Generation of barcoded plasmid libraries for massively parallel analysis of chromatin position effects

M.O. Lebedev[1, 2], L.A. Yarinich[1, 2], A.V. Ivankin[1], A.V. Pindyurin[1, 2] ✉

[1] Institute of Molecular and Cellular Biology, SB RAS, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
✉ e-mail: a.pindyurin@mcb.nsc.ru

The discovery of the position effect variegation phenomenon and the subsequent comprehensive analysis of its molecular mechanisms led to understanding that the local chromatin composition has a dramatic effect on gene activity. To study this effect in a high-throughput mode and at the genome-wide level, the Thousands of Reporters Integrated in Parallel (TRIP) approach based on the usage of barcoded reporter gene constructs was recently developed. Here we describe the construction and quality checks of high-diversity barcoded plasmid libraries supposed to be used for high-throughput analysis of chromatin position effects in *Drosophila* cells. First, we highlight the critical parameters that should be considered in the generation of barcoded plasmid libraries and introduce a simple method to assess the diversity of random sequences (barcodes) of synthetic oligonucleotides using PCR amplification followed by Sanger sequencing. Second, we compare the conventional restriction-ligation method with the Gibson assembly approach for cloning barcodes into the same plasmid vector. Third, we provide optimized parameters for the construction of barcoded plasmid libraries, such as the vector:insert ratio in the Gibson assembly reaction and the voltage used for electroporation of bacterial cells with ligation products. We also compare different approaches to check the quality of barcoded plasmid libraries. Finally, we briefly describe alternative approaches that can be used for the generation of such libraries. Importantly, all improvements and modifications of the techniques described here can be applied to a wide range of experiments involving barcoded plasmid libraries.

Key words: barcode; reporter construct; plasmid library; DNA cloning; Gibson assembly; massively parallel analysis; regulation of gene expression; regulatory DNA elements; chromatin position effects; cultured *Drosophila* cells.

Создание штрихкодированных плазмидных библиотек для множественного одновременного анализа эффекта положения гена

М.О. Лебедев[1, 2], Л.А. Яринич[1, 2], А.В. Иванкин[1], А.В. Пиндюрин[1, 2] ✉

[1] Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
✉ e-mail: a.pindyurin@mcb.nsc.ru

Открытие явления мозаичного эффекта положения гена и последующий тщательный анализ его молекулярных механизмов привели к пониманию того, что локальный состав хроматина оказывает существенное влияние на активность генов. Разработанный недавно метод Thousands of Reporters Integrated in Parallel (TRIP) основан на использовании штрихкодированных генов-репортеров и позволяет выполнять высокопроизводительный анализ эффекта положения гена в масштабе всего генома. В настоящей работе мы описываем конструирование и проверку качества штрихкодированных плазмидных библиотек высокой сложности, которые предполагается использовать для высокопроизводительного анализа эффекта положения гена в клетках дрозофилы. Во-первых, рассмотрены наиболее важные параметры, которые следует учитывать при создании штрихкодированных плазмидных библиотек, и предложен простой метод оценки сложности вырожденного фрагмента (штрихкода) в синтетических олигонуклеотидах с использованием ПЦР-амплификации и последующего секвенирования по методу Сэнгера. Далее проведено сравнение традиционного метода клонирования посредством рестрикции–лигирования с подходом сборки по методу Гибсона для клонирования штрихкодов в один и тот же плазмидный вектор. Кроме того, описаны оптимизированные параметры для создания штрихкодированных плазмидных библиотек, такие как соотношение вектор:вставка в реакции сборки методом Гибсона и напряжение, используемое для электропорации бактериальных клеток продуктами лигирования. Сравниваются также различные подходы для проверки качества штрихкодированных плазмидных библиотек. Наконец, кратко описаны альтернативные подходы,

которые можно использовать для создания таких библиотек. Важно отметить, что все улучшения и модифи-
кации методов, приведенные в данной работе, могут быть применены к широкому кругу экспериментов,
в которых используются штрихкодированные плазмидные библиотеки.

Ключевые слова: штрихкод; репортерная конструкция; плазмидная библиотека; клонирование ДНК; сборка
методом Гибсона; множественный одновременный анализ; регуляция экспрессии генов; регуляторные эле-
менты ДНК; эффект положения гена; культивируемые клетки дрозофилы.

Introduction

In a eukaryotic cell, an enormous number of specialized pro-
teins are responsible for dense packaging of very long genomic
DNA molecules into a nucleus. These proteins also play a
crucial role in the maintenance and replication of the genome
as well as in gene expression. These DNA-protein complexes
are referred to as chromatin, which is classically subdivided
into loosely packed, transcriptionally active euchromatin and
more tightly packed, transcriptionally inactive heterochroma-
tin (Babu, Verma, 1987; Grewal, Moazed, 2003; Huisinga
et al., 2006). Recent genome-wide localization studies of
chromatin components revealed up to 15 principal chromatin
types depending on the nature of cells. These chromatin types
differ in composition and gene activity levels. Furthermore, the
activity of genes greatly varies even within the same chromatin
type (Filion et al., 2010; Ernst et al., 2011; Kharchenko et al.,
2011; Riddle et al., 2011).

To exclude the input of unique DNA regulatory elements
associated with each gene and thus to identify the pure influ-
ence of the local chromatin environment on gene activity,
it was proposed to insert the same reporter construct (as a
sensor) in different genomic loci. The idea of this assay was
based on the position effect variegation (PEV) phenomenon,
originally described in *Drosophila* (Elgin, Reuter, 2013). Such
analysis performed in cells of different species substantially
contributed to our understanding of the mechanisms control-
ling gene activity throughout the genome, although only up to
a few hundred/thousand of unique genomic sites were tested
in each study due to the laborious and time-consuming na-
ture of the assay (Gierman et al., 2007; Babenko et al., 2010;
Ruf et al., 2011; Chen et al., 2013). Subsequently, usage of
unique identifiers (barcodes, or tags, or indexes, or unique
identification DNA sequences) (Kinde et al., 2011; Blundell,
Levy, 2014; Quail et al., 2014; Vvedenskaya et al., 2015;
Kebschull, Zador, 2018) for labeling the reporters enabled the
development of a high-throughput mode of the analysis. The
modified method provided a possibility of studying thousands
integrated into the genome reporters in one relatively simple
and short experiment and eliminated the bias towards selec-
tion of reporter insertions in active chromatin regions. The
approach was named Thousands of Reporters Integrated in
Parallel (TRIP), and its pilot application to mouse embryonic
stem cells (mESCs) allowed characterization of chromatin
position effects at more than 27,000 distinct genomic loci
(Akhtar et al., 2013, 2014).

From a technical point of view, TRIP is a specialized va-
riant of massively parallel reporter assays (Patwardhan et
al., 2009, 2012; Melnikov et al., 2012), so its performance
relies on barcoded plasmid libraries. Such libraries consist
of reporter plasmid constructs, which are identical except
random and *a priori* unknown barcode sequences. Within a
library, all barcodes are typically of the same length, which
can vary from few bp up to several dozen bp depending on

the experimental setup. Thus, as soon as the plasmid vector
carrying all necessary components except the barcode (e. g.,
functional transposon terminal repeats, a promoter, a coding
DNA sequence, a transcriptional terminator, etc.) is avail-
able, the next technical task is to add the barcodes into such
plasmid. Several different cloning strategies can be used for
this purpose, but the easiest way to get a DNA fragment car-
rying barcodes (hereafter barcoded fragment or insert) is to
PCR-amplify it using synthetic oligonucleotides containing
a fragment of random sequence (the barcode).

Theoretically, the maximum complexity of a barcoded
plasmid library or, in other words, the total number of unique
molecules or clones in the library is determined by the barcode
length and its nucleotide composition. More specifically, this
parameter depends on whether all four nucleotides or just a
subset of them are used at each position within the barcode.
For example, a random 4-nucleotide 18-bp sequence can
generate over $\sim 6.8 \times 10^{10}$ (4^{18}) different barcodes, while a
random 3-nucleotide 20-bp sequence can give rise to only
$\sim 3.5 \times 10^{9}$ (3^{20}) unique sequences. Although such high
numbers are never achieved in practice due to the necessity
to propagate plasmids in bacterial cells that typically limits
the complexity to 10^{6}–10^{7} clones, the diversity of synthetic
barcoded oligonucleotides is nevertheless crucial for the
construction of high-quality barcoded plasmid libraries.
Indeed, a strong overrepresentation of one or two particular
nucleotides within the barcode in synthetic oligonucleotides
extremely reduces the complexity of the final plasmid library.
In addition, it leads to the appearance of very similar barcode
sequences, which can substantially complicate unambiguous
identification of barcodes during the processing and analysis
of raw TRIP datasets generated by high-throughput DNA
sequencing (HTS). Furthermore, note that barcoded plasmid
libraries frequently contain some amounts of the original
plasmid vectors due to limitations of cloning techniques. Such
contamination should be kept at the lowest possible level or,
ideally, completely eliminated to increase the proportion of
useful reads in raw TRIP datasets. All these parameters, the
total number of unique barcodes, the randomness and diversity
of their sequences, and the level of the original vector contami-
nation, determine the quality of barcoded plasmid libraries.

We aimed to find out the simplest and most efficient way
to construct high-quality barcoded plasmid libraries for sub-
sequent TRIP experiments. In doing this, we first developed a
fast and simple test to check whether the quality of synthetic
oligonucleotides containing a barcode meets the expectations
before the generation of plasmid libraries. Next, we found
that the Gibson assembly approach was superior over the
conventional restriction-ligation cloning for the construc-
tion of barcoded plasmid libraries in terms of both the total
number of unique clones obtained per the same amount of
the vector used and the level of original vector contamina-
tion. Finally, we provide a set of quality checks of barcoded

М.О. Лебедев, Л.А. Яринич
А.В. Иванкин, А.В. Пиндюрин

Создание штрихкодированных плазмидных
библиотек для анализа эффекта положения гена

2019
23·2

plasmid libraries to verify that they are suitable for subsequent TRIP experiments.

Materials and methods

Barcoded oligonucleotide primers and quality check of the randomness of their barcode sequence. Synthetic oligonucleotides containing 18-nt barcodes, within which all four nucleotides (A, C, G, and T) were supposed to be equally represented at each position, were ordered either from IDT (USA) with standard desalting purification (the PB-Barcode1-SalI-F primer) or from DNK-Sintez (Moscow, Russia) with PAGE purification (the PB-Barcode1-SalI-F primer) or from Biosynthesis (Novosibirsk region, Russia) with PAGE purification (the PB-Barcode-PI-X-Gibson-F primers). Barcoded PCR products were amplified with the following combinations of the forward barcoded/reverse primers – PB-Barcode1-SalI-F/3′-TR-EagI-R or PB-Barcode-PI-X-Gibson-F/3′-TR-EagI-R (sequences of all primers used in the study are listed in the Suppl. Table S1)[1] – using plasmids with the appropriate library indexes (Suppl. Fig. S1) as templates. Five microliters of the PCR products were incubated with Exonuclease I (NEB, Cat. No. M0293S) and Shrimp Alkaline Phosphatase (NEB, Cat. No. M0371S) at 37 °C for 30 min to remove primers and nucleotides. The treated products were subjected to direct Sanger sequencing with the 3′-TR-EagI-R primer.

Preparation of PCR products for construction of barcoded plasmid libraries. All DNA fragments used in the ligation or Gibson assembly reactions were amplified using Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Cat. No. F530L) according to manufacturer's recommendations at the following cycling conditions: 98 °C for 60 sec followed by 35 cycles with 98 °C for 30 sec, 58 °C for 30 sec, 72 °C for 1 min and a final extension step of 72 °C for 5 min. Amplified fragments were then column-purified with GeneJET PCR Purification Kit (Thermo Fisher Scientific, Cat. No. K0702) before use in subsequent reactions.

Construction of barcoded plasmid libraries by restriction enzyme cloning. To generate the barcoded plasmid libraries, the 3′-terminal repeat (TR) of the piggyBac transposon was amplified with primers PB-Barcode1-SalI-F and 3′-TR-EagI-R using the pPB-Promoter1-eGFP-LI plasmid as a template. The PCR product was digested with SalI-HF (NEB, Cat. No. R3138L) and EagI-HF (NEB, Cat. No. R3505L). Then 0.9 pmol (240 ng) of PCR product were ligated with 0.3 pmol (1000 ng) of pPB-Promoter1-eGFP-LI vector digested with SalI-HF and NotI (Thermo Fisher Scientific, Cat. No. FD0593) restriction enzymes, using T4 DNA ligase (Roche, Cat. No. 10799009001). The ligation mixture was digested with NotI restriction enzyme at 37 °C for 4 h in a volume of 100 μL to destroy the original pPB-Promoter1-eGFP-LI vector molecules and then purified with MinElute PCR Purification Kit (Qiagen, Cat. No. 28004); DNA was eluted with 10 μL of pre-warmed elution buffer. Five microliters of the purified ligation mixture were used to transform *Escherichia coli* TOP10 electrocompetent cells.

Gibson assembly of barcoded plasmid libraries. pPB-PromoterX-eGFP-LI vectors (where X is 0 through 6) were digested with SalI (Thermo Fisher Scientific, Cat. No. FD0644) and NotI (Thermo Fisher Scientific, Cat. No. FD0593) restriction enzymes to prepare linearized vector backbones. The 3′-TR of the piggyBac transposon was amplified with primers PB-Barcode-PI-X-Gibson-F and PB-Gibson-R1 using a pPB-Promoter0-eGFP-LI plasmid as a template. The PCR products were digested with DpnI restriction enzyme (NEB, Cat. No. R0176) at 37 °C for 4 h in a total volume of 100 μL in order to destroy plasmid template and thus minimize non-barcoded vector contamination in the libraries. Next, 60 fmol (200 ng) of the linearized vector and 0.30 pmol (~80 ng) of each DpnI-treated and column-purified PCR product were incubated with Gibson Assembly Master Mix (NEB, Cat. No. M5510AA) at 50 °C for 1 h. Then, the reaction mixture was diluted 10-fold with nuclease-free water and purified with MinElute PCR Purification Kit; DNA was eluted with 10 μL of pre-warmed elution buffer. Five microliters of the purified Gibson reaction products were used to transform *E. coli* TOP10 electrocompetent cells.

Electrotransformation of bacterial cells. *E. coli* TOP10 strain (F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac* ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G λ⁻) was used as the host to clone plasmid libraries. Electrocompetent cells with transformation efficiency of ~1.5×10^9 colony forming units (CFU) per 1 μg of supercoiled plasmid DNA were prepared using the previously reported protocol (Morrison, 2001). For transformation, 50 μL of the electrocompetent cells were mixed with 5 μL of purified DNA sample, transferred into a 0.1 cm cuvette (Bio-Rad), and electroporated at the following settings: 1.8 kV, 25 mF, and 200 Ω using the Gene Pulser electroporator with the capacitance extender (Bio-Rad) unless otherwise stated. Next, cells were grown in 1 mL of pre-warmed Luria–Bertani (LB) medium at 37 °C, 220 rpm for 1 h. Then, aliquots of cells (1/5,000) were spread on LB-ampicillin plates (LB medium supplemented with 15 g/L bacto agar and 100 μg/mL ampicillin) and allowed to grow overnight at 37 °C before manual counting of colonies, whereas the rest of cells were transferred into 500 mL of LB medium supplemented with 100 μg/mL ampicillin and cultured at 37 °C and 220 rpm overnight before plasmid isolation with GeneJET Plasmid Maxiprep Kit (Thermo Fisher Scientific, Cat. No. K0491).

Analysis of the complexity and quality of barcoded plasmid libraries. The complexity of each plasmid library was defined as the average number of colonies on LB-ampicillin plates (CFU per plate; see above) multiplied by 5,000. To measure the contamination of each library by the original plasmid vector, the appropriate control ligation or Gibson assembly reaction lacking the barcoded insert DNA fragment was always set up and processed in parallel with the main samples, except the plasmid isolation step. The contamination is reported as the proportion (percentage) of the number of colonies obtained for the control sample with respect to the number of colonies obtained for the main samples. The DNA sequences of barcodes in the plasmid libraries and randomly selected individual clones were analyzed by Sanger sequencing with the EGFP-int3 primer.

Preparation of barcoded plasmid libraries for Illumina sequencing. An equimolar mixture of barcoded plasmid libraries was used to transform *E. coli* TOP10 cells in order to obtain several thousand clones. The isolated mixture of such

---

[1] Supplementary Materials are available in the online version of the paper: http://www.bionet.nsc.ru/vogis/download/pict-2019-23/appx4.pdf

low-complexity plasmid libraries was used as a template for amplification of barcode sequences for their subsequent Illumina HTS. The amplification was done in duplicate by two rounds of PCR with Phusion High-Fidelity DNA Polymerase and 200 μM of each dNTP in a total volume of 25 μL. In the first round, 2 ng of the plasmid mixture and primers Libr-cDNA-for/Libr-cDNA-AN-rev (where N was 18 or 19) were used at the following cycling conditions: 98 °C for 60 sec followed by 15 cycles with 98 °C for 30 sec, 70 °C for 30 sec, 72 °C for 30 sec and a final extension step of 72 °C for 5 min. In this round, each of the duplicates was marked by a unique HTS index present within the Libr-cDNA-AN-rev primer. In the second round, 0.5 μL of the unpurified first round PCR products and primers Libr-P5-for and Libr-P7-rev were used at the following cycling conditions: 98 °C for 60 sec followed by 23 cycles with 98 °C for 30 sec, 61 °C for 30 sec, 72 °C for 30 sec and a final extension step at 72 °C for 5 min. The samples were purified with Monarch PCR & DNA Cleanup Kit (NEB, Cat. No. T1030S), quantified using the KAPA Library Quantification Kit (Roche, Cat. No. 07960255001), and sequenced on an Illumina MiSeq sequencer using a 150 cycle MiSeq Reagent Kit v3 (Illumina, Cat. Nos. 15043893 and 15043894). The expected structure of the DNA fragments subjected to HTS is shown in Suppl. Fig. S2.

HTS data analysis. The FASTQ files were analyzed with custom Python scripts. First, only reads having the expected structure (see Suppl. Fig. S2) were filtered for the downstream analysis. Briefly, no mismatches were allowed within sequences of HTS and library indexes, while up to 4 mismatches were allowed within the first constant part (CGCCAGGGTTTTCCCAGTCACAAGGGCCGGCCACAACTC) and 1 mismatch was allowed within the second constant part (CTCGATC). The barcodes were defined as sequences of any length present between (i) the first constant part and (ii) the sequence consisting of the CCTC motif, the library index, and the second constant part. Next, the reverse complement sequences of barcodes present in the filtered reads were divided into individual pools according to the associated HTS and library indexes. For each pool, unique barcode sequences were identified and subjected to length distribution analysis. To analyze the frequency distribution of the four nucleotides per position within barcodes, only unique 18-nt barcodes were considered.

Results and discussion

Quality control of barcoded oligonucleotides
The main problems associated with preparations of long (>30 nt) synthetic barcoded oligonucleotides, which are the key element in the cloning of barcoded plasmid libraries, are (i) single-base deletions/insertions arising from incomplete coupling of nucleotide monomers during oligonucleotide synthesis and (ii) unequal representation of different nucleotides at each position along the barcode. PAGE purification of the synthesized oligonucleotides can partially reduce the first problem, but, to the best of our knowledge, no fast, simple, and cheap approach to assess the randomness of barcode sequences in an oligonucleotide preparation has been described so far. Instead, Sanger sequencing of plasmids isolated from dozens of randomly selected bacterial colonies, as well as
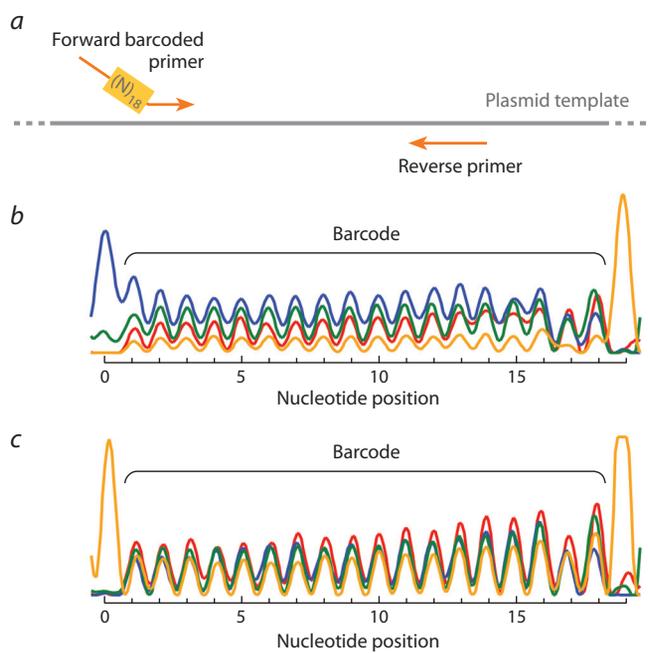
Fig. 1. The fast and simple approach to assess the diversity of barcoded oligonucleotides.

(*a*) Scheme of the DNA template and oligonucleotide primers used to prepare the barcoded PCR product. Note that the forward primer contains a random sequence of 18-nt long, the barcode. The obtained PCR product is purified and Sanger sequenced from the reverse primer. (*b*, *c*) Sanger sequencing chromatograms of the representative PCR products generated using (*b*) low-quality and (*c*) high-quality barcoded primers.

Sanger or HTS of barcoded plasmid libraries were used in previous studies (Akhtar et al., 2013; Wong et al., 2016). In addition to being laborious, these approaches require a ready barcoded plasmid library.

In this study, we developed a simple and fast protocol to check the randomness of barcode sequences within synthesized primers by using Sanger sequencing prior to the generation of plasmid libraries. The procedure includes amplification of the barcoded PCR product from a non-barcoded plasmid template followed by enzymatic removal of nucleotides and primers from the PCR mixture by Exo-SAP (Exonuclease I with Shrimp Alkaline Phosphatase) treatment, and direct Sanger sequencing of the purified product (Fig. 1). If the barcode is designed to consist of the four nucleotides equally represented at each position, then in the case of ideally synthesized oligonucleotides, the height of Sanger sequencing chromatogram peaks for each of nucleotides along the barcode should be equal with only slight variations.

Generation of barcoded plasmid libraries
Many approaches have been developed for cloning or assembly of barcoded plasmid libraries, including traditional cloning by restriction enzyme digestion, self-ligation of PCR products, polymerase cycling assembly (PCA), dA/dT ligation, Gibson assembly and others (see Table 1 and references therein). All these approaches consist of several steps: (i) preparation of a plasmid vector and barcoded insert, (ii) ligation of the vector and insert, (iii) highly efficient transformation of bacterial cells, (iv) plasmid propagation in bacteria and its subsequent

М.О. Лебедев, Л.А. Яринич
А.В. Иванкин, А.В. Пиндюрин

Создание штрихкодированных плазмидных
библиотек для анализа эффекта положения гена

2019
23•2

Table 1. Common methods to construct barcoded plasmid libraries

Method	Advantages "+"/Disadvantages "–"	Requirements	Efficiency*	References
Cloning by restriction enzyme digestion and ligation	"+" (1) Simple primer design; (2) Cost effective "–" (1) Restriction site-depending; (2) Time-consuming; (3) Assembly of multiple fragments is complicated/inefficient	Restriction endonucleases, High-fidelity DNA polymerase, Antarctic phosphatase, T4 DNA ligase, DNA purification columns	1.5×10^3	Akhtar et al., 2014
Self-ligation of barcoded PCR product	"+" (1) Simple primer design; (2) Cost effective "–" (1) PCR-induced mutations in the vector sequence; (2) Not applicable for large DNA molecules (>8 kb); (3) Time-consuming; (4) Formation of concatemers	Phosphorylated oligonucleotide, High-fidelity DNA polymerase, T4 DNA polymerase, DpnI restriction enzyme, T4 DNA ligase, DNA purification columns	5.8×10^7	Brueckner et al., 2016; Chen et al., 2017
dA/dT ligation	"+" (1) Simple primer design; (2) Sequence-independent "–" (1) Nondirectional; (2) Expensive; (3) Time-consuming	T-tailed vector, DNA polymerase, T4 DNA ligase, DNA purification columns	20	Seguin-Orlando et al., 2013; Davidsson et al., 2016
Golden Gate assembly	"+" (1) Multiple fragments assembly; (2) Single step; (3) Possible transfer of a DNA fragment library from one vector to another; (4) Fast; (5) Tolerant to repetitive sequences "–" (1) Complicated primer design; (2) Restriction site-depended; (3) Expensive	Nonpalindromic restriction endonucleases, High-fidelity DNA polymerase, Special assembly kit, DNA purification columns	2.7×10^4	Woodruff et al., 2017
Gateway cloning	"+" (1) Single step; (2) Possible transfer of a DNA fragment library from one vector to another; (3) Fast "–" (1) Complicated primer design; (2) Sequence-dependent; (3) Necessity to use specific plasmid backbones; (4) Expensive; (5) Scar att regions	Special assembly kit, High-fidelity DNA polymerase, DNA purification columns	2.0×10^4	Dickel et al., 2014; Lehtonen et al., 2015; Davidsson et al., 2016
Gibson assembly	"+" (1) Single step; (2) Multiple fragments assembly; (3) Sequence-independent; (4) Seamless; (5) Fast "–" (1) Complicated primer design; (2) Self-ligation of linearized vector backbone, if only one restriction site is used for the linearization; (3) Expensive; (4) Not tolerant to repetitive sequences	Special assembly kit, Restriction endonucleases, DNA purification columns	6.9×10^3	Hill et al., 2018

* Efficiency is defined as the number of unique bacterial clones obtained per 1 ng of vector DNA used in an assembly/ligation reaction.

isolation, and (v) assessment of the plasmid library quality. Each approach has its own advantages and disadvantages (see Table 1 and Suppl. Table S2).

For example, cloning by restriction enzyme digestion and ligation is simple and widely used, but it has the following limitations. First, since restriction enzyme recognition sites are short sequences, they are certainly present in a subset of randomly synthesized barcodes. During the cloning, such barcodes will be cut into pieces, thus leading to generation of cloned barcodes of undesired shorter lengths. Accordingly, the number of plasmid molecules with unique barcode sequences of the expected length will be reduced. The longer is the barcode, the higher is the probability of restriction site occurrence in its sequence. For example, a random 18-nt barcode is expected to contain one of two 6-bp restriction sites used for cloning at a frequency of ~1/160. Second, restriction enzymes are not hundred-percent effective, and some of them are extremely ineffective in cutting their recognition sites located close to the ends of linear DNA fragments (in particular, typical PCR products). This results in reduced ligation efficiency, whose compensation requires large amounts of the prepared vector (~1 μg) and insert (~0.2 μg) at the ligation step.

The aim of the present study was to make a set of barcoded plasmid libraries for high-throughput analysis of chromatin position effects (Elgin, Reuter, 2013; Yankulov, 2013) on different types of promoter sequences in *Drosophila* cells. Since our ultimate goal requires integration of barcoded reporter constructs at randomly chosen genomic loci in the studied cells, we employed the piggyBac transposon-based gene delivery system (Handler, Harrell, 1999; Cadinanos, Bradley, 2007), which was successfully used for the same purpose in cultured mouse embryonic stem cells (mESCs) (Akhtar et al., 2013). We selected six different *Drosophila* promoters (here referred to as Promoter1 through 6; their details will be described somewhere else), and cloned each of them upstream of the coding sequence for enhanced green fluorescent protein (eGFP) placed between the piggyBac TRs. The constructs also contained polyadenylation signals immediately before and within the piggyBac 3′-TR. As the control, we used a similar construct without any promoter element (hereafter Promoter0; see Suppl. Fig. S1). Next, two unique 5-bp long library indexes were separately cloned downstream of the eGFP coding sequence into each of the aforementioned construct to generate 14 different indexed plasmids, which were

Table 2. Barcoded plasmid libraries constructed
by Gibson assembly

#	Barcoded plasmid library	Library index	Efficiency*
1	pPB-Promoter0-eGFP-Libr1	TCGCT	2.2×10^4
2	pPB-Promoter0-eGFP-Libr2	CCGAG	2.6×10^3
3	pPB-Promoter1-eGFP-Libr3	TTGAG	3.0×10^3
4	pPB-Promoter1-eGFP-Libr4	GATGG	5.7×10^3
5	pPB-Promoter2-eGFP-Libr5	GGCTT	1.8×10^3
6	pPB-Promoter2-eGFP-Libr6	TGTGT	5.2×10^3
7	pPB-Promoter3-eGFP-Libr7	AGTCA	3.0×10^3
8	pPB-Promoter3-eGFP-Libr8	TGGCC	3.0×10^3
9	pPB-Promoter4-eGFP-Libr9	CTAGT	3.1×10^3
10	pPB-Promoter4-eGFP-Libr10	TTATC	4.7×10^3
11	pPB-Promoter5-eGFP-Libr11	CTGCT	4.0×10^2
12	pPB-Promoter5-eGFP-Libr12	CAATT	2.9×10^3
13	pPB-Promoter6-eGFP-Libr13	GTCAA	8.4×10^3
14	pPB-Promoter6-eGFP-Libr14	GGGAT	3.2×10^3

* Efficiency is defined as the number of unique bacterial clones obtained per
1 ng of vector DNA used in the assembly/ligation reaction.

named pPB-PromoterX-eGFP-LI vectors (for exact sequences
of library indexes, see Table 2). Such indexing of plasmids
allows combinations of (even all) barcoded constructs to be
studied simultaneously in a single TRIP experiment.

At the final step, it was necessary to clone random barcode
sequences immediately downstream of library indexes to ge-
nerate barcoded plasmid libraries. We chose the length of the
barcodes to be 18 bp, which meant that, theoretically, there
might be up to $4^{18} \approx 68$ billion unique barcode sequences. First,
we tried to construct a set of barcoded plasmid libraries using
traditional cloning by restriction enzyme digestion and sub-
sequent ligation following the previously described protocol
(Akhtar et al., 2014) with minor modifications. Specifically,
we designed the barcoded oligonucleotide PB-Barcode1-
SalI-F, which contains the SalI restriction site at its 5′ end and
is suitable to construct all 14 barcoded plasmid libraries due
to the presence of a unique SalI restriction site immediately
downstream of library indexes in pPB-PromoterX-eGFP-LI
vectors. Next, we amplified the DNA fragment containing
piggyBac 3′-TR using this barcoded primer and the reverse
primer 3′-TR-EagI-R containing the EagI restriction site.
We digested barcoded PCR products with SalI and EagI and
cloned them into the pPB-Promoter1-eGFP-LI vector cut with
SalI and NotI. We used the property of the EagI and NotI re-
striction sites, which produce compatible cohesive ends restor-
ing only the EagI site, to remove the original (non-barcoded)
plasmid molecules from the barcoded plasmid library by NotI
digestion. Overall, we found this way of barcode cloning
simple, cheap, and convenient. However, we could not pro-
duce plasmid libraries with complexity of more than 30 clones
per 1 ng of the digested vector due to incomplete digestion of
the barcoded PCR product presumably by the SalI restriction
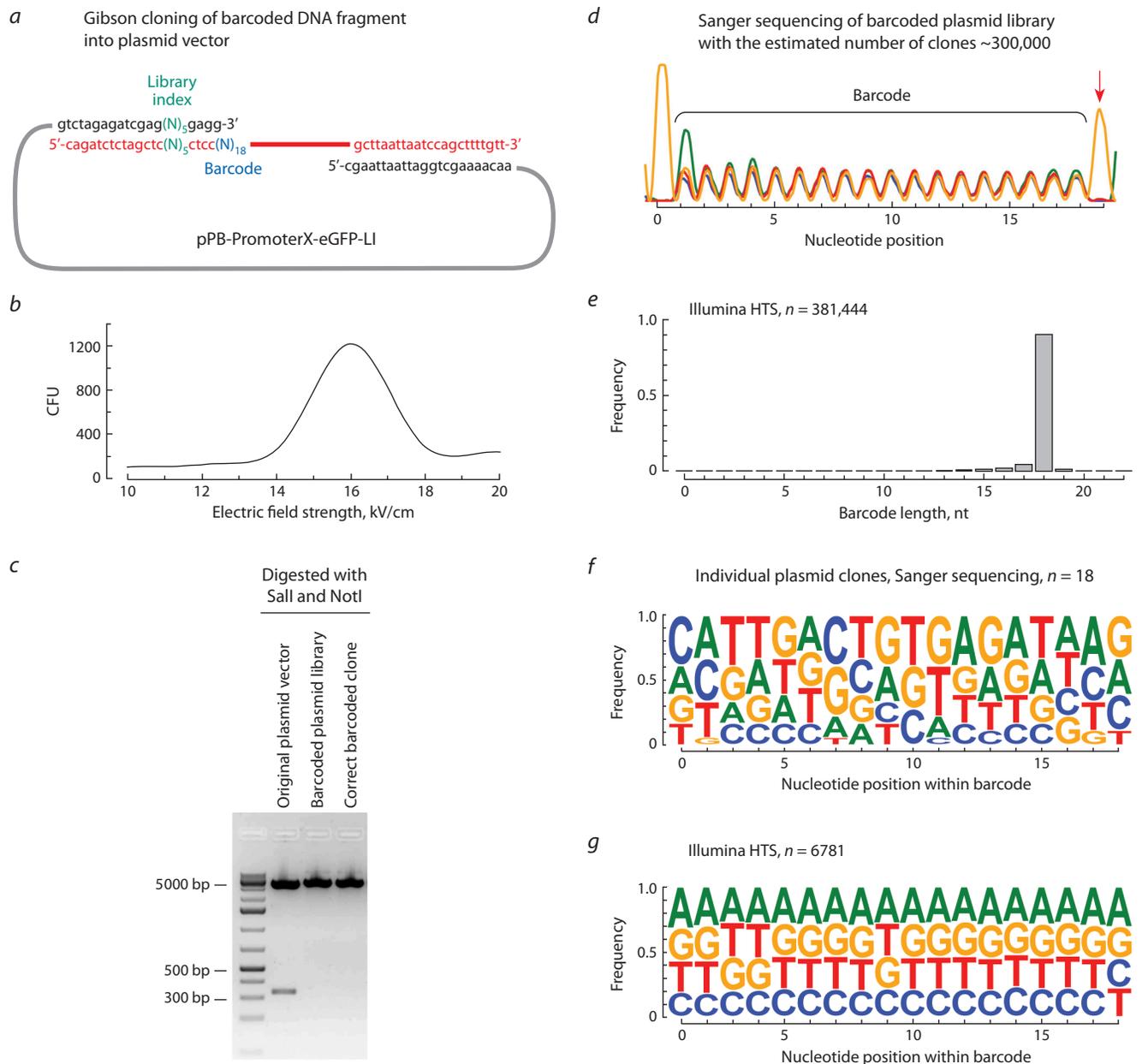enzyme. The protocol consisted of six steps (PCR, restriction

digestion, dephosphorylation of 5′-ends, DNA ligation, treat-
ment of the ligation mixture by restriction endonuclease, and
electrotransformation of bacteria) with DNA purification after
each step. As a result, using about 1000 ng of the digested
vector and 10 individual electrotransformations, we obtained
a plasmid library consisting of ~300,000 clones. Thus, the
conventional restriction-ligation method appeared to be too
laborious especially for the generation of large numbers of
different barcoded plasmid libraries, while its cloning accu-
racy, calculated as the percentage of the number of colonies
obtained for the "vector with barcoded insert" ligation with
respect to the number of colonies obtained by the control
"vector only" ligation, was satisfactory (about 0.6 %).

To optimize the construction of barcoded plasmid librar-
ies, we decided to use the Gibson assembly approach, which
employs three commonly used enzymes: (i) 5′ exonuclease
that digests the ends of dsDNA, exposing the ssDNA, (ii) high-
fidelity DNA polymerase, and (iii) T4 DNA ligase (Gibson
et al., 2009). The assembly fragments were designed so that
they contained at their ends overlapping sequences of 23 bp
in length, which were expected to anneal at the temperature
60–65 °C (Fig. 2, a). We amplified the barcoded DNA frag-
ments containing piggyBac 3′-TR using PB-Barcode-PI-X-
Gibson-F/PB-Gibson-R1 primers. The primer PB-Barcode-PI-
X-Gibson-F contains a random 18-bp barcode but lacks the
SalI restriction site in order to allow elimination of the original
plasmid molecules from the barcoded plasmid library by SalI
digestion. As vector fragments, we used NotI/SalI-linearized
pPB-PromoterX-eGFP-LI plasmids. Next, to enhance the ef-
ficiency of the Gibson assembly reaction, we tested a vector to
insert ratios of 1:3, 1:5, and 1:10 and found that the maximum
number of individual bacterial colonies was obtained with the
1:3 ratio (570,000 clones versus 470,000 and 380,000 clones
for the 1:5 and 1:10 ratios, respectively). Thus, an excess of
the insert reduced the Gibson assembly efficiency and conse-
quently yielded fewer clones possibly due to competition of
DNA molecules for enzymes. Furthermore, we revealed that
electroporation of *E. coli* cells at 1.6 kV increased the number
of bacterial colonies more than fourfold and greatly reduced
arcing compared to the predefined 1.8 kV in the Gene Pulser
electroporator (Bio-Rad) (see Fig. 2, b).

A total of 14 barcoded plasmid libraries were generated by
the Gibson assembly approach (see Table 2), which consists of
four steps (PCR, vector linearization, 1-h assembly reaction,
and electrotransformation of bacteria) with DNA purification
after each step. As a result, using about 200 ng of the linearized
vector and 2 individual electrotransformations, we obtained
plasmid libraries consisting of ~300,000–2,000,000 clones.

The complexity and quality of barcoded plasmid libraries

To generate populations of cells carrying thousands of uni-
quely barcoded reporter constructs integrated at different ge-
nomic loci, it is desirable to obtain appropriate plasmid lib-
raries with the highest possible complexity. This would mi-
nimize the chances of integration of the identically barcoded
reporter constructs in different genomic locations. Note that
the presence of the same barcode at different unique genomic
sites (we do not consider here integration of a reporter within
repetitive elements) makes this barcode useless for the down-
stream analysis, since its characteristics (e. g., the expression

М.О. Лебедев, Л.А. Яринич
А.В. Иванкин, А.В. Пиндюрин

Создание штрихкодированных плазмидных
библиотек для анализа эффекта положения гена

2019
23•2

a Gibson cloning of barcoded DNA fragment
into plasmid vector

b

c

d Sanger sequencing of barcoded plasmid library
with the estimated number of clones ~300,000

e Illumina HTS, *n* = 381,444

f Individual plasmid clones, Sanger sequencing, *n* = 18

g Illumina HTS, *n* = 6781

Fig. 2. Generation and quality control of barcoded plasmid libraries.

(*a*) Schematic presentation of the Gibson assembly of barcoded plasmid libraries. (*b*) Dependence of the number of bacterial colonies obtained on the electric field strength used in the electrotransformation. (*c–g*) Characterization of barcode diversity: (*c*) Agarose gel electrophoresis showing a negligible amount, if any, of the original vector contamination in the barcoded plasmid library. The original plasmid vector and correct barcoded clone serve as positive and negative controls of the presence of the original vector, respectively. (*d*) Representative Sanger sequencing chromatogram of a barcoded plasmid library generated by Gibson assembly approach. Note the presence of only guanine (orange line), but not thymine (red line) expected in the case of extensive original vector contamination at the position indicated by the red arrow. (*e*) Distribution of the barcode length revealed by HTS; *n*, number of reads analyzed. (*f*, *g*) Frequency distributions of the four nucleotides per position within the barcode revealed by (*f*) Sanger sequencing of individual plasmids and (*g*) by HTS. Weblogo 3, a web based application (http://weblogo.threeplusone.com/create.cgi), was used for visualization of the frequencies. In panel (*f*), *n* is the number of clones, and in panel (*g*), unique barcodes analyzed.

level) cannot be unambiguously associated with a genomic locus. Also, such ambiguous barcodes reduce the number of informative HTS reads used to identify the genomic locations of reporter constructs and their expression levels. In addition, barcode sequences should be easily distinguishable from each other to avoid mistakes in the interpretation of the results. Such diversity among barcode sequences can be achieved by equal representation of each nucleotide at each position along

a barcode in the plasmid library (not to mention the effects of barcode length and the number of different nucleotides constituting the barcode). Moreover, it is crucial to avoid substantial levels of contamination of barcoded plasmid libraries with the original (non-barcoded) vector molecules, as the latter would also reduce the number of informative HTS reads.

To assess the quality of the barcoded plasmid libraries obtained in the present study, we performed a set of experiments.

First, to evaluate the level of contamination by the original vector, we digested one of the barcoded plasmid libraries with the NotI and SalI restriction enzymes. Recognition sites for both these enzymes were present in the original vectors but not in plasmids with properly cloned barcodes. The analysis did not reveal any substantial amount of the original vector molecules within the library (see Fig. 2, *c*). Consistently, the transformation of bacterial cells with the control "vector only" samples also showed that the level of the original vector contamination in the libraries was no higher than 0.5 %. Second, we analyzed plasmids isolated from 20 randomly selected bacterial colonies of one barcoded plasmid library by Sanger sequencing. It showed that all the 20 plasmids had the expected structure, although one plasmid carried a 17 bp long barcode and another colony appeared to contain two plasmids with different barcodes. The barcodes of the remaining 18 plasmids were of the expected length (18 bp) and demonstrated satisfactory representation of different nucleotides at each position (see Fig. 2, *f*). Third, we subjected all barcoded plasmid libraries to Sanger sequencing to assess the diversity of the barcode sequences and possible non-barcoded plasmid contamination. This analysis demonstrated almost equal representation of all four nucleotides at the most positions within the barcode and confirmed the absence of contamination by the original vector (see Fig. 2, *d*). Fourth, we performed Illumina HTS for more precise characterization of barcoded plasmid libraries. For this purpose, we mixed all prepared barcoded plasmid libraries and reduced the total number of clones by restrictive transformation of *E. coli* cells. Analysis of the HTS reads showed that (i) the length of vast majority of barcodes was 18 bp (see Fig. 2, *e*), (ii) the representation of different nucleotides along the barcodes was almost perfect (see Fig. 2, *g*), and (iii) the contamination of the barcoded plasmid libraries by the original vectors was about 0.4 % on the average.

Finally, as described above, we quantified the complexity of each barcoded plasmid library by plating a small aliquot of the transformation mixture in duplicate on LB-ampicillin plates and counting the resulting number of bacterial colonies that supposedly corresponded to the number of individual plasmid molecules. Unfortunately, we noticed that this method was not always accurate due to its dependence on bacterial concentration in liquid culture, the quality of LB-ampicillin plates, and plating techniques. Thus, alternative simple methods for the estimation of the complexity of barcoded plasmids libraries need to be developed.

Conclusions

Generation of barcoded plasmid libraries is challenging. The main difficulties are associated with the diversity of barcode sequences and their efficient cloning in a plasmid vector, including obtaining a sufficient number of individual bacterial colonies. We describe a low-cost, fast, and simple technique to check the quality of barcoded oligonucleotides by broadly used Sanger sequencing. We also present an improved method to create barcoded plasmid libraries in a single step with the commercially available Gibson assembly mixture. It is worth mentioning that the protocols developed in this study are suitable for generation of barcoded plasmid libraries for a wide range of applications, including but not limited to functional analysis of DNA regulatory elements in different species.

References

Akhtar W., de Jong J., Pindyurin A.V., Pagie L., Meuleman W., de Ridder J., Berns A., Wessels L.F., van Lohuizen M., van Steensel B. Chromatin position effects assayed by thousands of reporters integrated in parallel. Cell. 2013;154:914-927. DOI 10.1016/j.cell. 2013.07.018.

Akhtar W., Pindyurin A.V., de Jong J., Pagie L., Ten Hoeve J., Berns A., Wessels L.F., van Steensel B., van Lohuizen M. Using TRIP for genome-wide position effect analysis in cultured cells. Nat. Protoc. 2014;9:1255-1281. DOI 10.1038/nprot.2014.072.

Babenko V.N., Makunin I.V., Brusentsova I.V., Belyaeva E.S., Maksimov D.A., Belyakin S.N., Maroy P., Vasil'eva L.A., Zhimulev I.F. Paucity and preferential suppression of transgenes in late replication domains of the *D. melanogaster* genome. BMC Genomics. 2010;11: 318. DOI 10.1186/1471-2164-11-318.

Babu A., Verma R.S. Chromosome structure: euchromatin and heterochromatin. Int. Rev. Cytol. 1987;108:1-60.

Blundell J.R., Levy S.F. Beyond genome sequencing: lineage tracking with barcodes to study the dynamics of evolution, infection, and cancer. Genomics. 2014;104:417-430. DOI 10.1038/nature14279.

Brueckner L., van Arensbergen J., Akhtar W., Pagie L., van Steensel B. High-throughput assessment of context-dependent effects of chromatin proteins. Epigenetics Chromatin. 2016;9:43. DOI 10.1186/ s13072-016-0096-y.

Cadinanos J., Bradley A. Generation of an inducible and optimized piggyBac transposon system. Nucleic Acids Res. 2007;35:e87. DOI 10.1093/nar/gkm446.

Chen H.C., Martinez J.P., Zorita E., Meyerhans A., Filion G.J. Position effects influence HIV latency reversal. Nat. Struct. Mol. Biol. 2017; 24:47-54. DOI 10.1038/nsmb.3328.

Chen M., Licon K., Otsuka R., Pillus L., Ideker T. Decoupling epigenetic and genetic effects through systematic analysis of gene position. Cell Rep. 2013;3:128-137. DOI 10.1016/j.celrep.2012.12.003.

Davidsson M., Diaz-Fernandez P., Schwich O.D., Torroba M., Wang G., Bjorklund T. A novel process of viral vector barcoding and library preparation enables high-diversity library generation and recombination-free paired-end sequencing. Sci. Rep. 2016;6:37563. DOI 10.1038/srep37563.

Dickel D.E., Zhu Y., Nord A.S., Wylie J.N., Akiyama J.A., Afzal V., Plajzer-Frick I., Kirkpatrick A., Gottgens B., Bruneau B.G., Visel A., Pennacchio L.A. Function-based identification of mammalian enhancers using site-specific integration. Nat. Methods. 2014;11:566-571. DOI 10.1038/nmeth.2886.

Elgin S.C., Reuter G. Position-effect variegation, heterochromatin formation, and gene silencing in *Drosophila*. Cold Spring Harb. Perspect. Biol. 2013;5:a017780. DOI 10.1101/cshperspect.a017780.

Ernst J., Kheradpour P., Mikkelsen T.S., Shoresh N., Ward L.D., Epstein C.B., Zhang X., Wang L., Issner R., Coyne M., Ku M., Durham T., Kellis M., Bernstein B.E. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 2011;473:43-49. DOI 10.1038/nature09906.

Filion G.J., van Bemmel J.G., Braunschweig U., Talhout W., Kind J., Ward L.D., Brugman W., de Castro I.J., Kerkhoven R.M., Bussemaker H.J., van Steensel B. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. Cell. 2010; 143:212-224. DOI 10.1016/j.cell.2010.09.009.

Gibson D.G., Young L., Chuang R.Y., Venter J.C., Hutchison C.A. 3rd, Smith H.O. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods. 2009;6:343-345. DOI 10.1038/ nmeth.1318.

Gierman H.J., Indemans M.H., Koster J., Goetze S., Seppen J., Geerts D., van Driel R., Versteeg R. Domain-wide regulation of gene expression in the human genome. Genome Res. 2007;17:1286-1295. DOI 10.1101/gr.6276007.

Grewal S.I., Moazed D. Heterochromatin and epigenetic control of gene expression. Science. 2003;301:798-802. DOI 10.1126/science. 1086887.

М.О. Лебедев, Л.А. Яринич
А.В. Иванкин, А.В. Пиндюрин

Создание штрихкодированных плазмидных
библиотек для анализа эффекта положения гена

2019
23•2

Handler A.M., Harrell R.A. 2nd. Germline transformation of *Drosophila melanogaster* with the piggyBac transposon vector. Insect Mol. Biol. 1999;8:449-457. DOI 10.1046/j.1365-2583.1999.00139.x.

Hill A.J., McFaline-Figueroa J.L., Starita L.M., Gasperini M.J., Matreyek K.A., Packer J., Jackson D., Shendure J., Trapnell C. On the design of CRISPR-based single-cell molecular screens. Nat. Methods. 2018;15:271-274. DOI 10.1038/nmeth.4604.

Huisinga K.L., Brower-Toland B., Elgin S.C. The contradictory definitions of heterochromatin: transcription and silencing. Chromosoma. 2006;115:110-122. DOI 10.1007/s00412-006-0052-x.

Kebschull J.M., Zador A.M. Cellular barcoding: lineage tracing, screening and beyond. Nat. Methods. 2018;15:871-879. DOI 10.1038/s41592-018-0185-x.

Kharchenko P.V., Alekseyenko A.A., Schwartz Y.B., Minoda A., Riddle N.C., Ernst J., Sabo P.J., Larschan E., Gorchakov A.A., Gu T., Linder-Basso D., Plachetka A., Shanower G., Tolstorukov M.Y., Luquette L.J., Xi R., Jung Y.L., Park R.W., Bishop E.P., Canfield T.K., Sandstrom R., Thurman R.E., MacAlpine D.M., Stamatoyannopoulos J.A., Kellis M., Elgin S.C., Kuroda M.I., Pirrotta V., Karpen G.H., Park P.J. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. Nature. 2011;471:480-485. DOI 10.1038/nature09725.

Kinde I., Wu J., Papadopoulos N., Kinzler K.W., Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. Proc. Natl. Acad. Sci. USA. 2011;108:9530-9535. DOI 10.1073/pnas.1105422108.

Lehtonen S.I., Taskinen B., Ojala E., Kukkurainen S., Rahikainen R., Riihimaki T.A., Laitinen O.H., Kulomaa M.S., Hytonen V.P. Efficient preparation of shuffled DNA libraries through recombination (Gateway) cloning. Protein Eng. Des. Sel. 2015;28:23-28. DOI 10.1093/protein/gzu050.

Melnikov A., Murugan A., Zhang X., Tesileanu T., Wang L., Rogov P., Feizi S., Gnirke A., Callan C.G., Jr., Kinney J.B., Kellis M., Lander E.S., Mikkelsen T.S. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. Nat. Biotechnol. 2012;30:271-277. DOI 10.1038/nbt.2137.

Morrison S.L. Transformation of *E. coli* by electroporation. Curr. Protoc. Immunol. 2001;21(1):A.3N.1-A.3N.4. DOI 10.1002/0471142735.ima03ns21.

Patwardhan R.P., Hiatt J.B., Witten D.M., Kim M.J., Smith R.P., May D., Lee C., Andrie J.M., Lee S.I., Cooper G.M., Ahituv N., Pennacchio L.A., Shendure J. Massively parallel functional dissection of mammalian enhancers *in vivo*. Nat. Biotechnol. 2012;30:265-270. DOI 10.1038/nbt.2136.

Patwardhan R.P., Lee C., Litvin O., Young D.L., Pe'er D., Shendure J. High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. Nat. Biotechnol. 2009;27:1173-1175. DOI 10.1038/nbt.1589.

Quail M.A., Smith M., Jackson D., Leonard S., Skelly T., Swerdlow H.P., Gu Y., Ellis P. SASI-Seq: sample assurance Spike-Ins, and highly differentiating 384 barcoding for Illumina sequencing. BMC Genomics. 2014;15:110. DOI 10.1186/1471-2164-15-110.

Riddle N.C., Minoda A., Kharchenko P.V., Alekseyenko A.A., Schwartz Y.B., Tolstorukov M.Y., Gorchakov A.A., Jaffe J.D., Kennedy C., Linder-Basso D., Peach S.E., Shanower G., Zheng H., Kuroda M.I., Pirrotta V., Park P.J., Elgin S.C., Karpen G.H. Plasticity in patterns of histone modifications and chromosomal proteins in *Drosophila* heterochromatin. Genome Res. 2011;21:147-163. DOI 10.1101/gr.110098.110.

Ruf S., Symmons O., Uslu V.V., Dolle D., Hot C., Ettwiller L., Spitz F. Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. Nat. Genet. 2011;43:379-386. DOI 10.1038/ng.790.

Seguin-Orlando A., Schubert M., Clary J., Stagegaard J., Alberdi M.T., Prado J.L., Prieto A., Willerslev E., Orlando L. Ligation bias in Illumina next-generation DNA libraries: implications for sequencing ancient genomes. PLoS One. 2013;8:e78575. DOI 10.1371/journal.pone.0078575.

Vvedenskaya I.O., Zhang Y., Goldman S.R., Valenti A., Visone V., Taylor D.M., Ebright R.H., Nickels B.E. Massively Systematic Transcript End Readout, "MASTER": transcription start site selection, transcriptional slippage, and transcript yields. Mol. Cell. 2015;60:953-965. DOI 10.1016/j.molcel.2015.10.029.

Wong A.S., Choi G.C., Cui C.H., Pregernig G., Milani P., Adam M., Perli S.D., Kazer S.W., Gaillard A., Hermann M., Shalek A.K., Fraenkel E., Lu T.K. Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. Proc. Natl. Acad. Sci. USA. 2016;113:2544-2549. DOI 10.1073/pnas.1517883113.

Woodruff L.B.A., Gorochowski T.E., Roehner N., Mikkelsen T.S., Densmore D., Gordon D.B., Nicol R., Voigt C.A. Registry in a tube: multiplexed pools of retrievable parts for genetic design space exploration. Nucleic Acids Res. 2017;45:1553-1565. DOI 10.1093/nar/gkw1226.

Yankulov K. Dynamics and stability: epigenetic conversions in position effect variegation. Biochem. Cell Biol. 2013;91:6-13. DOI 10.1139/bcb-2012-0048.

ORCID ID

M.O. Lebedev orcid.org/0000-0001-7461-0977
L.A. Yarinich orcid.org/0000-0003-0469-0371
A.V. Ivankin orcid.org/0000-0003-3309-6363
A.V. Pindyurin orcid.org/0000-0001-6959-0641