microbeads working Mix » with 75 beads/µL in 1mL of TMAC 1.5X hybridization Solution (Add 1.5 µL of each coupled microbeads stock to 998,5 µL of TMAC 1.5X hybridization Solution).

6. Mix the « Microbeads working Mix » by vortexing 20 s and sonication during 20 s.

7. In each assayed well of a 96-well plate, including the negative control add 33µL of « Microbeads working Mix ».

8. in the negative control well(s), add 17µL of TE 1X, pH=8.

9. In each tested well, add the mixture of biotinylated complementary oligonucleotides (5 to 200 fmoles) and TE 1X, pH=8 up to a total volume of 17µL.

10. Smoothly mix the tested wells by pipeting up and down a couple of times.

11. Cover the 96-well plate to prevent evaporation and incubate at 95-100°C during 3 min to break all oligonucleotides secondary structures.

12. Incubate the 96-well plate at the same temperature as the one used for the PCR-product hybridization assay (52°C for TB-SPOL and 59°C for STM-CRISPOL) during 20 min (cf. section 4, Note 8).

13. Centrifuge the plate during 1 min, eliminate as much as possible of the supernatant (25-35µl) by pipeting carefully, replace with 25-35 µl of TMAC 1X.

14. Prepare a fresh « Reporter Mix » by adding 4µL of Streptavidin-R-phycoerythrin in 996µL of TMAC 1X hybridization solution to obtain a 4µg/ml solution (4/1000 of 1mg/ml stock dilution).

15. Add 25µL of « Reporter Mix » to each well and mix smoothly by pipeting up and down a few times

16. Incubate the 96-well plate at 52°C during 10mn

17. Analyze 50µL at 52°C with the Luminex 200/BioPlex analyzer using user's manual (cf. section 4, Note 9).

3.2.2. High-Throughput MTC and STM spoligotyping protocols on Luminex® 200 or BioPlex®.

Step 1 : PCR protocols for MTC-Spoligotyping (DRa-DRb) and STM-CRISPOL (DRSTMA-DRSTMB)

A. PCR protocol for MTC-Spoligotyping (Dra-DRb)

DNA (CTAB-extracted or thermolyzates)	2µL
dNTPs 2µM	2.5µL
biotinylated-Dra 5µM	2.5µL
Drb 5µM	2.5µL
Betain 5X*	5µL
Q Buffer 10X*	2.5µL
H_2O	8µL
Home made or commercial Taq Pol*(1U)	0.1µL
Total volume	25µL

(*cf. Section 4, Note 10)

Cycling conditions :

96°C-3mn; (96°C-30s; 55°C-30s; 72°C-15s)*; 72°C-5mn; *repeat 20 cycles for CTAB, 25

cycles for thermolysates DNA (cf. Section 4, Note 11)

B. PCR protocol for STM-CRISPOL

DNA (thermolysate)	1µL
dNTPs	200µM
biotinylated-DRSTMB	50pmol
DRSTMA	50pmol
MgCl2	1.5 mM
Go Taq Promega	1.25 U
Go Taq Promega buffer 5X	10µL
H_2O up to a total volume of	50µL

Cycling conditions :

95°C-2mn; (95°C-1mn; 59°C-30s; 72°C-15s)* repeat 20 cycles

(cf. section 4, Note 12)

Step 2 : Generic high-throughput hybridization protocol in 96 wells plates

1. Choose the appropriate set of oligonucleotides-coupled microbeads.

2. Resuspend the microbeads by vortexing 20 s and sonicating 20 s.

3. Prepare a « microbead working mix » at 75 microbeads/µL in a total volume of 1mL of TMAC 1.5X (hybridization solution).

Add 1.5µL of each individual coupled-microbead stock (at 50.000 coupled microbeads/µl) within 1000-(1.5*N) µL of TMAC 1.5X where N is the multiplexing level.

(cf. Section 4, Note 13)

4. Mix the « microbead working mix» by vortexing 20 s and sonicating for 20 s.

5. Distribute 33µL of the « microbead working mix » in each sample well and controls.

6. Dispense 17µL of TE 1X, pH=8 in the negative control well.

7. In the wells which contains samples, add biotinylated PCR-amplified DNA and TE 1X, pH=8 in a total volume of 17µL.

(cf. Section 4, Note 14)

8. Mix gently by pipeting up and down a few times.

9. Place a lid on the reaction plate to prevent evaporation and incubate at 95-100°C for 10 min to denature the amplified biotinylated DNA (PCR product).

10. Incubate at 52°C (MTC-Spoligotyping) or 59°C (STM-CRISPOL) for 20 min.

11. Centrifugation and supernatant elimination.

For TB-SPOL, Centrifuge the plate during 1 min, eliminate as much as possible of the supernatant (25-35µl) by pipeting carefully. This step is done only when working with MicroPlex® (polystyrene) beads. This is used to lower the background and can be replaced by filtration if working with filter plates. Alternatively If using MagPlex® beads, you can use a magnet and simple upside down move of the plates. Filter plates and filtration can also be used.

For STM-CRISPOL, Centrifuge the plate during 3 min, eliminate as much as possible of the supernatant by pipeting carefully.

12. Prepare a fresh « reporter mix» by adding 4µL (TB-SPOL) or 1.25µL (STM-CRISPOL) of Streptavidin-R-phycoerythrin (1mg/ml stock) to 996µL (TB-SPOL) or 999µL (STM-CRISPOL) of TMAC 1X Hybridization solution.

13. Detection step

For TB-SPOL**,** Resuspend the beads by adding as much of TE 1X (cf. Section 4, Note 15) than the TMAC quantity you removed at the step 11, in general 25-35 µl. Add 25 µl of this fresh « reporter mix » to each well and mix smoothly by pipeting up and down.

Resuspend the beads by adding 90 µl of this fresh « reporter mix » to each well and mix smoothly by pipeting up and down.

Step 4 : Results interpretation ; basic knowledge, advanced knowledge

The signals generated by the instrument are of two kinds : (1) real-time acquisition of data, which may allow direct control of the success/failure of the experiences run on the instrument (2) final results data points files with quantitative MFI (mean fluorescence intensities) measures (Figure 4A). The raw output is a .csv file (Figure 4A) that can be easily transferred to .xls files, that are processed and analyzed using specifically designed macros, transforming the analogical signals (MFI) into digital values (positive/negative) after cut-off computation For CRISPR data analysis, numerical results are converted into binary states results (presence/absence) of a given sequence and translated into characters (white or empty squares, to create a string or « spoligotype », Figure 2 and 4B). The distribution of negatives RFI compared to positive RFI shows a bimodal distribution (Figure 4C), the full raw data file after cut-off calculation may be translated into a colour-code (pink= presence, white= absence as shown on Figure 4D).

The computation of cut-offs will vary depending on instrument fine tuning and on techniques. Briefly, cut-offs are defined statistically using Mean Fluorescence Intensities (MFI) ± 2 standard errors when possible and ROC (Receiver Operating Characteristic) curves. If the experimental lowest reference positive results are always superior to the highest reference negative results, the technique is optimal. In some cases (e.g. 1-2% of data points for the TB-SPOL technique) and only for suboptimal techniques, it may happen that a gray zone (zone for which the results may be positive or negative) has to be defined. In that case, rerunning the samples with doubtful results, and/or expert interpretation is required. This interpretation

is often achieved, given (1) the level of signals obtained for other spacers that allow to provide indication on the quality/quantity of the starting material, (2) the neighborhood of the spacer, (3) available databases. Unique patterns should be re-run many times beforing inferring too rapid wrong conclusions.

For the sophisticated CRISPR-SNPs based method (not described in detail here), an adequate interpretation is also achieved thanks to the existence of internal wild-type/mutated alleles controls, since our probe design always uses two beads to check results, one with the wild-type nucleotide, one with the mutation assessed. It is out of the scope of this technical review to enter into all details and the reader will refer to our latest developments for further information [2, 28-32].

The full informativeness of spoligotyping profiles is then achieved through local or global database comparisons and mathematical similarity analysis. For MTC-Spoligotyping (TB-SPOL), updated SpolDB4.xls files are freely available from the author (CS). Alternatively the Institut Pasteur of Guadeloupe is managing a world-wide spoligotyping database that accepts scientific collaborations and new large data sets (*www.pasteur-guadeloupe.fr:8081/SITVITDemo*) and that, by the time of writing contains more than 110.000 patterns from more than 160 countries of origin (N. Rastogi, personal communication). For STM-CRISPOL, the CRISPOL database of the Pasteur Institute, Paris contains more than 7,000 strains resulting in more than 750 different CRISPOL types (CTs). This database will be soon made publicly available through a web tool. In the meantime, readers could contact the authors (SLH and FXW) to obtain the different validated CTs.

To conclude, we described in this chapter some of the technological, scientific and practical aspects of high-throughput CRISPR typing for *Mycobacterium tuberculosis* complex (TB-SPOL, 43-Plex method) and for *Salmonella enterica* serotype Typhimurium (STM-

CRISPOL, 72-Plex method). Spoligotyping, even being a classic typing method was highly successful to describe the *Mycobacterium tuberculosis* genetic diversity, as well as within the molecular epidemiology research field. Molecular epidemiology is a moving field and Whole Genome Sequencing (WGS) will play an increasing role in the future [33-36]. However the advent of new technologies such as microbead-based suspension arrays, high-resolution melting or mass spectrometry will also provide complementary tools to give a second youth to spoligotyping and SNPs typing [26, 27, 37]. Moreover, the recently launched MagPix® new generation microbead-based assay with a decreasing cost and no need for air conditioning of the laboratory will undoubtedly create a growing environment of new Point Of Care diagnostic applications, similarly as what the iPhone has been for information technologies developers [38, 39]. Indeed, our increasing knowledge of CRISPR diversity on new bacterial targets could boost surveillance tools alternative to serotyping in combination with microbead-based hybridization systems [7].

Acknowledgments

The authors would like to thank Jian Zhang, Marie Accou-Demartin, Lucile Sontag, Saïana de Romans, Catherine Lim and Laëtitia Fabre. M. François Topin and M. Jan van Gils from Luminex BV, The Netherlands, are also acknowledged for technical support.

4. Notes

Note 1 : Microbeads should be protected as much as possible from prolonged exposure to light throughout this procedure.

Note 2: 0.1 M MES buffer provides adequate pH conditions for efficient carboxylic functions on microbeads -oligonucleotides-amino moiety coupling.

Note 3 : Addition of oligonucleotide and EDC. There are two rounds of EDC addition to

maximize coupling efficiency. EDC has a very short half-life in solution; a single EDC addition is not sufficient to complete oligonucleotide coupling.

Note 4: the EDC solution must not be reused.

Note 5: 0.02% Tween 20 is a washing buffer that allows to discard non coupled material. It also helps to recover microbeads from tube walls.

Note 6 : 0.1% SDS is also a washing buffer that prevents microbeads aggregation

Note 7 : TE 1X pH=8 is a standard storage buffer of nucleic acids (TRIS 10 mM, EDTA 1mM, prepared by extemporaneous dilution of stocks: TRIS 1M pH=8 (1/100), EDTA 0.5M pH=8 (1/500)

Note 8: step 11 and 12 may be run in a thermocycler : Hold at 95°C, 3min ; Hold at 52°C or 59°C (see above), 20min

Note 9 : Preparing a microbead working mix is possible, this mix can be kept at +4°C for a couple of months (up to 6 months) if protected from light.

Note 10 : *Go Taq Promega buffers and Polymerases can also be used efficiently as an alternative.

Note 11 : the final 72°C-5mn elongation step is likely to be skipped in MTC since the products have a very short size, as this step was suppressed for the STM PCR Protocol

Note 12 : The PCR products can be checked by electrophoresis in 1.2% agarose gels (smears sizing 100-300 bp should be visible) and if not analysed by the Luminex® platform on the same day, they might be stored at −20°C for no more than three days. Negative (no PCR product, replaced with water) and Positive controls should always be used -H37Rv, *M. bovis* BCG P3 for MTC, SARA8, 81-784, 02-7015, 07-1777 (4 STM strains with a known spacer content covering all the spacers in the assay) for STM.

Note 13 : As an Example, if you run the 43-Plex assay (TB-SPOL), add 935.5 µl TMAC 1.5X to (43*1.5=64.5 µl of beads): if you run the 72-plex assay (STM-CRISPOL), add 892 µl

TMAC 1.5X to (72*1.5=108 µl of beads)

Note 14 : we routinely use 2 µl (1-7µL) of biotinylated PCR-amplified DNA for MTC-Spoligotyping and 7µL for STM-CRISPOL

Note 15 : the decrease in ionic strength is not a problem; at this step, you should normally add TMAC 1X, however evaporation during reading is much more intense with TMAC than with TE. We verified that TE substitutes without problem to TMAC 1X at this step.

5. References

1. Cowan, L.S., et al., *Transfer of a Mycobacterium tuberculosis genotyping method, Spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multianalyte profiling system.* J Clin Microbiol, 2004. **42**(1): p. 474-7.
2. Fabre, L., et al., *CRISPR Typing and Subtyping for Improved Laboratory Surveillance of Salmonella Infections.* PLoS One, 2012. **7**(5): p. e36995.
3. Groenen, P.M., et al., *Nature of DNA polymorphism in the direct repeat cluster of Mycobacterium tuberculosis; application for strain differentiation by a novel typing method.* Mol Microbiol, 1993. **10**(5): p. 1057-65.
4. Kamerbeek, J., et al., *Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology.* J Clin Microbiol, 1997. **35**(4): p. 907-14.
5. Jansen, R., et al., *Identification of a novel family of sequence repeats among prokaryotes.* Genomics, 2002. **6**(1): p. 23-33.
6. Makarova, K.S., et al., *A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action.* Biol Direct, 2006. **1**: p. 7.
7. Sorek, R., V. Kunin, and P. Hugenholtz, *CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea.* Nat Rev Microbiol, 2008. **6**(3): p. 181-6.
8. Makarova, K.S., et al., *Evolution and classification of the CRISPR-Cas systems.* Nat Rev Microbiol, 2011. **9**(6): p. 467-77.
9. Viswanathan, P., et al., *Regulation of dev, an operon that includes genes essential for Myxococcus xanthus development and CRISPR-associated genes and repeats.* J Bacteriol, 2007. **189**(10): p. 3738-50.
10. Pourcel, C., G. Salvignol, and G. Vergnaud, *CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies.* Microbiology, 2005. **151**(Pt 3): p. 653-63.
11. Mokrousov, I., et al., *Efficient discrimination within a Corynebacterium diphtheriae epidemic clonal group by a novel macroarray-based method.* J Clin Microbiol, 2005. **43**(4): p. 1662-8.
12. Weill, F.X., et al., *Molecular typing and subtyping of Salmonella by identification of the variable nucleotide sequences of the CRISPR loci*, 2009.
13. Fabre, L., et al., *Improving laboratory surveillance of Salmonella infections by fast*

typing based on CRISPR polymorphisms. IS2, International Symposium on Salmonella, 2010.
14. Liu, F., et al., *Subtyping Salmonella enterica serovar enteritidis isolates from different sources by using sequence typing based on virulence genes and clustered regularly interspaced short palindromic repeats (CRISPRs).* Appl Environ Microbiol, 2011. **77**(13): p. 4520-6.
15. Ginevra, C., et al., *Legionella pneumophila sequence type 1/Paris pulsotype subtyping by spoligotyping.* J Clin Microbiol, 2012. **50**(3): p. 696-701.
16. Lopez-Sanchez, M.J., et al., *The highly dynamic CRISPR1 system of Streptococcus agalactiae controls the diversity of its mobilome.* Mol Microbiol, 2012. **85**(6): p. 1057-71.
17. Hermans, P.W.M., et al., *Insertion element IS987 from Mycobacterium bovis BCG is located in a hot-spot integration region for insertion elements in Mycobacterium tuberculosis complex strains.* Infect. and Immun., 1991. **59**: p. 2695-2705.
18. van Embden, J.D.A., et al., *Genetic variation and evolutionary origin of the Direct repeat locus of Mycobacterium tuberculosis complex bacteria.* J. Bacteriol., 2000. **182**: p. 2393-2401.
19. Filliol, I., et al., *Snapshot of moving and expanding clones of Mycobacterium tuberculosis and their global distribution assessed by spoligotyping in an international study.* J Clin Microbiol, 2003. **41**(5): p. 1963-70.
20. Gagneux, S., et al., *Variable host-pathogen compatibility in Mycobacterium tuberculosis.* Proc Natl Acad Sci U S A, 2006. **103**(8): p. 2869-73.
21. Gagneux, S. and P.M. Small, *Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development.* Lancet Infect Dis, 2007. **7**(5): p. 328-37.
22. Reed, M.B., et al., *Major Mycobacterium tuberculosis Lineages Associate with Patient Country of Origin.* J Clin Microbiol, 2009. **47**(4): p. 1119-1128.
23. Brudey, K., et al., *Mycobacterium tuberculosis complex genetic diversity : mining the fourth international spoligotyping database (SpolDB4) for classification, Population Genetics, and Epidemiology.* BMC Microbiol., 2006. **6**(6): p. 23.
24. Dale, J.W., et al., *Spacer oligonucleotide typing of Mycobacterium tuberculosis: recommendations for standardized nomenclature.* Int. J. Tuberc. Lung. Dis., 2001. **5**: p. 216-219.
25. Dunbar, S.A., *Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection.* Clin Chim Acta, 2006. **363**(1-2): p. 71-82.
26. Zhang, J., et al., *Mycobacterium tuberculosis complex CRISPR genotyping: improving efficiency, throughput and discriminative power of 'spoligotyping' with new spacers and a microbead-based hybridization assay.* J Med Microbiol, 2010. **59**(Pt 3): p. 285-94.
27. Honisch, C., et al., *Replacing reverse line blot hybridization spoligotyping of the Mycobacterium tuberculosis complex.* J Clin Microbiol, 2010. **48**(5): p. 1520-6.
28. Zhang, J., et al., *A first assessment of the genetic diversity of Mycobacterium tuberculosis complex in Cambodia.* BMC Infect Dis, 2011. **11**(1): p. 42.
29. Abadia, E., et al., *Resolving lineage assignation on Mycobacterium tuberculosis clinical isolates classified by spoligotyping with a new high-throughput 3R SNPs based method.* Infect Genet Evol, 2010. **10** (7): p. 1066-1074.
30. Abadia, E., et al., *The use of microbead-based spoligotyping for Mycobacterium tuberculosis complex to evaluate the quality of the conventional method: providing guidelines for Quality Assurance when working on membranes.* BMC Infect Dis, 2011. **11**: p. 110.

31. Gomgnimbou, M.K., et al., << *Spoligoriftyping* >> *a DPO-based direct-hybridization assay for TB control on a multianalyte microbead-based hybridization system.* J Clin Microbiol, 2012. **50**(10): p. 3172-9.
32. Gomgnimbou, M.K., et al., *"TB-SPRINT: TUBERCULOSIS-SPOLIGO-RIFAMPIN-ISONIAZID TYPING"; an All-in-One assay technique for surveillance and control of multi-drug resistant tuberculosis on Luminex® devices.* J Clin Microbiol, 2013. **51**(11): p. 3527-3534.
33. Ioerger, T.R., et al., *Genome analysis of multi- and extensively-drug-resistant tuberculosis from KwaZulu-Natal, South Africa.* PLoS One, 2009. **4**(11): p. e7778.
34. Schurch, A.C., et al., *High resolution typing by integration of genome sequencing data in a large tuberculosis cluster.* J Clin Microbiol, 2010. **48**(9): p. 3403-3406.
35. Schurch, A.C., et al., *Mutations in the regulatory network underlie the recent clonal expansion of a dominant subclone of the Mycobacterium tuberculosis Beijing genotype.* Infect Genet Evol, 2011.
36. Gardy, J.L., et al., *Whole-genome sequencing and social-network analysis of a tuberculosis outbreak.* N Engl J Med, 2011. **364**(8): p. 730-9.
37. Chen, X., et al., *Rapid Detection of Isoniazid, Rifampin and Ofloxacin Resistance in Mycobacterium tuberculosis Clinical Isolates Using High Resolution Melting Analysis.* J Clin Microbiol, 2011. **49**(10): p. 3450-3457.
38. Lin, A., et al., *Rapid O serogroup identification of the ten most clinically relevant STECs by Luminex microbead-based suspension array.* J Microbiol Methods, 2011. **87**(1): p. 105-110.
39. Bergval, I., et al., *Combined Species Identification, Genotyping, and Drug Resistance Detection of Mycobacterium tuberculosis Cultures by MLPA on a Bead-Based Array.* PLoS One, 2012. **7**(8): p. e43240.

Legends to Figures

Figure 1 : Principle of the MTC spoligotyping technique ; a single couple of primers (Dra-Drb) allows to amplify by PCR a set of overlapping fragments that will be further detected by hybridization on a 2D or 3D device if one primer (Dra) is biotinylated and a detectable reporter is added (Streptavidin-Peroxydase using chemoluminescence, streptavidin-Phycoerythrin using microbead-based laser detection)
Figure 2 : Example of an MTC membrane spoligotyping experiment (A) and a binary transcription of results done on Excel (B)

Figure 3 : Principle of laser-based multiplexed experiment. (A) A set of 100 different commercial beads in the red region are available (B) the L1 laser recognizes the beads and the L2 laser quantifies the signals that are present on « hairy » (oligonucleotide-coupled) and hybridized beads, by detection of SA-PE after previous hybridization with a single strand-biotinylated PCR amplified DNA fragment. (C) artistic drawing of the analyzer showing the power of lasers combined to microfluidic systems

Figure 4 : (A) raw output results file of the Luminex 200 (preliminary trial experiment on 9 out of a total 43 targets) compared to (B) reference membrane-based spoligotypes. Raw values in (A) can intuitively be correlated with patterns shown on (B). After running experiments with enough samples with presence/absence of spacers it is shown that the experimental distribution of negative results does not overlap with the distribution of positive results (C). After introduction of automatized, cut-off based, interpretation of results, a

coloured pattern (pink/presence of the spacer sequence, white/absence of the spacer sequence) is obtained on Excel spreadsheet files. These files are further transcribed using Macros into final black/white patterns, as shown in Figure 2. The same principles are applied to STM-CRISPOL results production.

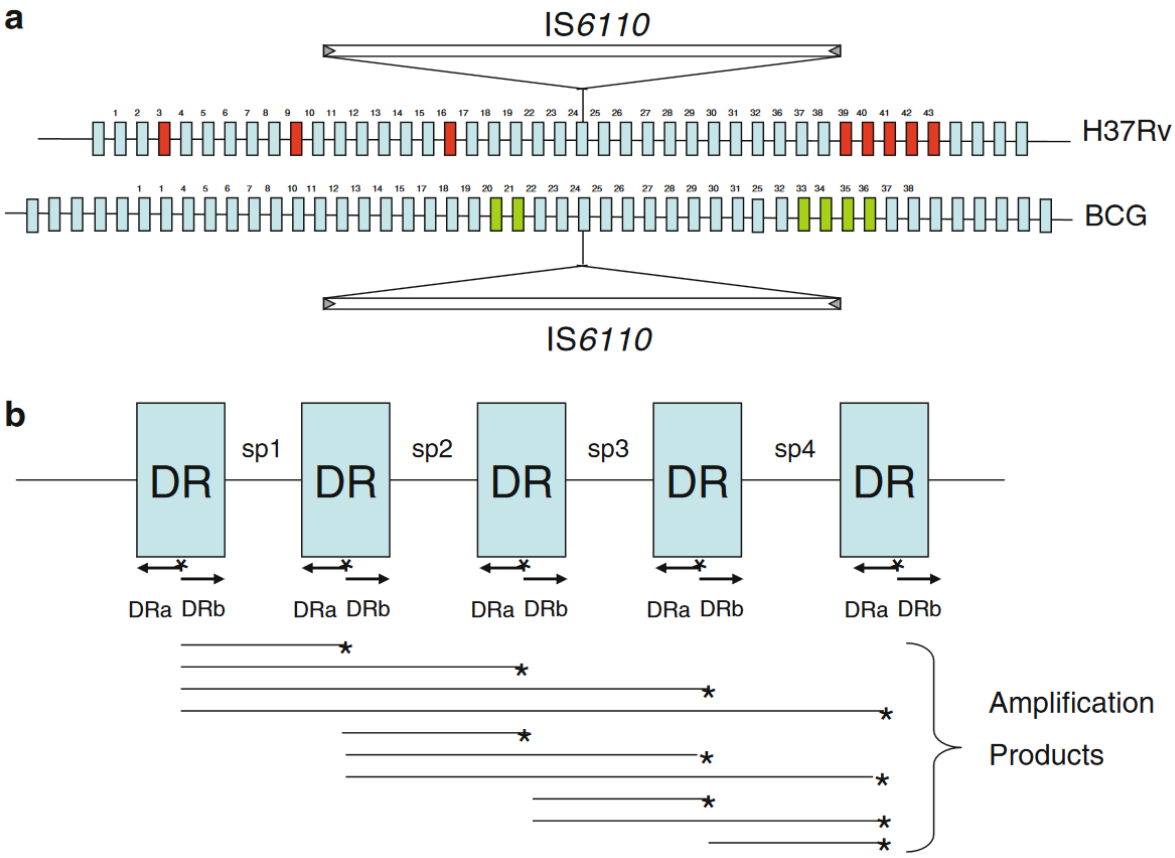
a
IS6110
H37Rv
BCG
IS6110
b
DR
sp1
DR
sp2
DR
sp3
DR
sp4
DR
DRa DRb
DRa DRb
DRa DRb
DRa DRb
DRa DRb
Amplification
Products

Figure 1.

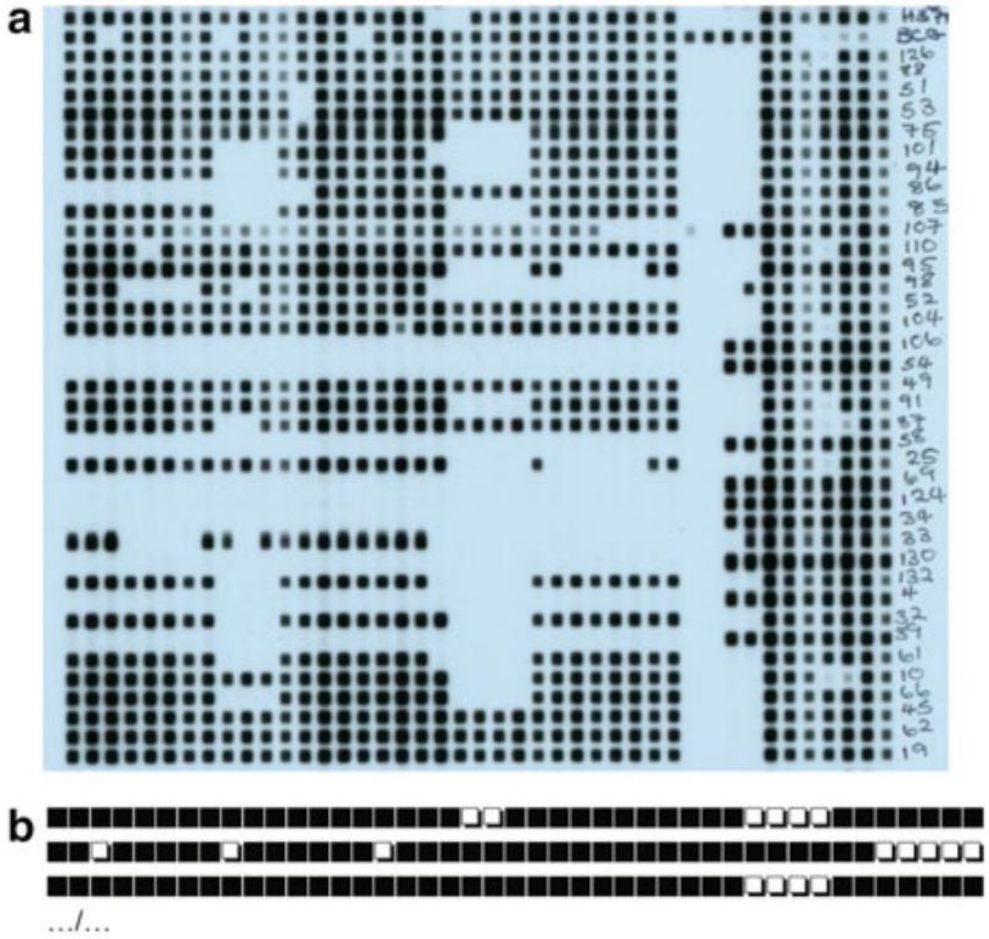

Figure 2.

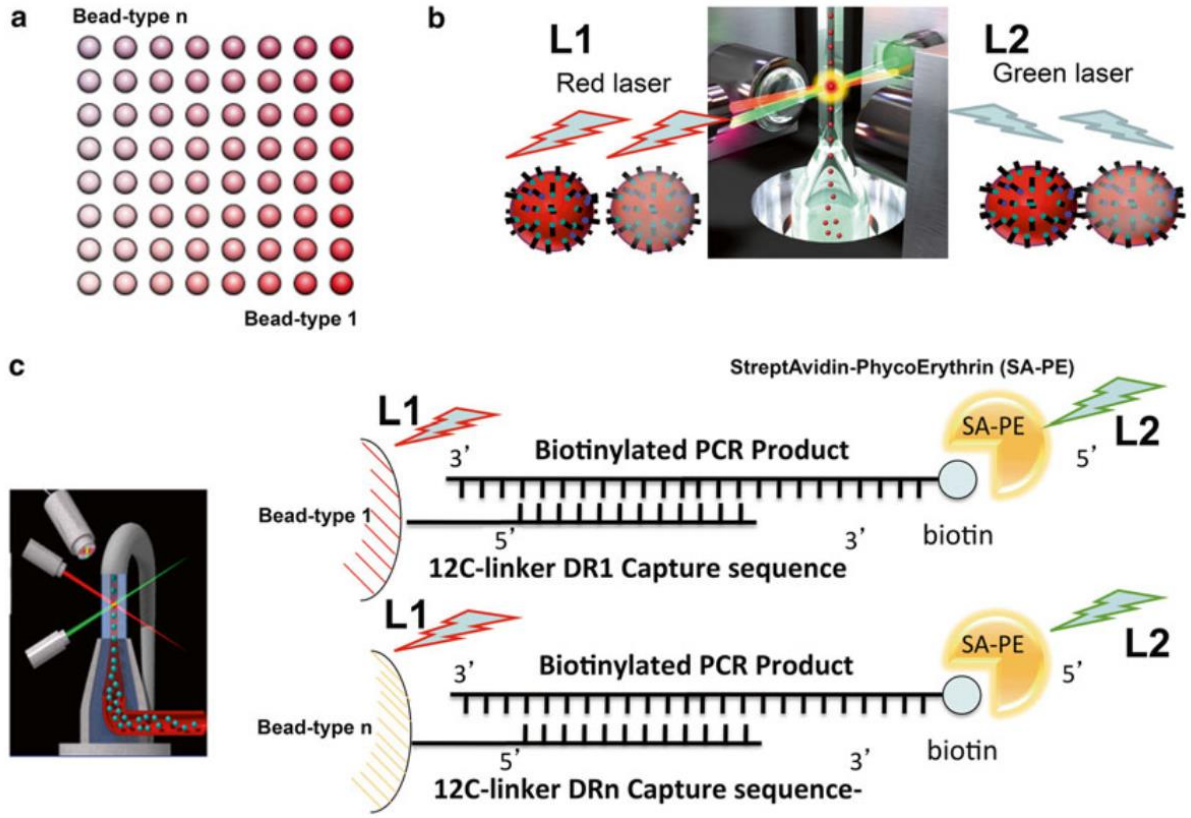

Figure 3.

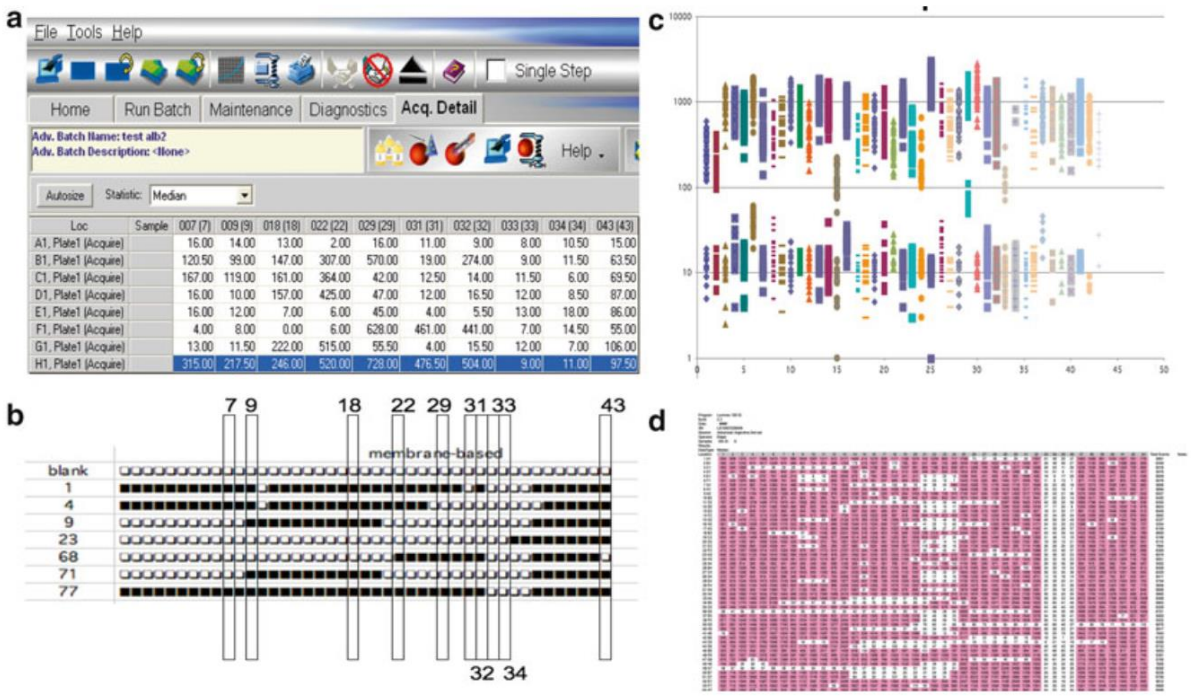

Figure 4.