

Philipps



Universität

Marburg

Charakterisierung der DNA-Replikation
in genetisch modifizierten *Escherichia coli* Stämmen
durch neue Methoden der Synthetischen Biologie

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich der Biologie

der Philipps-Universität Marburg

vorgelegt von

M.Sc. Biology

Sarah Milbredt

aus Bielefeld

Marburg, Oktober 2016

Originaldokument gespeichert auf dem Publikationsserver der
Philipps-Universität Marburg
<http://archiv.ub.uni-marburg.de>



Dieses Werk bzw. Inhalt steht unter einer
Creative Commons
Namensnennung
Keine kommerzielle Nutzung
Weitergabe unter gleichen Bedingungen
3.0 Deutschland Lizenz.

Die vollständige Lizenz finden Sie unter:
<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>

Die Untersuchungen zur vorliegenden Arbeit wurden von Mai 2012 bis Oktober 2016 am LOEWE-Zentrum für Synthetische Mikrobiologie in Marburg unter Leitung von Prof. Dr. T. Waldminghaus durchgeführt.

Vom Fachbereich Biologie der Philipps-Universität Marburg
als Dissertation angenommen am:

Erstgutachter: Prof. Dr. T. Waldminghaus

Zweitgutachter: Prof. Dr. H.-U. Mösch

Tag der mündlichen Prüfung: 16.12.2016

Die in dieser Dissertation vorgestellten Ergebnisse sind in den folgenden Originalpublikationen veröffentlicht oder werden zur Veröffentlichung vorbereitet:

Milbredt, S., Farmani, N., Sobetzko, P. and Waldminghaus, T., 2016. **DNA replication in engineered *Escherichia coli* genomes with extra replication origins.** ACS Synth Biol. 5, 1167-1176.

Schindler, D., Milbredt, S., Sperlea, T. and Waldminghaus, T., 2016. **Design and assembly of DNA sequence libraries for chromosomal insertion in bacteria based on a set of modified MoClo vectors.** ACS Synth Biol. DOI: 10.1021/acssynbio.6b00089. [Epub ahead of print].

Milbredt, S. and Waldminghaus, T., **BIFCROS: A low background Fluorescence Repressor Operator System.** (prepared for submission).

Inhaltsverzeichnis

Abkürzungsverzeichnis.....	1
1. Einleitung.....	2
1.1 Synthetische Biologie	2
1.2 DNA-Replikation in <i>Escherichia coli</i>	3
1.2.1 Der Replikationsursprung von <i>Escherichia coli</i> : <i>oriC</i>	3
1.3 Eignen sich ektopische <i>oriC</i> als Grundlage für sekundäre Chromosomen in <i>Escherichia coli</i> ? .	6
1.4 Der Modellorganismus <i>Vibrio cholerae</i> : Ein Bakterium mit zwei Chromosomen	6
1.5 <i>orill</i> : Ein prädestinierter Replikationsursprung für sekundäre Chromosomen in <i>Escherichia coli</i>	8
1.6 Die Kopienzahl eines genetischen Lokus auf Einzelzellniveau als Reporter für die DNA-Replikation	9
1.7 Moderne Klonierungssysteme	13
1.7.1 Homologie-basierte Methoden zur DNA-Assemblierung	13
1.7.2 Restriktionsenzym-basierte Methoden zur DNA-Assemblierung.....	14
1.7.3 Modulares Klonieren mit MoClo	15
2. Ziele dieser Arbeit	16
3. Ergebnisse.....	17
3.1 DNA replication in engineered <i>Escherichia coli</i> genomes with extra replication origins	18
3.2 Design and assembly of DNA sequence libraries for chromosomal insertion in bacteria based on a set of modified MoClo vectors	21
3.3 BiFCROS: A low background Fluorescence Repressor Operator System.....	23
4. Diskussion.....	26
4.1 Fortschritt auf Einzelzellniveau: Entwicklung von Methoden zur Kopienzahlbestimmung	26
4.1.1 Entwicklung eines fluoreszenzbasierten Reportersystems für die Analyse von Genloki .	27
4.1.2 Charakterisierung der DNA-Replikation mit Hilfe von Click-Chemistry	29
4.2 Ein neuer Ansatz um die Kopienzahl von chromosomalen Elementen anhand von kurzen, individuellen Sequenzen zu bestimmen	30
4.2.1 Nukleotide als Testsystem für Linker-counting.....	34
4.3 Die „Eukaryotisierung“ von <i>Escherichia coli</i> – ein Gedankenexperiment	35
4.3.1 Welche “Chromosome maintenance systems” werden auf linearen Chromosomen benötigt?	39
5. Zusammenfassung	43
5. Summary	45
6. Literaturverzeichnis.....	46
Erklärung der selbstständigen Erarbeitung der Dissertation.....	58

Abkürzungsverzeichnis

bp	Basenpaare
DNA	Desoxyribonukleinsäure
et. al	und andere
kb	Kilobasen
Mbp	Mega-Basenpaare
PCR	Polymerasekettenreaktion (polymerase chain reaction)
RNA	Ribonukleinsäure
tRNA	Transfer-RNA

Allgemein gebräuchliche Abkürzungen und Maßeinheiten sind nicht gesondert aufgeführt.

1. Einleitung

1.1 Synthetische Biologie

Die Synthetische Biologie befasst sich mit dem Ziel, biologische Systeme mit neuen Eigenschaften zu konstruieren sowie komplexe biologische Systeme zu vereinfachen. Um diese Ziele zu erreichen, können mit den Ansätzen der Synthetischen Biologie Proteine, Zellen und Organismen gezielt entworfen werden. Neue Eigenschaften beschreiben unter anderem die Produktion biologisch relevanter Wirkstoffe durch Modulierung von Stoffwechselwegen, wie die heterologe Herstellung von Artemisinin - ein Antimalaria Medikament - in der Hefe oder die Produktion von Biodiesel und Benzin in *Escherichia coli* (Choi and Lee, 2013; Paddon et al., 2013; Steen et al., 2010). Neue oder verbesserte Aktivitäten von Proteinen können durch die Erweiterung des genetischen Codes erzielt werden. Seltene oder unbenutzte Codons können für den Einbau von nicht-natürlichen Aminosäuren genutzt werden, um Designerproteine mit neuen Eigenschaften zu generieren (Liu and Schultz, 2010). Mit Hilfe dieses Ansatzes konnten bereits medizinische Therapeutika wie Antikörper und Antikoagulanzen hergestellt werden (Grunewald et al., 2009; Liu and Schultz, 2006). Ein großer Durchbruch gelang der Synthetischen Biologie mit der Herstellung synthetischer Chromosomen. Venter *et al.* generierten mit Hilfe neuer DNA-Assemblierungsmethoden das synthetische Genom von *Mycoplasma mycoides* und transplantierten es in *Mycoplasma capricolum* (Gibson et al., 2008a; Gibson et al., 2008b; Gibson et al., 2010). So ergaben sich lebensfähige Zellen, die durch das synthetische Chromosom kontrolliert wurden (Gibson et al., 2010). Die Herstellung eines synthetischen Chromosoms der Hefe konnte ebenfalls erfolgreich bewerkstelligt werden (Annaluru et al., 2014). Synthetische Chromosomen bieten den Vorteil, dass sie frei entworfen und nach dem „bottom-up“-Prinzip erweitert werden können. Alternativ besteht die Möglichkeit native Chromosomen gezielt zu verändern. Genom-Rekonstruktionen können den Umbau von zirkulären Chromosomen in lineare beinhalten oder aber das Unterteilen des Chromosoms in kleinere Untereinheiten (Cui et al., 2007; Itaya and Tanaka, 1997). So bieten kleine Einheiten den Vorteil, dass sie aufgrund ihrer Größe besser genetisch manipulierbar sind (Schindler and Waldminghaus, 2015). Die Reduktion eines Genoms kann ebenfalls als Ansatz verwendet werden, um komplexe Systeme zu vereinfachen (Hutchison et al., 2016). Allen Chromosomen, synthetisch oder nativ, ist gemein, dass für ihre Vererbung die Duplikation essentiell ist.

1.2 DNA-Replikation in *Escherichia coli*

Die DNA-Replikation bezeichnet den essentiellen Prozess der Duplikation der genetischen Information einer Zelle vor der Zellteilung. Somit kann an jede Tochterzelle ein Chromosom vererbt werden. Findet dieser Prozess nicht statt, erhält die Tochterzelle kein Chromosom, was zum Absterben dieser Zelle führt. Es ist wichtig, dass das Chromosom reguliert repliziert wird, also nur einmal pro Zellzyklus (Boye et al., 2000). *E. coli*, ein gramnegatives γ -Proteobakterium, enthält die genetische Information auf einem einzigen, zirkulär-organisierten Chromosom mit einer Größe von 4.6 Mbp (Blattner et al., 1997). Die DNA-Replikation startet an einem definierten Punkt auf dem Chromosom, dem sogenannten Replikationsursprung, *oriC* (*origin of replication*). Vermittelt wird die DNA-Replikation durch das in Bakterien hochkonservierte Initiatorprotein DnaA, welches an spezifische Sequenzmotive innerhalb des *oriC*, den DnaA-Boxen, bindet (Fuller et al., 1984; Messer, 2002; Ozaki and Katayama, 2009). Das beschriebene Protein führt zum Aufbrechen der Wasserstoffbrückenbindungen zwischen den Basenpaaren in der AT-reichen Region innerhalb des *oriC* und somit zum Aufschmelzen der DNA. DnaA sorgt auch für das Rekrutieren replikationsspezifischer Proteine (Kaguni, 1997; Langer et al., 1996; Skarstad and Boye, 1994). Die eigentliche DNA-Replikation wird vom Replisom, einem Multiproteinkomplex, an einer Y-geformten Replikationsgabel ausgeführt. Die Schlüsselenzyme dieses Komplexes umfassen bei *E. coli* die Helikase DnaB, die Primase DnaG und das Polymerase III Holoenzym, Pol III (McHenry, 2011; O'Donnell, 2006). Von *oriC* aus verläuft die DNA-Replikation bidirektional (Bird et al., 1972). Die Replisomen laufen dabei in entgegengesetzte Richtungen und treffen erst wieder in der Terminations-Domäne (*ter*) aufeinander, nachdem sie jeweils die Hälfte des Chromosoms repliziert haben (Hill et al., 1987). Dort wird die DNA-Replikation beendet, wobei die Replisomen durch den Tus-*ter*-Komplex in ihrem Fortlaufen gehindert werden (Duggin et al., 2008). Als Folge disassemblieren die Replisomen und ermöglichen so die Segregation der beiden Chromosomen auf die Tochterzellen. Welche Charakteristika eines Replikationsursprungs für eine erfolgreiche Initiation der DNA-Replikation benötigt werden beschreibt Kapitel 1.2.1.

1.2.1 Der Replikationsursprung von *Escherichia coli*: *oriC*

Seit ungefähr 39 Jahren ist bekannt, dass sich der Replikationsursprung von *E. coli*, *oriC*, bei 84.3 min auf der Genomkarte befindet (Marsh and Worcel, 1977). Um die Sequenz von *oriC* genauer einzugrenzen, wurden sogenannte *oriC*-basierte Minichromosomen verwendet (Hiraga, 1976; Yasuda and Hirota, 1977). Dabei handelt es sich um Replikons mit *oriC* als Replikationsursprung, die autonom in der Zelle replizieren (Lobner-Olesen et al., 1987; Messer et al., 1978). Sie können auch als sekundäre Chromosomen bezeichnet werden, weil sie einen chromosomalen Replikationsursprung enthalten und in *E. coli* zusätzlich zum primären Chromosom eingebracht werden (Schindler and Waldminghaus, 2015). Durch das Einfügen von variablen *oriC*-Sequenzen auf dem Minichromosom konnte die

minimale Sequenz, die ein Replikon zur erfolgreichen DNA-Replikation *in vivo* befähigt, zunächst auf 422 bp und später auf 232-246 bp eingegrenzt werden (Sugimoto et al., 1979; Tabata et al., 1983; von Meyenburg et al., 1979). Eine Übersicht der *oriC*-Sequenz wird in Abbildung 1 dargestellt. In Abschnitt 1.2 wurde bereits deutlich, dass der Replikationsursprung von *E. coli* unter anderem eine AT-reiche Region und Bindestellen für das Initiatorprotein DnaA aufweisen muss. DnaA bindet eine asymmetrische, 9 bp bzw. 6 bp lange Konsensussequenz, die als DnaA-Box bezeichnet wird (Fuller et al., 1984; Schaper and Messer, 1995). DnaA bindet ATP sowie ADP (Sekimizu et al., 1987). Aufgrund der Bindeaffinität wird bei den DnaA-Boxen zwischen starken und schwachen Boxen differenziert: Starke Boxen werden von DnaA^{ATP} und DnaA^{ADP} gleichermaßen gebunden (Fuller et al., 1984). Schwache Boxen hingegen werden nur von DnaA^{ATP} erkannt (Rozgaja et al., 2011). In der Sequenz von *oriC* sind insgesamt zehn DnaA-Boxen (R1 - R10) zu finden (Skarstad and Katayama, 2013). Nur drei davon, R1, R2 und R4, sind starke Boxen (Kawakami and Katayama, 2010; McGarry et al., 2004; Rozgaja et al., 2011). Zusätzlich enthält *oriC* Bindestellen für DNA strukturverändernde Proteine (histonähnliche Proteine), wie dem IHF (Integration Host Factor) und dem FIS Protein (Factor for Inversion Stimulation) (Grimwade et al., 2000; Roth et al., 1994). Die Bindung beider Proteine führt zu einer Konformationsänderung der DNA, wodurch das Aufschmelzen der DNA und einhergehend die Assemblierung des Replisoms begünstigt wird (Gille et al., 1991; Polaczek, 1990). Die *oriC*-Sequenz wird von zwei hochkonservierten Genen, *gidA* und *mioC*, flankiert. *gidA*, auch bekannt als *mnmG*, kodiert für ein Protein, welches in der Modifikation von tRNAs involviert ist (Bregon et al., 2001; Moukadiri et al., 2009). Der *gidA*-Promotor befindet sich neben der AT-reichen Region, sodass die Transkription von *gidA* nicht innerhalb der *oriC*-Sequenz stattfindet (siehe Abb. 1) (Kolling et al., 1988). *mioC* kodiert für ein Protein, dessen biologische Funktion noch nicht genau identifiziert werden konnte (Lies et al., 2015).

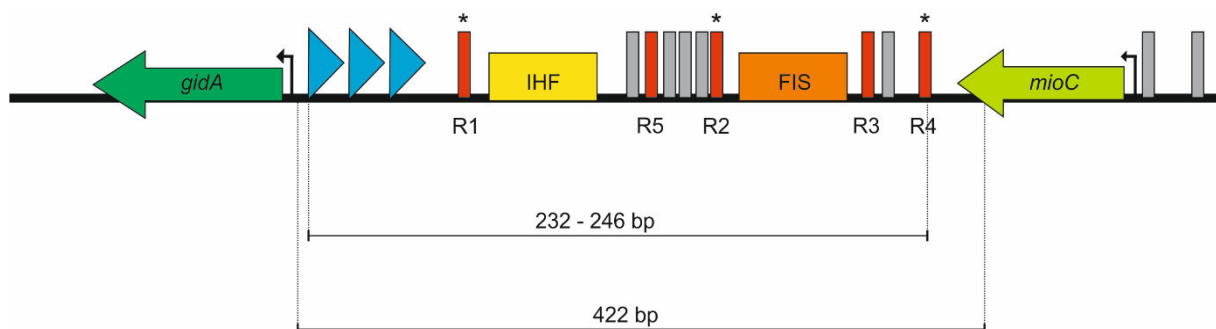


Abbildung 1: Der Replikationsursprung *oriC* und seine flankierenden Bereiche. Die DNA-Elemente der *oriC*-Region sind wie folgt dargestellt: Blaue Dreiecke, 13-mer in AT-reicher Region; rote Rechtecke, 9 bp DnaA-Boxen; graue Rechtecke, 6 bp DnaA-Boxen. Starke DnaA-Boxen sind mit einem Stern gekennzeichnet. Bindestellen für IHF und FIS sind gelb bzw. orange markiert. Flankierende Gene sind in grün dargestellt. Die Längen der verschiedenen *oriC*-Fragmente sind indiziert (Sugimoto et al., 1979; Tabata et al., 1983; von Meyenburg et al., 1979).

Jedoch konnten zwei DnaA-Boxen innerhalb der Promotorregion von *mioC* gefunden werden (Hansen et al., 2006). Zudem endet die Transkription von *mioC* innerhalb der *oriC*-Sequenz (Hansen et al., 2007; Messer and Weigel, 1997). *oriC*-basierte Minichromosomen werden seit längerer Zeit verwendet, um die Initiation der DNA-Replikation genauer zu analysieren (Leonard and Helmstetter, 1986; Lobner-Olesen et al., 1987). Die Attraktivität dieses Ansatzes kann darauf zurückgeführt werden, dass ein *oriC*-basiertes Minichromosom ähnliche Eigenschaften für die DNA-Replikationsinitiation mit dem primären Chromosom teilt; (i) die DNA-Replikation verläuft in beiden Fällen bidirektional ab (Messer and Weigel, 1996) und (ii) beide benötigen die Transkription und eine *de novo* Proteinsynthese (Messer and Weigel, 1996). Dennoch konnten signifikante Unterschiede zwischen der Regulation der Initiation der DNA-Replikation auf einem Minichromosom und dem nativen Chromosom gefunden werden (Asai et al., 1998). Die Aktivität von *oriC* auf Minichromosomen kann durch benachbarte Gene auf dem jeweiligen Replikon oder durch die Länge der verwendeten *oriC*-Sequenz beeinflusst werden (Asai et al., 1998). Wird die DnaA-Box R4 (Abb. 1) aus der *oriC*-Region ausgeschlossen, ist ein Minichromosom im Gegensatz zum Chromosom, nicht mehr in der Lage zu replizieren (Woelker and Messer, 1993). Ein primäres Chromosom mit einer mutierten DnaA-Box R4 zeigt lediglich ein asynchrones Initiationsverhalten (Bates et al., 1995). Ein ähnliches Bild ergibt sich bei der näheren Analyse von IHF und FIS. Mutationen, die das Binden von IHF und FIS an ihre Bindestellen verhindern, führen dazu, dass Minichromosomen nicht mehr in der Lage sind zu replizieren, wohingegen sich der Effekt bei primären Chromosomen auf die asynchrone Initiation der DNA-Replikation beschränkt (Filutowicz et al., 1992; Kano et al., 1991). Für die Genprodukte der flankierenden Bereiche von *oriC*, *MioC* und *GidA*, konnte keine essentielle Funktion nachgewiesen werden (Leonard and Helmstetter, 1986; Tanaka and Hiraga, 1985). Trotzdem konnte für Minichromosomen gezeigt werden, dass, auch wenn die Transkription dieser Gene nicht essentiell ist, durch Anwesenheit von *mioC* in direkter Nachbarschaft zu *oriC* die Kopienzahl und somit auch die Stabilität der Minichromosomen erhöht wird (Bates et al., 1997; Lobner-Olesen and Boye, 1992). Es wird angenommen, dass die Transkription von *gidA* und *mioC* maßgeblich die Struktur des Replikationsursprungs verändert und diese Veränderung auf Minichromosomen einen höheren Einfluss hat (Asai et al., 1998). Zusammengefasst zeigen die Studien, dass die Anforderungen von *oriC* deutlich voneinander abweichen, wenn es darum geht die DNA-Replikation eines primären Chromosoms oder die eines Minichromosoms zu initiieren. Ob *oriC* einen geeigneten Replikationsursprung für sekundäre Chromosomen darstellt, wird im folgenden Kapitel erläutert.

1.3 Eignen sich ektopische *oriC* als Grundlage für sekundäre Chromosomen in

Escherichia coli?

Umstrukturierungen des Genoms können dafür sorgen, dass es für genetische Manipulationen besser zugänglich ist (Schindler and Waldminghaus, 2015). So war es bereits möglich einen *E. coli* Stamm herzustellen, der statt eines zirkulären Chromosoms ein lineares Chromosom stabil prozessiert (Cui et al., 2007). Des Weiteren konnte das *E. coli* Chromosom in zwei lineare Chromosomen unterteilt werden, die stabil weitervererbt werden (Liang et al., 2013). Für das Unterteilen von einem Chromosom in mehrere Untereinheiten, das sogenannte „Genome-splitting“, ist es essentiell, dass diese stabil von einem Replikationsursprung repliziert werden können. So könnte zunächst ein ektopischer Replikationsursprung, also ein zweiter *oriC* zusätzlich zum nativen *oriC*, auf dem Chromosom inseriert werden. Es konnte gezeigt werden, dass *E. coli* einen zweiten *oriC* innerhalb des Chromosoms toleriert und dass dieser ektopische *oriC*, *oriZ* genannt, in der Lage, ist als alleiniger Replikationsursprung auf dem *E. coli* Chromosom zu fungieren (Ivanova et al., 2015; Wang et al., 2011). *OriZ*, enthält nicht nur die Sequenz von *oriC*, sondern mit 1172 bp bzw. 5160 bp auch die flankierenden Bereiche *gidA* und *mioC* (Wang et al., 2011). Grundlegend ist jedoch zu beachten, dass zwei *oriC*-basierte Chromosomen innerhalb einer Zelle immer miteinander konkurrieren werden. Dies kann dazu führen, dass das sekundäre Chromosom nicht stabil an Tochtergenerationen weitervererbt wird oder aber, dass es über die *oriC*-Sequenz in das primäre Chromosom integriert (Lobner-Olesen, 1999; Messerschmidt et al., 2015). Deshalb ist es notwendig einen anderen, kompatiblen Replikationsursprung zu finden, mit dem es möglich ist ein sekundäres Chromosom herzustellen und stabil aufrecht zu erhalten. *Vibrio cholerae* enthält ein natives, sekundäres Chromosom und bietet so eine sehr gute Vorlage für das Design sekundärer Chromosomen.

1.4 Der Modellorganismus *Vibrio cholerae*: Ein Bakterium mit zwei Chromosomen

Die meisten Prokaryoten tragen ihre genetischen Informationen auf nur einem einzigen, zirkulären Chromosom (Ochman, 2002; Wang et al., 2013). Im Gegensatz dazu gibt es jedoch einige Bakterien, in denen das Erbgut auf multiple Chromosomen verteilt ist (Egan et al., 2005; Jha et al., 2012). Ein Modellorganismus der Gruppe der multi-chromosomalen Bakterien ist das humanpathogene, γ -Proteobakterium *V. cholerae*, welches Auslöser der schweren Infektionskrankheit Cholera ist (Ansaruzzaman et al., 2004; Mohapatra et al., 2007). *V. cholerae* enthält zwei unterschiedlich große Chromosomen: Chromosom I mit 2,96 Mbp und Chromosom II mit 1,07 Mbp (Heidelberg et al., 2000; Trucksis et al., 1998). Chromosom I wird auch als primäres Chromosom bezeichnet. Der Replikationsursprung von Chromosom I, *oril*, gleicht *oriC* aus *E. coli* und wird von dem hochkonservierten Protein DnaA initiiert (Duigou et al., 2006; Egan and Waldor, 2003). Der Replikationsursprung von Chromosom II, *orill*, ähnelt aufgrund vieler Sequenzwiederholungen, die

auch als Iterone bezeichnet werden, stark der Replikationssequenz von Plasmiden (Egan and Waldor, 2003). Deshalb wurde postuliert, dass Chromosom II ursprünglich als Plasmid in *V. cholerae* eingeführt wurde und sich mit der Zeit angepasst hat (Jha et al., 2012; Venkova-Canova and Chattoraj, 2011). Chromosom II enthält essentielle Gene und repliziert zu einem bestimmten Zeitpunkt während des Zellzyklus, womit es zwei charakteristische Merkmale eines Chromosoms aufweist (Egan et al., 2004; Heidelberg et al., 2000; Rasmussen et al., 2007). Hingegen replizieren Plasmide laut Definition unabhängig vom Zellzyklus und weisen keine essentiellen Gene auf (del Solar et al., 1998; Harrison et al., 2010; Nordstrom and Dasgupta, 2006). Deshalb wird Chromosom II in *V. cholerae* als Chromosom bezeichnet. Die Initiation der DNA-Replikation an *oriII* wird von dem Initiatorprotein RctB reguliert, welches für die Gattung der *Vibrionaceae* einzigartig ist (Duigou et al., 2006; Egan and Waldor, 2003). Zudem kodieren beide Chromosomen für Orthologe von Plasmid-Partitionierungssystemen (Heidelberg et al., 2000). Diese beinhalten die Segregationsproteine ParA1/B1 für Chromosom I und ParA2/B2 für Chromosom II (Yamaichi et al., 2007a; Yamaichi et al., 2007b). Wie die Regulation der DNA-Replikation der beiden Chromosomen innerhalb einer Zelle abläuft, war lange Zeit Gegenstand der Forschung und lieferte kontroverse Ergebnisse (Egan et al., 2004; Rasmussen et al., 2007; Stokke et al., 2011). Methoden, die für die Analyse von Bakterien mit einem einzigen Chromosom eingesetzt wurden, stellten sich bei *V. cholerae* als nicht anwendbar dar. So wurde zunächst postuliert, dass beide Chromosomen gleichzeitig mit der DNA-Replikation beginnen, was als Initiationssynchronie bezeichnet wurde (Abb. 2 A) (Egan et al., 2004). 2007 wurde diese These durch Rasmussen et al. widerlegt (Rasmussen et al., 2007). Durch eine kombinatorische Analyse aus Durchflusszytometrie, Marker-Frequenzanalysen und einer Computer-basierten Simulation der DNA-Replikation in *V. cholerae* konnten sie ableiten, dass zunächst Chromosom I und später im Zellzyklus auch Chromosom II mit der DNA-Replikation beginnt. Dieses Modell wird als Terminationssynchronie bezeichnet (Abb. 2 B) (Rasmussen et al., 2007; Stokke et al., 2011) und wird weiterhin durch die Erkenntnis untermauert, dass Chromosom II nach ca. 2/3 der DNA-Replikation von Chromosom I startet (Stokke et al., 2011; Val et al., 2016). Eine neue Studie konnte zeigen, dass die zeitliche Koordinierung der DNA-Replikation von Chromosom II von einer kurzen, intergenischen Sequenz auf Chromosom I abhängig ist (Val et al., 2016). Diese ist eine 150 bp lange Sequenz und wird als *crts*-Sequenz (Chromosome II replication triggering site) bezeichnet (Val et al., 2016). Es wird vermutet, dass die Verdopplung der *crts*-Sequenz das Initiatorprotein RctB aktiviert und so die DNA-Replikation an *oriII* auslöst. Des Weiteren konnte in dieser Studie mittels 3C-Analysen (chromosome conformation capture) ein physikalischer Kontakt zwischen Chromosom I und Chromosom II ausgemacht werden, was auf eine mechanische, funktionelle Interaktion der beiden Chromosomen hinweisen könnte (Val et al., 2016).

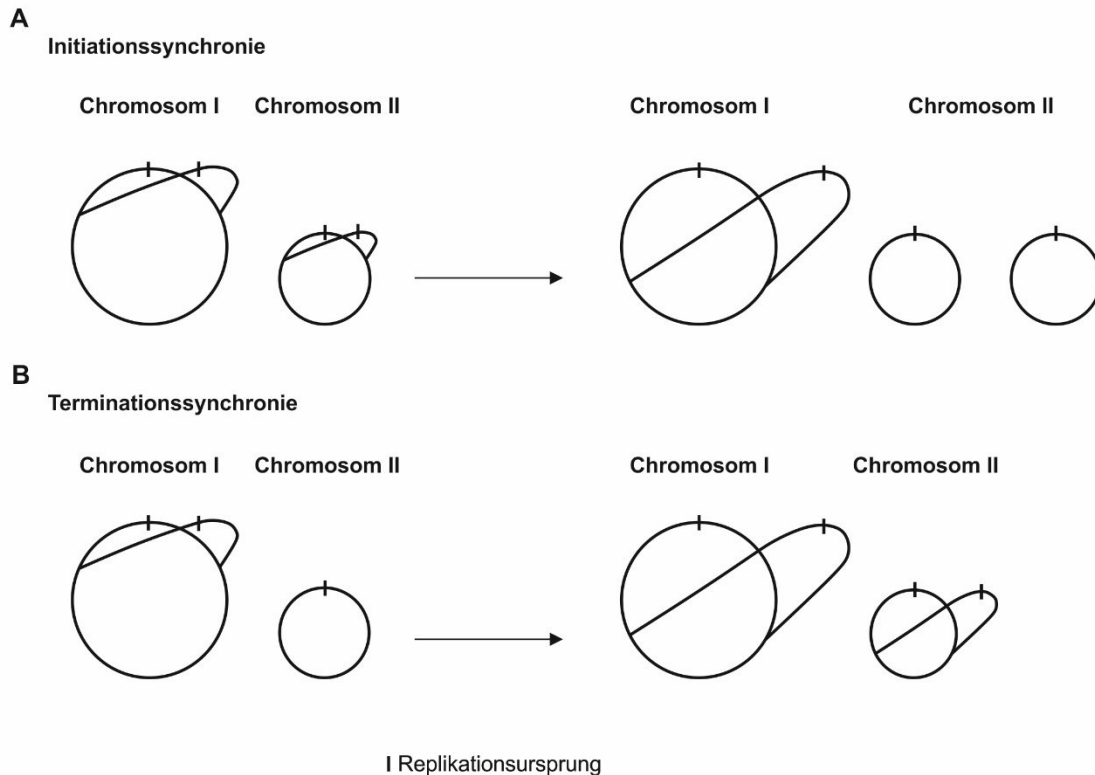


Abbildung 2: Hypothesen zur Regulation der DNA-Replikation der beiden Chromosomen in *V. cholerae*. Der Verlauf der DNA-Replikation von Chromosom I und II wird schematisch dargestellt. Pfeile indizieren den Verlauf des Zellzyklus. **A** Initiationssynchronie. Beide Chromosomen starten zum gleichen Zeitpunkt innerhalb des Zellzyklus mit der DNA-Replikation. Chromosom II beendet die DNA-Replikation vor Chromosom I. **B** Terminationssynchronie. Chromosom I startet zuerst mit der DNA-Replikation. Chromosom II startet im späteren Verlauf im Zellzyklus mit der DNA-Replikation. Die Termination der DNA-Replikation beider Chromosomen findet gleichzeitig statt.

Diese einzigartige Kommunikation zwischen den beiden Chromosomen wird auch als eine Art bakterieller Zellzyklus „Checkpoint“ in *V. cholerae* bezeichnet (Val et al., 2016). Anhand der aufgeführten Studien zur Analyse der DNA-Replikation in *V. cholerae* wird deutlich, dass die zur Verfügung stehenden Methoden nicht hinreichend genutzt werden können, um die Chromosomen multi-chromosomaler Bakterien unabhängig voneinander zu analysieren.

1.5 *orill*: Ein prädestinierter Replikationsursprung für sekundäre Chromosomen in

Escherichia coli

Sekundäre Chromosomen (vergl. Kapitel 1.2.1) können genutzt werden, um grundsätzliche Fragen bezüglich der Charakteristika natürlicher, bakterieller Chromosomen zu beantworten. Mutationen des primären Chromosoms müssen zwingend lebensfähig sein um sie untersuchen zu können. Sekundäre Chromosomen hingegen können manipuliert und analysiert werden, ohne dass die Wirtszelle darunter leidet (Schindler and Waldminghaus, 2015). Sekundäre Chromosomen können ebenfalls genutzt werden um Genome umzustrukturieren. Aus einem großen Chromosom könnten mehrere kleine Untereinheiten generiert werden, die so besser genetisch manipulierbar sind. So war es für *E. coli*

bereits möglich, einen Stamm zu generieren, der statt eines großen, zirkulären Chromosoms zwei kleinere, lineare Chromosomen prozessiert (Liang et al., 2013). Um grundsätzliche Fragen zur Chromosomenbiologie mit sekundären Chromosomen zu klären, ist es essentiell, dass sich diese auch wie Chromosomen verhalten. Ein sehr wichtiges Charakteristikum eines Chromosoms ist, dass es nur einmal zu einem bestimmten Zeitpunkt im Zellzyklus repliziert (Boye et al., 2000). Deshalb ist bei dem Design sekundärer Chromosomen zu beachten, dass ein Replikationsursprung gewählt wird, der dieses Kriterium erfüllt. Chromosom II von *V. cholerae* stellt ein natives Modell dar, um sekundäre Chromosomen in andere Organismen zu generieren. Der Replikationsursprung von Chromosom II, *orill*, ist in der Lage, erfolgreich in dem heterologen Wirtsorganismus *E. coli* zu replizieren (Egan and Waldor, 2003; Liang et al., 2013; Messerschmidt et al., 2015). Dies konnte anhand von *orill*-basierten Replikons festgestellt werden (Egan and Waldor, 2003). Mit Hilfe dieses Modells konnte ferner definiert werden, welche Faktoren für eine erfolgreiche DNA-Replikation von *orill* in *E. coli* notwendig sind. Neben der Sequenz von *orill* benötigt es dazu das Initiatorprotein RctB, welches nicht in *E. coli* vorhanden ist. Des Weiteren werden die Gene *parAB2* für eine stabile Vererbung von *orill*-basierten Replikons benötigt (Egan and Waldor, 2003; Yamaichi et al., 2007a). Messerschmidt et al. konnten ein *orill*-basiertes synthetisches Chromosom, *synVicII*, mit den genannten Faktoren in *E. coli* konstruieren (Messerschmidt et al., 2015). Analysen zum Initiationszeitpunkt von *synVicII* in *E. coli* suggerieren, unter der Annahme, dass das Replikon Zellzyklus-spezifisch repliziert, dass es nach ca. 53% der DNA-Replikation des primären Chromosoms mit der DNA-Replikation starten könnte (Messerschmidt et al., 2015). Für das native, sekundäre Chromosom in *V. cholerae* konnte gezeigt werden, dass es nach 54-66% der DNA-Replikation des primären Chromosoms initiiert (Stokke et al., 2011). Eine weitere, wichtige Eigenschaft eines Chromosoms, nämlich die stabile Vererbung an die Tochtergenerationen, konnte auch für *synVicII* festgestellt werden. Im direkten Vergleich dazu zeigte ein *oriC*-basiertes Replikon eine geringere Stabilität (Messerschmidt et al., 2015). Aufgrund der unterschiedlichen Initiatorproteine von *oriC* und *orill*-basierten Replikons sollte es nicht zu einer konkurrierenden Situation in *E. coli* kommen. *orill* bietet somit ideale Voraussetzungen, um als Replikationsursprung von synthetischen, sekundären Chromosomen in *E. coli* zu fungieren. Wie die DNA-Replikation sekundärer Chromosomen untersucht werden könnte, wird im nächsten Kapitel aufgezeigt.

1.6 Die Kopienzahl eines genetischen Lokus auf Einzelzellniveau als Reporter für die DNA-Replikation

Seitdem wir technisch in der Lage sind das humane Genom zu sequenzieren, wird deutlich, dass die Genomsequenz zwischen den Menschen stärker variiert, als bisher angenommen (Beckmann et al., 2007). Die Kopienzahlvariation (CNV), die nach dem bisherigen Erkenntnisstand 10% des humanen Genoms umfasst, spielt dabei eine wichtige Rolle (Shlien and Malkin, 2009). Als CNV wird ein DNA

Segment von 1 kb oder mehr beschrieben, das in variabler Kopienzahl im Vergleich zu einem Referenzgenom vorkommt (Feuk et al., 2006). CNVs können einfache Duplikationen oder auch Deletionen beinhalten. Sie können dazu genutzt werden Individuen voneinander zu differenzieren und werden beim Menschen auch mit zahlreichen Erkrankungen wie beispielsweise Schizophrenie oder Krebs in Verbindung gebracht (Redon et al., 2006; Sebat et al., 2004). CNVs haben somit einen hohen Stellenwert was die Prädisposition, die genetisch bedingte Ausbildung von Krankheiten, angeht und werden meist anhand von Microarray-basierter komparativer genomischer Hybridisierung (CGH) oder Genomsequenzierung identifiziert (Mills et al., 2011; Pollack et al., 1999). Neben diesen CNVs kann die Kopienzahl eines Gens innerhalb der gesamten Zelle auch Aufschluss über den Status der DNA-Replikation geben. Während der DNA-Replikation wird das gesamte Chromosom dupliziert. Vor der DNA-Replikation liegt ein Genloki in einfacher Kopie und nach der Replikation in doppelter Kopienzahl in der Zelle vor (Abb. 3). Auf Grundlage dieser Annahme kann die DNA-Replikation anhand der Kopienzahl der Gene auf dem Chromosom verfolgt werden.

So ist die Kopienzahlbestimmung ein interessanter Ansatz für die Analyse von Pro- als auch Eukaryoten. Anhand eines multi-chromosomalen Bakteriums wird in Abbildung 3 dargestellt, wie der Status der DNA-Replikation anhand der Kopienzahl von definierten Genloki und der Anzahl an Replisomen verfolgt werden kann. Eine Methode um den Replikationsstatus einer Zelle genauer zu charakterisieren, ist die Bestimmung des Kopienzahlverhältnis von Replikationsursprung zu Terminationsregion (Bremer and Churchward, 1977; Stokke et al., 2011). Bei einem Verhältnis von 1 liegen beide Regionen in gleicher Kopienzahl vor (Abb. 3 A). Somit wurde je nach Zellstadium die DNA-Replikation noch nicht gestartet oder bereits beendet. Ist das Verhältnis größer als 1, liegt der Replikationsursprung in höherer Kopienzahl vor als die Terminusregion und wurde deshalb bereits repliziert (Abb. 3 B + C). Mittels quantitativer PCR (qPCR) können die zu vergleichenden genetischen Regionen spezifisch amplifiziert und anschließend das Verhältnis dieser bestimmt werden (Abb. 3 A, B, C) (Hill et al., 2012). Nachteil dieser Methode ist, dass das Multiplexen, die Analyse von mehreren Genen in einer Reaktion, limitiert ist (Klein, 2002). PCR-basierte Methoden beinhalten zudem das Problem der ungleichen Anreicherung von DNA-Fragmenten, was zu falschen Ergebnissen führen kann (Smith and Osborn, 2009). CGH kann ebenfalls genutzt werden, um Kopienzahlen genetischer Loki zu bestimmen (Pinkel et al., 1998). Im Gegensatz zur qPCR umfasst CGH die Analyse des gesamten Genoms, wodurch ein detaillierteres Bild der Kopienzahlen entsteht und der Verlauf der DNA-Replikation erkennbar wird. Diese Methode basiert auf der simultanen Hybridisierung der Proben- und Referenz-DNA, die mit unterschiedlichen Fluoreszenzfarbstoffen markiert sind, an immobilisierte DNA-Fragmente (Govindarajan et al., 2012). Das sich ergebende Verhältnis der Fluoreszenzintensitäten von Probe zu Referenz ist proportional zum Verhältnis der Kopienzahl. Das Ausgangsmaterial für eine qPCR oder CGH ist bisher meist die Gesamt-DNA einer Zellpopulation. Das Ergebnis spiegelt also den

Durchschnittszustand der Population wieder. Doch jede Zelle ist individuell und kann Unterschiede in der Genomsequenz, der Genexpression und der Mengen an Proteinen aufweisen (Lidstrom and Konopka, 2010). Um diese Heterogenität näher zu untersuchen, ist die Analyse von Einzelzellen unabdingbar. So kann der Replikationsstatus einer einzigen Zelle anhand der Anzahl von Replisomen bestimmt werden. Pro Replikationsursprung assemblieren zwei Replisomen (Abb. 3 B + C). Replisomen können visualisiert werden, indem eine Untereinheit, beispielsweise DnaB oder Proteine, die sich in der Nähe des Replisoms befinden, mit einem Fluoreszenzprotein markiert werden (Waldminghaus et al., 2012; Wang and Shapiro, 2004). Mittels Fluoreszenzmikroskopie können die Replisomen als distinkte Signale innerhalb der Zelle ausgemacht und gezählt werden.

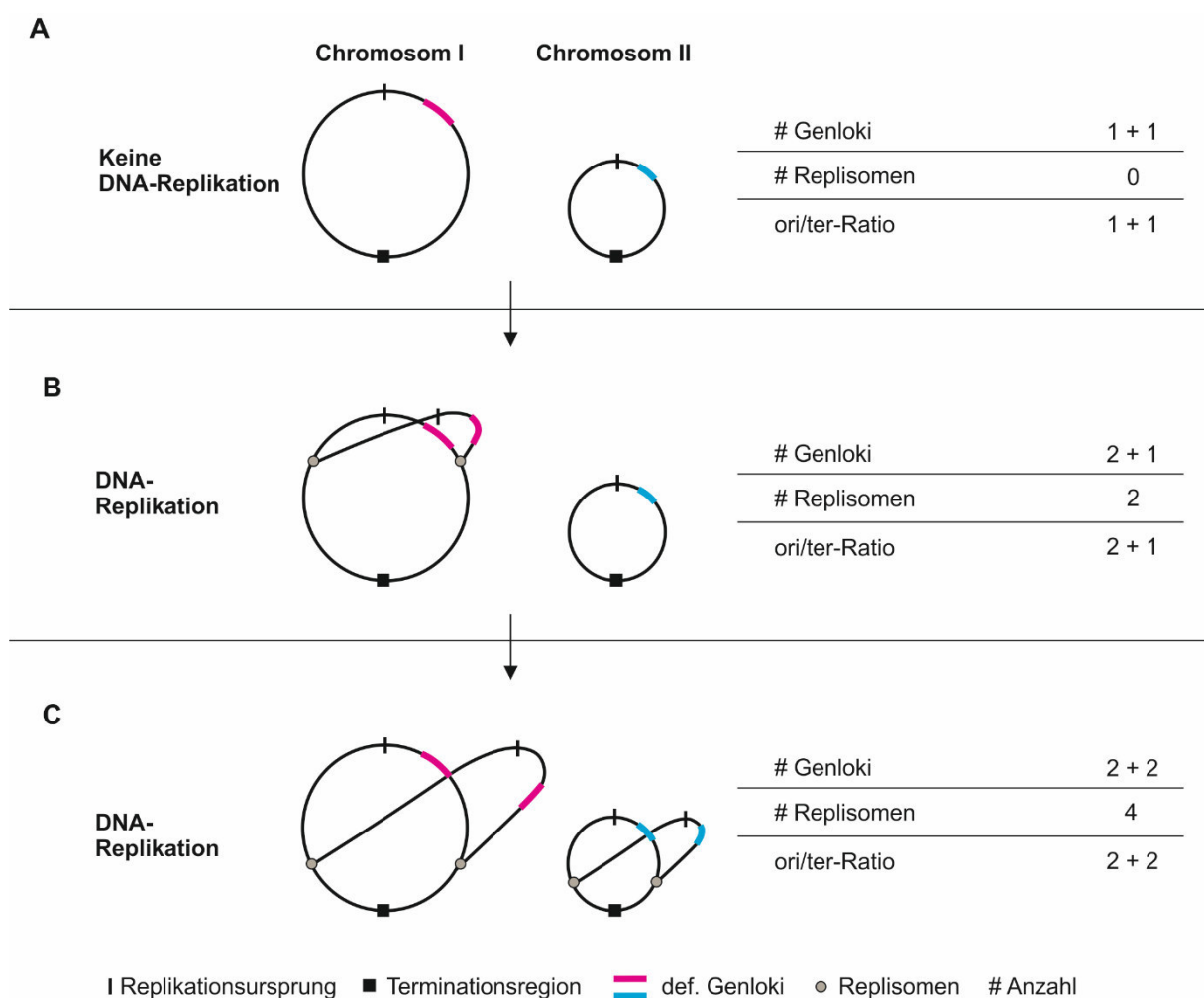


Abbildung 3: Die Kopienzahl eines genetischen Lokus als Reporter für den Status der DNA-Replikation. A, B, C Verschiedene Stadien der DNA-Replikation am Beispiel des Zellzyklus eines multi-chromosomalen Bakteriums. Replikationsursprünge, genetische Loci, Replisomen und Terminationsregionen sind hervorgehoben. Die Anzahl von Genloki, Replisomen und das Kopienzahlverhältnis von Replikationsursprung zu Terminationsregion (ori/ter-Ratio) sind für das jeweilige Replikationsstadium (A, B, C) aufgeführt. Pfeile indizieren den Verlauf des Zellzyklus.

Für die Quantifizierung von Bakterien mit zwei Chromosomen ergäbe sich folgendes Bild: Kann kein Replisom innerhalb der Zelle detektiert werden, wurde die DNA-Replikation noch nicht gestartet oder bereits abgeschlossen (Abb. 3 A). Zellen, die zwei Foci aufweisen, spiegeln die DNA-Replikation von nur einem Chromosom wieder (Abb. 3 B), wohingegen vier Foci auf die DNA-Replikation beider Chromosomen hinweisen (Abb. 3 C). Nachteil dieser Methode ist jedoch, dass die markierten Replisomen nicht immer gut zu differenzieren sind. Es ist demnach möglich, dass beide Replisomen eines Chromosoms überlagern. Eine detailliertere Auflösung der DNA-Replikation kann mit der digitalen PCR (dPCR) erreicht werden. Mit Hilfe dieser Methode kann die Kopienzahl von einzelnen DNA-Molekülen bestimmt werden (Sykes et al., 1992). Dafür wird eine DNA-Probe verdünnt und auf viele, einzelne, parallele PCR-Reaktionen aufgeteilt, von denen einige das zu analysierende DNA-Molekül enthalten und einige nicht (Vogelstein and Kinzler, 1999). Durch das Zählen von positiven und negativen Amplifizierungsreaktionen, die proportional zur eingesetzten DNA-Menge auftreten, kann die Kopienzahl der DNA-Moleküle in der analysierten Probe bestimmt werden (Baker, 2012). Neben dem Verdünnen von DNA-Proben um einzelne Moleküle zu erhalten, kann auch die DNA einer einzigen Zelle verwendet werden. Ein grundsätzliches Problem der Analyse auf Einzelzellniveau ist die Zellgröße. Der Arbeitsmaßstab bewegt sich im Bereich von Mikrometern und Pikolitern (Zhao et al., 2008). Trotzdem gibt es bereits zahlreiche Methoden um eukaryotische und prokaryotische Einzelzellen zu isolieren (Blainey, 2013). Dazu zählen die klassische Verdünnung einer Kultur bis zu einem Grenzwert, das Mikropipettieren, bei dem mit einer Pipette und einem Mikroskop einzelne Zellen transferiert werden oder auch FACS (Fluorescence-Activated Cell Sorting) (Grindberg et al., 2011; Navin et al., 2011; Raghunathan et al., 2005; Zhang et al., 2006). Zudem gibt es viele Methoden, welche die Gewinnung von Einzelzellen mit Hilfe des Fortschritts der Nanotechnologie durchführen. Dazu zählen unter anderem „microfluidic“-Systeme (Kortmann et al., 2011; Yin and Marshall, 2012; Zare and Kim, 2010). „microfluidic“-Systeme bieten viele Vorteile für das Arbeiten mit einzelnen Zellen, weil sie die Parallelisierung und Automatisierung der Analysen und eine erhöhte Sensitivität sowie Minimierung von Kontaminationen durch das Arbeiten in geringen Reaktionsvolumina ermöglichen (Gawad et al., 2016; Lecault et al., 2012; Marcus et al., 2006). Je nachdem welche Analysen mit Einzelzellen durchgeführt werden sollen, muss eine Zellyse stattfinden, um Komponenten wie DNA, RNA oder Proteine zugänglich zu machen (Blainey, 2013). Dabei ist zu beachten, dass nachfolgende Reaktionen nicht durch Zell- oder Pufferreste inhibiert werden. Eine *E. coli* Zelle enthält ungefähr 5 fg DNA (Raghunathan et al., 2005). Für Analysen, die mehr Ausgangsmaterial benötigen, wie beispielsweise Genomsequenzierungen, muss die DNA der Einzelzelle angereichert werden (de Bourcy et al., 2014; Gonzalez-Roca et al., 2010; Kalisky et al., 2011; Tang et al., 2010). Mittels der sogenannten „Whole Genome Amplification“ (WGA) kann das komplette Genom einer Zelle amplifiziert werden, sodass am Ende DNA-Mengen vorliegen, mit denen nachfolgende Analysen durchgeführt werden können

(Blainey, 2013; Huang et al., 2015; Yilmaz and Singh, 2012). Bei der Amplifizierung der DNA von Einzelzellen kann es jedoch zu ungleichen Anreicherungen oder unspezifischen Produkten kommen (Pan et al., 2008). Diese Unterschiede in der Anreicherung können das Ergebnis nachfolgender Analysen verfälschen. Es liegt demnach Nahe, das Repertoire an Methoden zur Analyse von Einzelzellen zu erweitern.

1.7 Moderne Klonierungssysteme

Das Anordnen von kleinen DNA-Fragmenten zu großen Konstrukten, die DNA-Assemblierung, hat in den letzten Jahren bedeutende Weiterentwicklungen erlebt (Chao et al., 2014; Ellis et al., 2011). Die klassische Restriktions-Ligationsmethode, die in den 1970er Jahren entwickelt wurde, führte zu einer Revolution in der Biotechnologie (Cohen et al., 1973). Seitdem wuchsen die Ansprüche und die Komplexität von DNA-Konstrukten, die mehrere Gene und intergene Komponenten enthalten. Zusätzlich sinken die Kosten der DNA-Synthese stetig und der Synthesemaßstab nimmt zu (Kosuri and Church, 2014). Trotzdem ist es noch nicht möglich ganze Genome zu synthetisieren. DNA-Fragmente, die den Synthesemaßstab überschreiten, werden deshalb noch immer durch DNA-Assemblierungen hergestellt (Cobb et al., 2014; Gibson et al., 2008a). Deshalb wurden mit der Zeit immer mehr Methoden etabliert, die eine effiziente, modulare und schnelle DNA-Assemblierung erlauben. Diese DNA-Assemblierungsmethoden können in zwei Gruppen unterteilt werden: (i) Homologie-basierte Methoden und (ii) Restriktionsenzym-basierte Methoden. Anhand zwei bekannter Methoden werden diese in Abbildung 4 schematisch dargestellt.

1.7.1 Homologie-basierte Methoden zur DNA-Assemblierung

Diese DNA-Assemblierungsmethoden basieren auf der Hybridisierung von 20-40 bp langen, komplementären Überhängen zwischen den DNA-Fragmenten (Abb. 4 A). Zu den *in vivo* Methoden zählt die native, homologe Rekombination in *Saccharomyces cerevisiae* und *Bacillus subtilis* (Gibson et al., 2008b; Itaya et al., 2008). Zu den *in vitro* Methoden zählen die Overlapping-Extension PCR (OE-PCR), das Gibson Assembly und die Ligase chain reaction (LCR) (de Kok et al., 2014; Gibson et al., 2009; Horton et al., 1989). Das Gibson Assembly basiert auf einem Enzym-Mix bestehend aus einer T5-Exonuklease, einer DNA-Ligase und einer DNA-Polymerase (Gibson et al., 2009). Die T5-Exonuklease generiert einzelsträngige Überhänge an den DNA-Fragmenten, die dann aufgrund ihrer komplementären Sequenz hybridisieren. Im Anschluss füllt die DNA-Polymerase die DNA wieder auf und die DNA-Ligase verknüpft die DNA-Stränge miteinander (Gibson et al., 2009). Mit Hilfe der aufgeführten Methoden ist es in einer einzigen Reaktion möglich mehrere DNA-Fragmente zielgerichtet zusammenzubauen (Chao et al., 2014; Ellis et al., 2011). Nachteil der Homologie-basierten Methoden ist, dass keine DNA-Fragmente mit repetitiven Sequenzen assembliert werden können, weil

die sich wiederholenden Regionen die Assemblierung stören und zu falschen Endprodukten führen können. Ein weiterer Nachteil ist, dass für die meisten DNA-Fragmente, die mittels PCR amplifiziert werden, Oligonukleotide benötigt werden, um die homologen Bereiche anzufügen. Diese können meist nicht wiederverwendet werden, weil die enthaltenen homologen Bereiche für ein spezifisches Produkt entworfen wurden. Oligonukleotide sind mittlerweile günstig, jedoch summieren sich die Kosten, wenn man den Klonierungsmaßstab auf DNA-Fragment-Bibliotheken ausweiten möchte. Gibson Assembly funktioniert prinzipiell auch mit einzelsträngiger DNA, wodurch die Kosten drastisch gesenkt werden könnten.

1.7.2 Restriktionsenzym-basierte Methoden zur DNA-Assemblierung

Diese Art der DNA-Assemblierung basiert auf dem klassischen Restriktions-Ligationsverfahren. DNA-Fragmente mit flankierenden Restriktionsenzym-schnittstellen werden verdaut und gemäß ihrer Überhänge miteinander ligiert. Zu diesen Methoden zählen unter anderem BioBrick™ und das Golden Gate cloning (Engler et al., 2009; Engler et al., 2008; Shetty et al., 2008). Das BioBrick™-System hat diese Grundlage standardisiert, in dem alle DNA-Fragmente eine gemeinsame Schnittstelle aufweisen, durch die sie sich kombinieren lassen. Durch wiederholende Restriktions-Ligationsschritte konnten so große DNA-Fragmente hergestellt werden (Sarrion-Perdigones et al., 2011; Smolke, 2009). Das BioBrick™-System verwendet Typ-II-Restriktionsenzyme wie *EcoRI*. Nachteil dieser Methode ist, dass die Erkennungssequenz zwischen den DNA-Fragmenten als sogenannte „Narbe“ erhalten bleibt und nicht erneut mit dem Enzym geschnitten werden kann. Typ-IIS-Restriktionsenzyme schneiden außerhalb ihrer Erkennungssequenz, wodurch die generierten Überhänge zwischen den DNA-Fragmenten theoretisch frei gestaltet werden können. Basierend auf der Eigenschaft von Typ-IIS-Restriktionsenzymen wurden hierarchische Assemblierungssysteme, wie das Golden Gate cloning, erstellt, die auf Vektoren mit definierten 4 bp Überhängen basieren, die sequentiell zueinander passen (Engler et al., 2009; Engler et al., 2008). Aufgrund der Kombinierbarkeit bleibt jedoch immer eine „Narbe“ zwischen den DNA-Fragmenten zurück. Die definierten Überhänge bieten Restriktionsenzym-basierten Methoden die Möglichkeit Sequenzen mit repetitiven Regionen zu klonieren (vgl. 1.7.1). Allen Restriktionsenzym-basierten Methoden bleibt jedoch gemein, dass die Erkennungssequenzen der verwendeten Enzyme möglicherweise in den zu assemblierenden DNA-Fragmenten auftauchen können. Diese zunächst zu entfernen führt zu einem deutlichen Mehraufwand. Auch die verbleibenden Überhänge der Restriktionsenzyme können problematisch sein. Eine exakte Zielsequenz kann mit diesen Systemen nicht nachgebaut werden.

1.7.3 Modulares Klonieren mit MoClo

Ein sehr bekanntes und effizientes DNA-Assemblierungssystem ist MoClo (Modular Cloning), welches auf dem bereits erwähnten Golden Gate cloning aufbaut und von der Gruppe um Marillonnet optimiert wurde (Weber et al., 2011). Es basiert auf drei grundlegenden Eigenschaften: (i) verschiedenen Vektorsets, dessen 4 bp Überhänge der Reihenfolge nach zueinander passen, (ii) dem Wechsel zwischen diesen Vektorsets (LevelM und Level1) aufgrund unterschiedlicher Resistenzen sowie Restriktionsenzymen und (iii) dem Verwenden von Endlinkersets, die die Brücke zwischen dem assemblierten DNA-Fragment und dem Vektor bilden und durch Enthalten der Erkennungssequenz des Typ-IIS-Enzyms eine Restriktion im nächsten Level ermöglichen. Zudem müssen keine DNA-Fragmente aufgearbeitet werden, weil die MoClo-Reaktion, welche ein Typ-IIS-Restriktionsenzym und eine DNA-Ligase enthält, mit Plasmiden durchgeführt wird.

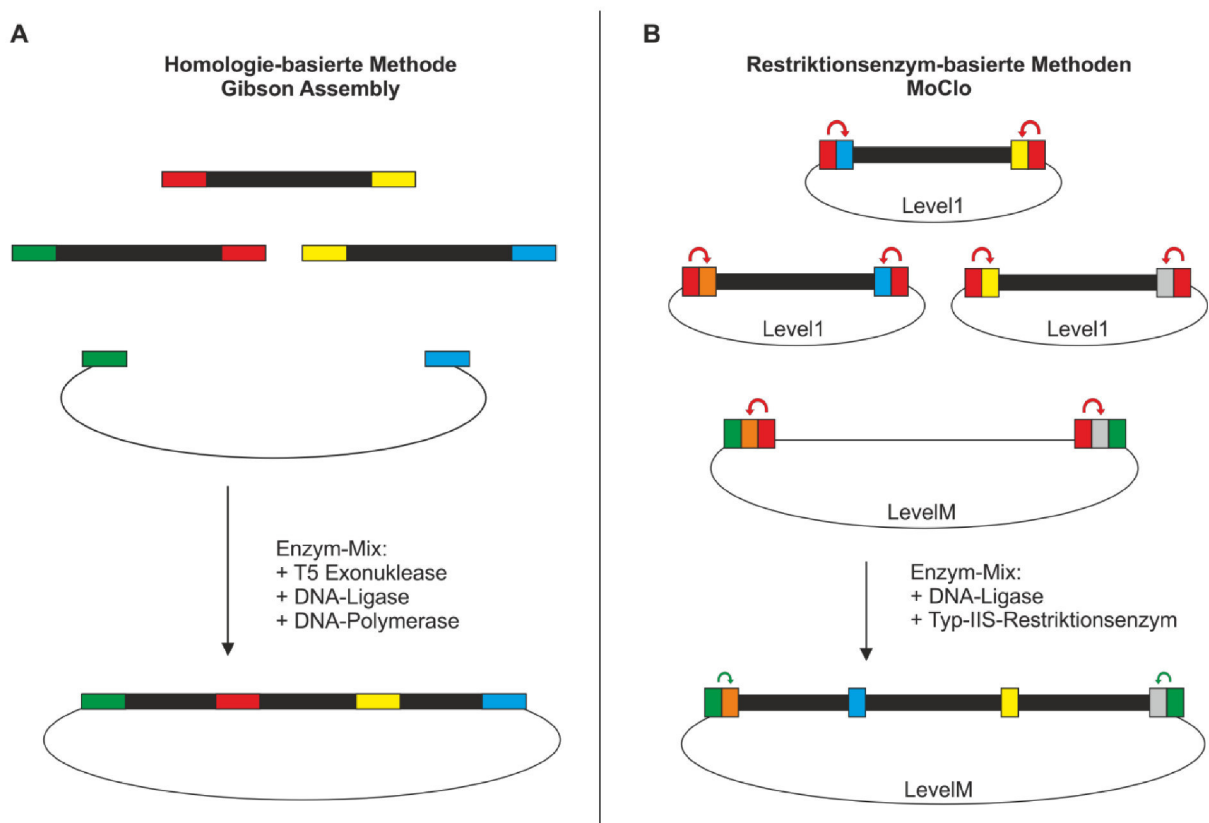


Abbildung 4: Schematische Darstellung verschiedener DNA-Assemblierungstechniken am Beispiel von Gibson Assembly und MoClo. A Die Hybridisierung der homologen Bereiche (farblich markiert) vermittelt die DNA-Assemblierung zwischen den verschiedenen DNA-Fragmenten. **B** 4 bp lange Überhänge (farblich markiert), die von Typ-IIS-Restriktionsenzymen generiert werden, vermitteln die Assemblierung der DNA-Fragmente.

Die DNA-Fragmente werden durch das Restriktionsenzym aus dem Plasmid herausgeschnitten, sodass alle DNA-Fragmente mit den jeweils passenden 4 bp Überhängen vorliegen und zusammen mit dem Endlinker in den Zielvektor ligiert werden können (Abb. 4 B). Durch den Wechsel zwischen den Vektorsets erlaubt das MoClo-System eine schnelle und effiziente DNA-Assemblierung.

2. Ziele dieser Arbeit

Bakterielle Chromosomen enthalten einen Replikationsursprung, von dem die DNA-Replikation gestartet wird (Ochman, 2002). Eukaryoten hingegen enthalten mehrere Replikationsursprünge pro Chromosom (Mechali, 2010). Zudem besteht das Genom von Eukaryoten aus mehreren linearen Chromosomen, wohingegen die Mehrheit der Bakterien ein einziges, zirkuläres Chromosom prozessiert (Bendich and Drlica, 2000; Stanier and Van Niel, 1962). Diese distinkte Verteilung zwischen den beiden Domänen wirft einige Fragen auf. Vor allem wie die DNA-Replikation in Bakterien mit den aufgeführten, eukaryotischen Charakteristika ablaufen könnte.

Die vorliegende Arbeit sollte zum Verständnis der DNA-Replikation von Bakterien mit multiplen Replikationsursprüngen sowie von Bakterien mit multiplen Chromosomen beitragen.

Zum einen sollten mittels Durchflusszytometrie und Microarray-basierter komparativer genomischer Hybridisierung (CGH) Bakterien mit multiplen Replikationsursprüngen sowie multiplen Chromosomen hinsichtlich ihres DNA-Gehalts und des Initiationsverhaltens analysiert werden um folgende Frage zu adressieren:

1. Wie erfolgt die Initiation der DNA-Replikation zweier gleicher bzw. unterschiedlicher Replikationsursprünge innerhalb einer Zelle?

Zum anderen sollte im Rahmen dieser Dissertation eine fluoreszenzbasierte Methode etabliert werden, die es ermöglicht, die Kopienzahl von Genloki und somit den Replikationsmechanismus von multi-chromosomalen Bakterien auf Einzelzellebene zu untersuchen. Die folgenden Fragestellungen sollten beantwortet werden:

1. Wie kann die Kopienzahl von Genloki auf Einzelzellebene bestimmt werden?
2. Kann ein Klonierungssystem für die Herstellung von repetitiven Elementen optimiert werden?

Zur Bestimmung der Kopienzahl auf Einzelzellebene sollte die neue Methode BiFCROS (Bimolekular Fluorescence complementation and Repressor Operator System) in *E. coli* etabliert und getestet werden. Dazu sollten neue Klonierungstechniken verwendet und fluoreszenzmikroskopische Analysen durchgeführt werden.

3. Ergebnisse

Dieser Teil der Arbeit gliedert sich in drei Kapitel. Dabei enthält jedes Kapitel unterschiedliche Studien, die bereits publiziert sind (Kapitel 3.1 und 3.2) oder zur Publikation vorbereitet werden (Kapitel 3.3). Vor jedem Abschnitt ist eine Zusammenfassung der relevanten Ergebnisse und eine Übersicht der Autoren sowie ihr Mitwirken an der jeweiligen Publikation angeführt.

3.1 DNA replication in engineered *Escherichia coli* genomes with extra replication origins

**Sarah Milbredt, Neda Farmani, Patrick Sobetzko
and Torsten Waldminghaus
(2016)**

ACS Synthetic Biology

3.1 DNA replication in engineered *Escherichia coli* genomes with extra replication origins

Eukaryotische Chromosomen enthalten multiple Replikationsursprünge (Robinson and Bell, 2005). Prokaryoten starten die DNA-Replikation des Chromosoms von nur einem einzigen Replikationsursprung in bidirektionaler Orientierung (Ochman, 2002). Dabei ist *E. coli* sehr wohl in der Lage, mehr als einen Replikationsursprung pro Chromosom zu tolerieren (Wang et al., 2011). Dieser zusätzliche, ektopische Replikationsursprung, *oriZ*, wurde 100 kb vom nativen Replikationsursprung inseriert (Wang et al., 2011). Es konnte zudem gezeigt werden, dass *oriZ* als alleiniger Replikationsursprung auf dem Chromosom fungieren kann (Ivanova et al., 2015; Wang et al., 2011). Ob die Aktivität des Replikationsursprungs von der Position auf dem Chromosom beeinflusst wird, könnte durch Stämme analysiert werden, die ektopische *oriC* an verschiedenen Positionen auf dem Chromosom enthalten (Liang et al., 2013). Der gleiche Arbeitskreis konstruierte *E. coli* Stämme mit einem ektopischen *oriII*, dem Replikationsursprung des zweiten Chromosoms von *V. cholerae* (Liang et al., 2013). Durch die Analyse dieser Stämme könnte untersucht werden, ob der wirtsfremde Replikationsursprung *oriII*, die DNA-Replikation in *E. coli* beeinflusst.

Die vorliegende Studie präsentiert die Analyse von *E. coli* Stämmen, die mehr als einen Replikationsursprung auf dem Chromosom enthalten, bezüglich ihrer DNA-Replikation. Analysen der *E. coli* Stämme von Liang et al. mittels Microarray-basierter komparativer genomischer Hybridisierung (CGH) zeigten, dass die ektopischen *oriC* auf dem Chromosom nicht in der Lage sind zu replizieren. Um herauszufinden, ob diese Inaktivität an der verwendeten *oriC*-Sequenz liegt, wurden drei neue *E. coli* Stämme generiert, die längere *oriC*-Sequenzen als ektopische Replikationsursprünge auf dem Chromosom enthalten. Die Initiation der DNA-Replikation konnte für alle drei Sequenzen festgestellt werden. Die von Liang et al. verwendete Sequenz von *oriC* war somit nicht ausreichend für die Replikation des gesamten Chromosoms. Stämme die *oriII* als ektopischen Replikationsursprung enthielten, zeigten im Gegensatz zum Wildtyp einen erhöhten DNA-Gehalt. CGH-Analysen konnten die Replikationsaktivität der ektopischen *oriII* nachweisen. *oriII* war somit in der Lage in *E. coli* auf dem Chromosom zu replizieren. Analysen von *oriII* als ektopischer Replikationsursprung wiesen auf eine Dysregulation der DNA-Replikation innerhalb dieser Zellen hin. Diese Dysregulation wurde aufgehoben, sobald sich *oriC* und *oriII* nicht mehr auf dem gleichen Chromosom, sondern auf zwei unterschiedlichen Chromosomen befanden. Um ein besseres Verständnis der Regulation der DNA-Replikation von *oriC* und *oriII* auf einem Chromosom zu erhalten, wurde eine *in silico* Simulation der DNA-Replikation durchgeführt. Zusammen mit den Ergebnissen der experimentellen Arbeit kann abgeleitet werden, dass die Regulation der DNA-Replikation von *oriC* und *oriII* abhängig von deren Lokalisierung ist.

In Zusammenarbeit mit Prof. Dr. T. Waldminghaus betreute ich die Masterarbeit von N. Farmani. Zudem generierte ich drei neue *E. coli* Stämme mit unterschiedlichen *oriC*-Konstrukten und analysierte diese mittels CGH. N. Farmani überprüfte die von Liang et al. erhaltenen *E. coli* Stämme mittels PCR und analysierte sie mittels Durchflusszytometrie und CGH. Die Auswertung der CGH-Daten und die Computer-basierte Simulation der DNA-Replikation erfolgte von Dr. P. Sobetzko. Das Manuskript wurde von Dr. P. Sobetzko, Prof. Dr. T. Waldminghaus und mir verfasst. Dr. P. Sobetzko, Prof. Dr. T. Waldminghaus und ich überarbeiteten ferner das Manuskript nach der ersten Revision.

**3.2 Design and assembly of DNA sequence libraries for
chromosomal insertion in bacteria based on a set of modified
MoClo vectors**

**Daniel Schindler, Sarah Milbredt, Theodor Sperlea
and Torsten Waldminghaus
(2016)**

ACS Synthetic Biology

3.2 Design and assembly of DNA sequence libraries for chromosomal insertion in bacteria based on a set of modified MoClo vectors

Es gibt zahlreiche moderne Klonierungssysteme mit denen es möglich ist, große DNA-Fragmente herzustellen. Eines dieser Systeme ist das MoClo-System, was zunächst für die Anwendung in Pflanzen etabliert und später auch für Arbeiten in Bakterien angepasst wurde (Weber et al., 2011). In der vorliegenden Studie wurde das MoClo-System bezüglich des Klonierens im Bibliotheken-Maßstab und der Genomintegration in *E. coli* optimiert. So wurde die Klonierungseffizienz durch das Einfügen eines *ccdB*-Gens verbessert, die Narbensequenz zwischen den assemblierten DNA-Fragmenten verkürzt und die Integration in das bakterielle Chromosom durch den Austausch des Replikationsursprungs mit einem konditionalen Ursprung vereinfacht. Die effiziente Anwendbarkeit des optimierten MoClo-Systems wurde anhand des Aufbaus eines LacO/TetO-Arrays gezeigt. Dieses Operator-Array enthält Tandemwiederholungen der Bindestellen für LacI und TetR und stellt aufgrund der repetitiven Sequenzen eine Herausforderung für diverse Klonierungsmethoden dar. Die initiale DNA-Fragmentbibliothek für die Klonierung des LacO/TetO-Arrays wurde von dem Programm MARSeG (Motif Avoiding Randomized Sequences Generator) entworfen. Dieses Computerprogramm bietet die Möglichkeit DNA-Sequenzen zu generieren, die eine hohe Diversität aufweisen, gleichzeitig aber diverse Motive ausschließen. Zu diesen Motiven können beispielsweise Erkennungssequenzen von Restriktionsenzymen gehören, mit denen die MoClo-Reaktion stattfindet und die zu Interferenzen während des Konstruktionsprozesses führen könnten. Basierend auf der von MARSeG generierten Oligonukleotidbibliothek wurde das LacO/TetO-Array mit Hilfe des MoClo-Systems generiert und in das Genom von *E. coli* integriert. Mit Hilfe fluoreszenzmikroskopischer Analysen konnte gezeigt werden, dass der mit mVenus fusionierte Repressor TetR an die Tet-Operatoren binden kann. Ein Stabilitätstest zeigte zudem, dass variable Sequenzen zwischen den Bindestellen essentiell für die Beständigkeit des LacO/TetO-Arrays sind. Ein LacO/TetO-Array mit variablen Zwischenregionen liegt auch nach 144 h Inkubation stabil in der Zelle vor. Gleiche Zwischensequenzen führten hingegen schon nach 72 h zum Abbau des LacO/TetO-Arrays.

Die Aufgabenbearbeitung erfolgte nach Absprache mit Prof. Dr. T. Waldminghaus. Ich stellte in Zusammenarbeit mit D. Schindler das LacO/TetO-Array her und überprüfte dessen Stabilität mittels PCR und Southern-Blot-Analysen. Zudem führte ich die fluoreszenzmikroskopischen Analysen und D. Schindler die Optimierung des MoClo-Systems durch. T. Sperlea schrieb das Programm MARSeG, führte Sequenzvergleiche durch und generierte die Oligonukleotidsequenz für das LacO/TetO-Array. Das Manuskript wurde von D. Schindler, T. Sperlea, Prof. Dr. T. Waldminghaus und mir verfasst. D. Schindler, T. Sperlea, Prof. Dr. T. Waldminghaus und ich überarbeiteten ferner das Manuskript nach der ersten Revision.

3.3 BiFCROS: A low background Fluorescence Repressor Operator System

**Sarah Milbredt and Torsten Waldminghaus
(2016)**

prepared for submission

3.3 BiFCROS: A low background Fluorescence Repressor Operator System

Die DNA-Replikation, Chromosomensegregation und Chromosomenfaltung sind wichtige, zentrale Prozesse aller lebenden Zellen. Um einen besseren Einblick in diese Vorgänge zu bekommen, kann die räumliche und zeitliche Organisation der DNA beobachtet werden. Für die Verfolgung von genetischen Regionen auf der DNA können diese markiert werden. Eine bereits etablierte Methode, FROS (Fluorescence Repressor Operator System), verwendet dafür Fusionen aus einem Fluoreszenzprotein und einem Repressorprotein, die spezifisch an ein Operator-Array binden, welches zahlreiche Kopien der Repressorbindestelle aufweist und in das Genom integriert wurde (Robinett et al., 1996; Straight et al., 1996). Mit Hilfe dieser Methode konnten bereits die Lokalisation, Replikation und Segregation von Chromosomen in verschiedenen Organismen visuell verfolgt werden (Lau et al., 2003; Sonntag Brown et al., 2011; Thompson and Compton, 2010). Nachteil dieser Methode ist jedoch, dass alle Fluoreszenzproteine, DNA-gebunden oder nicht, detektiert werden und so ein starkes Hintergrundsignal ergeben.

Das Manuskript präsentiert eine neue Einzelzell-Methode, die dieses Hintergrundsignal minimiert. BiFCROS (Bimolecular Fluorescence complementation and Repressor Operator System) basiert auf zwei verschiedenen Operatorsequenzen, die sich abwechselnd wiederholen. Dies wird als hybrid FROS-Array bezeichnet. Die zugehörigen Repressorproteine werden jeweils mit einem nicht fluoreszierenden Fragment eines Fluoreszenzproteins fusioniert. Binden die Repressoren an die Operatoren, gelangen die nicht-fluoreszierenden Fragmente in räumliche Nähe und es kommt zur Rekonstitution des funktionellen Fluoreszenzproteins, der Bimolekularen Fluoreszenzkomplementation (BiFC). Fluoreszenzsignale können demnach nur entstehen, wenn die Repressoren an das hybrid FROS-Array binden. Nichtgebundene Proteine weisen keine Fluoreszenz auf, wodurch das Hintergrundsignal minimiert wird. Für die Quantifizierung von Fluoreszenzsignalen bietet es den Vorteil, dass die Fluoreszenzintensität der gesamten Zelle bestimmt werden kann. Diese sollte proportional zur Kopienzahl des hybrid FROS-Arrays ansteigen, sodass eine Kopienzahlbestimmung des Operator-Arrays möglich sein sollte. Für die Etablierung von BiFCROS wurde ein Operator-Array bestehend aus der Bindestelle OL1, welche von dem Repressorprotein *ci* des Phagen λ gebunden wird und der Bindestelle UAS (Upstream Activating Sequences) des Modulatorproteins Gal4 aus der Hefe, hergestellt. Dieses OL1/UAS hybrid FROS-Array wurde in das Genom von *E. coli* integriert. Die genannten Proteine wurden zum einen mit dem Volllängen Fluoreszenzprotein mVenus oder mit dem N- bzw. C-terminalen Fragment von mVenus fusioniert. Die grundlegende Funktionalität von BiFCROS wurde anhand fluoreszenzmikroskopischer Analysen bestätigt. Alle Plasmid-kodierten Fusionsproteine waren in der Lage an ihre Bindestellen innerhalb des OL1/UAS hybrid FROS-Arrays zu binden. Auch die Komplementation der mVenus-Fragmente konnte erfolgreich nachgewiesen werden. Quantitative, fluoreszenzmikroskopische Analysen mit dem Volllängen Fluoreszenzprotein mVenus, dem

fragmentierten mVenus und dem Leerplasmid in Zellen ohne das OL1/UAS hybrid FROS-Array ergaben, dass die Hintergrundfluoreszenz des neuen Systems stark reduziert ist. Analysen von Zellen mit dem OL1/UAS hybrid FROS-Array zeigten zudem, dass das Verhältnis von Signal zu Hintergrund in Zellen mit dem fragmentierten mVenus im Gegensatz zum Vollängen mVenus deutlich verbessert wird. Außerdem zeigten Zellen mit zwei Kopien des OL1/UAS hybrid FROS-Arrays und den Fusionsproteinen mit dem N- bzw. C-Terminus von mVenus eine höhere Fluoreszenzintensität als Zellen mit nur einer Kopie. BiFCROS stellt somit eine neue Einzelzell-Methode dar, die das herkömmliche FROS verbessert und die Quantifizierung von Fluoreszenzsignalen erleichtert.

Ich führte nach Absprache mit Prof. Dr. T. Waldminghaus alle experimentellen Arbeiten durch. N. Schallop assemblierte das finale OL1/UAS hybrid FROS-Array und konstruierte die Plasmide pMA182 und pMA301. Die benutzerdefinierten R-Skripte zur Analyse der fluoreszenz-mikroskopischen Daten wurden von Prof. Dr. T. Waldminghaus verfasst. Das Manuskript wurde von Prof. Dr. T. Waldminghaus und mir verfasst.

BiFCROS: A low background fluorescence repressor operator system

*Sarah Milbredt and Torsten Waldminghaus**

Chromosome Biology Group, LOEWE Center for Synthetic Microbiology, SYNMIKRO, Philipps-Universität Marburg, Hans-Meerwein-Str. 6, D-35043 Marburg, Germany

Abstract

Fluorescence based methods are widely used to analyze elementary cell processes such as DNA replication or chromosomal folding and segregation. Labeling DNA with a fluorescent protein allows the visualization of its temporal and spatial organization. FROS (Fluorescence Repressor Operator System) is based on this approach. This method specifically labels DNA *in vivo* through binding of a fusion of a fluorescent protein and a repressor protein to an operator array, which contains numerous copies of the repressor binding site and has been integrated into the genome. Even though this and other fluorescence based methods are widely used, they are subjected to background fluorescence, which can reduce the measurement reading. Here, we present a new method that greatly reduces the background signal of FROS. BiFCROS (Bimolecular Fluorescence complementation and Repressor Operator System) is based on fusions of repressor proteins to a split fluorescent protein, which then bind to a hybrid FROS array, resulting in fluorescence signals due to bimolecular fluorescence complementation. Only proteins bound to the hybrid FROS array fluoresce. Fluorescence microscopic analyses of BiFCROS show reduced background fluorescence and an improved signal to noise ratio compared to FROS.

Introduction

The ability to visualize more and more details within organisms and cells has been a main driving force of biological science for many years starting with early microscopy studies of Anton van Leeuwenhoek. More recently, the discovery and application of fluorescence proteins as the small green fluorescence protein (GFP) have revolutionized life science and were consequently honored with the Nobel Prize to Osamu Shimomura, Martin Chalfie und Roger Tsien in 2008. Through genetic manipulation, the fluorescence can be added to individual proteins and then be visualized by fluorescence microscopy or quantified by flow cytometry or other methods (Chalfie et al., 1994; Wei and Dai, 2014). While this approach is straight forward for the analysis of proteins other cellular

components are less easy to visualize, including DNA or lipids. Especially the labelling of DNA is of high importance since processes as DNA replication as well as chromosome segregation and folding are central processes of all living cells (Joshi et al., 2011; Williams et al., 2010; Zickler and Kleckner, 1999). Different methods have been developed to fluorescently label the entire DNA of a cell by for example using thymidine-analog incorporation or labelling of proteins that show uniform binding to DNA (Salic and Mitchison, 2008; Wery et al., 2001). However, such approaches do not allow visualization of specific genetic loci. One approach to do so is Fluorescence In Situ Hybridization (FISH), a method that uses sequence complementarity to target a genetic locus by hybridization with a fluorescence probe (DeLong et al., 1989; Scherthan and Loidl, 2010). The disadvantage of FISH is that cells have to be fixed and therefore analysis of dynamics within living cells is not possible. One popular way to overcome this limitation is the use of so called Fluorescence Repressor Operator System (FROS)(Lau et al., 2003; Robinett et al., 1996). It is based on fusions between a fluorescent protein and a repressor that binds specifically to a respective operator sequence. An array of several such operator sequences is inserted at the locus of interest. This allows visual tracking of genetic loci by fluorescence microscopy because multiple labeled repressor proteins bound to the operator array appear as fluorescence focus. FROS was initially applied with a LacI-GFP fusion that bound to an array of Lac operators in CHO and yeast cells (Gordon et al., 1997; Robinett et al., 1996). Meanwhile, FROS arrays were established based on the tet-repressor protein TetR and cI from the λ -phage (Fekete and Chattoraj, 2005; Michaelis et al., 1997). FROS was successfully applied in various organisms to gain new insights into the localization, replication and segregation of chromosomes in individual cells (Lau et al., 2003; Matzke et al., 2005; Straight et al., 1996). In addition to studies on the spatial and temporal organization of genetic loci, FROS can also be used to determine copy numbers of genetic loci because the number of fluorescence foci should indicate the number of copies of the tagged locus (Jonas et al., 2011; Srivastava and Chattoraj, 2007). A problem with this approach is that genetic loci don't necessarily have to be spatially separated but could be held together as for example by sister chromosome cohesion (Lesterlin et al., 2012; Sherwood et al., 2010).

Generally fluorescence based analyses suffer from background fluorescence which reduces the true signal. Weak signals can be superimposed by background fluorescence, so that false negative results may occur. Background fluorescence is an unspecific fluorescence signal which can be caused by various factors. Media, agents like antibiotics, unbound fluorophores or fluorescence of the specimen itself can interfere with fluorescence analyses (Waters, 2009). Cells always exhibit a natural intrinsic fluorescence that is also known as auto-fluorescence. It is based on endogenous molecules which possess specific excitation and emission wavelengths which results in characteristic fluorescence spectra associated with specific species or strains (Ammor, 2007). Background

fluorescence can also be caused by experimental setup and imaging parameters, like light from the light source or camera noise (Joglekar et al., 2008). Different approaches can be used to solve the problem of high background signals. Examples are the treatment with chemicals as CuSO_4 that disrupt background but remain signal, new microscopy techniques as TIRF (total internal reflection fluorescence) microscopy or application of mathematical models (Axelrod, 2001; Schnell et al., 1999; Van de Lest et al., 1995). Also FROS suffers from background fluorescence that is caused by the freely diffusing fluorescent proteins not bound to the operator array. Fluorescence microscopic analyzes images arise with distinct foci paired with a diffuse fluorescence which is distributed throughout the cell. Conceptionally, elimination of this background should allow gene copy number analysis in individual cells by simply measuring the whole-cell fluorescence for example by flow cytometry.

Our current study describes the design and construction of a low background hybrid FROS in *E. coli*. It is based on the Bimolecular Fluorescence Complementation (BiFC) of two repressor fusions binding to a respective hybrid FROS array. This new Bimolecular Fluorescence Complementation Repressor Operator System (BiFCROS) shows significantly less background fluorescence compared to conventional FROS. We also elucidated its ability to determine genetic copy numbers by measurement of whole-cell fluorescence.

Material & Methods

Bacterial strains, plasmids, oligonucleotides, and culture conditions

All strains, replicons and oligonucleotides used in this study are listed in Table S1-S3. Bacteria were cultured in LB liquid medium under vigorous shaking or on LB plates at 37°C with the desired antibiotic if not differential indicated. Antibiotic selections were used at the following concentrations: ampicillin 100 µg/ml, chloramphenicol 12,5 µg/ml, kanamycin 35 µg/ml, tetracycline 15 µg/ml.

Fluorescence Microscopy

For fluorescence microscopy cells were grown in AB-glucose or AB-sodium acetate until $\text{OD}_{450} \sim 0.15$ (Messerschmidt et al., 2015). Production of fusions proteins was induced with 400 nM IPTG for 1h in AB-glucose and 2h in AB-sodium acetate. 1 ml of the culture was harvested by centrifugation, the supernatant was discarded and cells were resuspended in 20 µl of the same media. 1,5 µl of cells were transferred on 1% agarose pads based on respective growth media.

Fluorescence microscopy was performed with a Nikon Eclipse Ti-E microscope with a phase-contrast Plan Apo I oil objective (100; numerical aperture, 1.45) with the AHF HC filter set F36-528 YFP (excitation band pass [ex bp] 500/24-nm, beam splitter [bs] 515-nm, and emission [em] bp 535/30-nm filters) and an Argon Ion Laser (Melles Griot). Exposure times were adjusted for each sample. Images

were acquired with an Andor iXon3 885 electron-multiplying charge-coupled device (EMCCD) camera. Images were analyzed by Fiji using the MicrobeJ plugin (Jiang et al., 2014). Data was further processed with Excel and R statistics.

Preparation of the OL1/UAS hybrid FROS array

An initial DNA fragment was designed which contains an UAS and a split OL1 binding site (Fig.1 B). In between the binding sites a 10 bp random spacer sequence was integrated, in which 2 bps were determined to exclude appearance of restriction enzyme sites. *BsaI* and *BpiI* recognition sites flank binding sites for further assembly steps. For generation of the DNA fragment two oligonucleotides, 257 and 258, were synthesized and annealed. Annealing results in single stranded 5' overlaps through which the DNA fragment was ligated into an *XbaI* and *EcoRI* restriction digested pUC19 and pUC57-kan. Ligation products were transformed into *E. coli* XL1-Blue. All colonies were used for plasmid isolation. Resulting plasmid libraries were divided into two parts which (i) were restriction digested with *BpiI* and *NdeI* (vector) and (ii) were restriction digested with *BsaI* and *NdeI* (insert). pUC57 insert was ligated with pUC19 vector and *vice versa*. Ligation products were transformed in *E. coli* XL1-Blue. Plasmids with two inserts (2 initial oligos) will arise. Cycles from plasmid library isolation, restriction digestion and ligation were repeated until 64 inserts were assembled. One final clone was chosen and the construct was sequenced to verify length and accuracy of OL1 and UAS binding sites. It was named pMA164. For integration into the *E. coli* chromosome 500 bp of the upstream (primer 95 + 96) and downstream (primer 97 + 98) region of *tnaA* were amplified. Together with *SmaI* and *DraI* restriction digested pMA164 amplicons were used in a Gibson assembly (Gibson et al., 2009) that resulted in plasmid pMA165. In the next step the origin of replication of pMA165 was exchanged. A conditional origin of replication, *oriR6K*, was amplified from synVicII-0.1 with flanking *SmaI* recognition sites (primer 389 + 390). Via Gibson assembly the *oriR6K* fragment and pMA165, which was restriction digested with *SmaI*, were assembled resulting in pMA252. pMA252 was restriction digested with *SmaI* and used as template for integration into the *E. coli* chromosome via lambda red recombination using *E. coli* strain D50 (Datsenko and Wanner, 2000). Integration was first verified by colony PCR (primer 102 + 103) and then by southern blot analysis with a DIG labeled *mnmE* specific probe (409 + 410; PCR DIG probe synthesis kit (Roche)) (data not shown). Via P1 transduction the integrated array was transferred to MG1655.

Preparation of fusion proteins

All fusion constructs of λ cl and Gal4 to mVenus were done by Gibson Assembly (Gibson et al., 2009). For λ cl fusions an RSIAT (cgcagcattgcgacc) linker was integrated between λ cl and mVenus or N-terminal fragment of mVenus. Fusion constructs that include Gal4 DNA binding domain (amino acids 1-441) of *Saccharomyces cerevisiae* contain no linker. BiFC constructs contain a random linker sequence of 21

bp between λ cl fused to N-terminal fragment of mVenus (λ cl -VN) and Gal4 fused to the C-terminal fragment of mVenus (Gal4-VC). As N-terminal fragment of mVenus amino acids 1-154 were used and amino acids 155-238 as C-terminal fragment. For improved signal to noise ratio mutation I152L was introduced in N-terminal fragment of mVenus by PCR (Kodama and Hu, 2010). Resulting plasmids were verified by restriction digestion and sequenced for verification of coding sequences. A detailed description of the preparation of different DNA-fragments for Gibson assembly is listed in Table 1.

Table 1: Instructions for fusion constructs.

	λ cl	Gal4	Full length mVenus	N-terminal fragment mVenus	C-terminal fragment mVenus	Acceptor vector
pMA275 (λcl - mVenus)	646 + 656 (pPS19)	-	-	268 + 47 (pBad24-Lacl-venus)	46 + 656 (pBad24-Lacl-venus)	pWM2060 <i>Sma</i> I
pMA276 (Gal4- mVenus)	-	809 + 712 (pGBKT7)	714 + 656 (pMA275)	-	-	pWM2060 <i>Sma</i> I
pMA270 λcl -VN + Gal4-VC)	646 + 267 (pPS19)	270 + 758 (pGBKT7)	-	268 + 769 (pBad24-Lacl-venus)	757 + 656 (pMA275)	pWM2060 <i>Sma</i> I

Primer and template (in brackets) are indicated as well as the acceptor vector. All fragments contain sequential overlapping regions for Gibson assembly.

Generation of additional replicons with and without OL1/UAS hybrid FROS array

The ampicillin cassette of the synthetic chromosome synVicII-1.3 was exchanged with a cat cassette (primer 655 + 657) by homologous recombination in yeast (Messerschmidt 2014). The resulting replicon was named pMA182. For insertion of the OL1/UAS hybrid FROS array into pMA182 it was amplified with primers 1045 + 1046 and pMA182 was restriction digested with *I-SceI* and *NheI*. Both products were used in a Gibson reaction to assemble plasmid pMA310. pMA182 naturally contains a *gfp* which would interfere with mVenus fluorescence signals. *gfp* was removed by restriction digestion of pMA182 with *I-SceI* and *NheI* and ligated with annealed primers 1259 + 1260 resulting in pMA310.

Data Analysis with MicrobeJ and R statistics

The plugin MicrobeJ for Fiji was used to detect raw pixel intensities and intensities of whole cells. Data was imported into Excel and if necessary background signal was distracted. R custom scripts were used for further analysis. Data was imported to Excel and plotted.

Results

A Bimolecular Fluorescence Complementation Repressor Operator System (BiFCROS)

The background fluorescence in FROS systems is caused by the non DNA bound fluorescence repressor fusions. Consequently, it should be possible to reduce the background signal by rendering the freely diffusing fluorescence proteins. To this end we made use of the ability of split fluorescence proteins to complement one another if in close spatial proximity (Kerppola, 2008). The general logic of our new approach is illustrated in figure 1A.

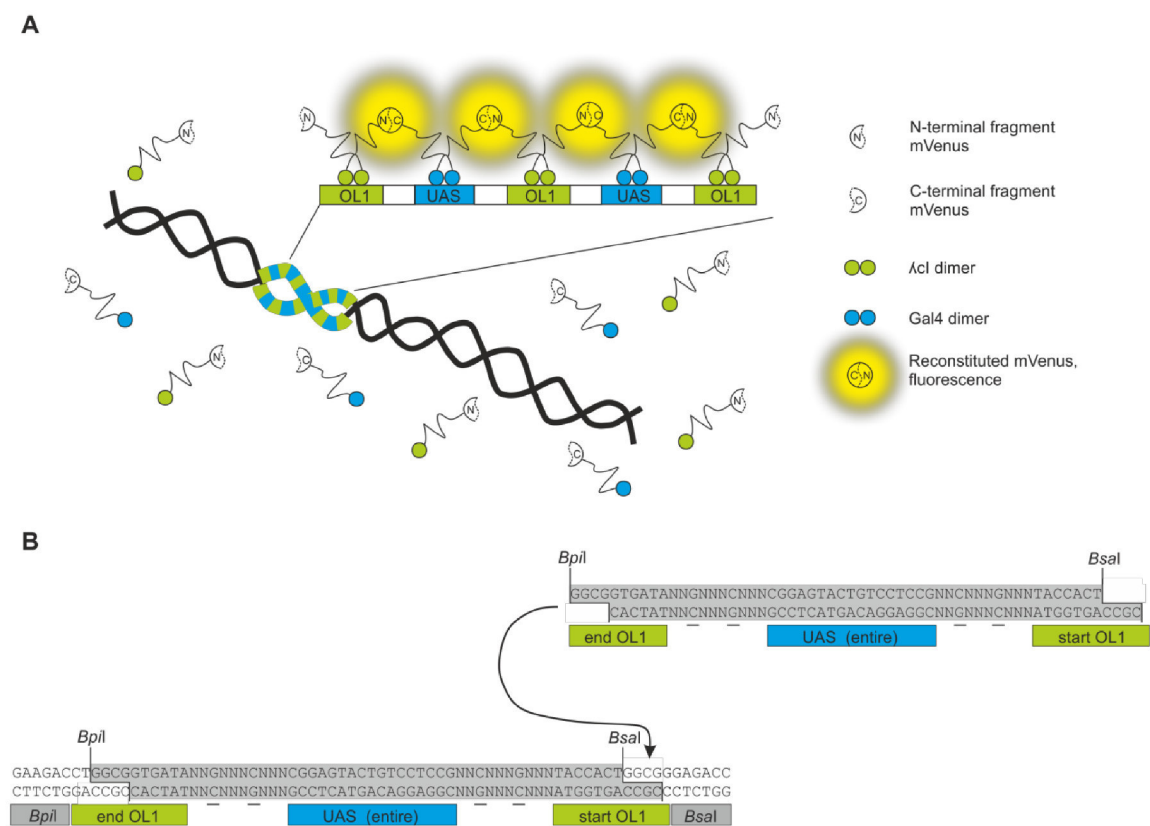


Fig. 1: Scheme of BiFCROS, an improved fluorescence operator system. **A** An array with alternating binding sites OL1 and UAS is integrated into the *E. coli* chromosome. Respective proteins λcl and Gal4 are fused to either the N- or C-terminal fragment of a fluorescence reporter (mVenus). Through binding of the proteins to the array, terminal fragments of mVenus come in close proximity and reconstitute resulting in a fluorescence signal. Proteins that are not bound do not fluoresce and cause no background signal. **B** Initial building blocks of the Operator array contain a split OL1 site, an entire UAS and recognition sequences for type II restriction enzymes *Bpil* and *Bsal*. Random spacer sequences between binding sites contain 2 determined base pairs (underlined). Restriction digestion with *Bpil* and *Bsal* results in DNA fragments with overhangs fitting to each other (highlighted in gray). Ligation of two DNA fragments leads to reconstitution of an entire OL1 site.

Instead of one transcriptional repressor protein fused to a full-length fluorescence protein the new system consists of two repressor proteins (C1 and Gal4), each fused to one of half of the fluorescence protein mVenus. These protein fusions should not generate any fluorescence signal unless they come in close spatial proximity. Such a signal would be desirable at a genetic locus of interest. To this end we constructed an array of operator sites with alternating binding sites for C1 (OL1) and Gal4 (UAS) (Fig. 1A). The design includes 64 instances of each operator type and was constructed with a seven-step hierarchical cloning scheme based on libraries of synthetic DNA oligos and *typIIS* restriction enzymes (Fig. 1B; see Material and Methods section for details).

Low background signals with BiFCROS

As a first test of array functionality we constructed an inducible system for expression of a full-length mVenus to the λ cl repressor. Expression of this fusion-protein in *E. coli* cells without the constructed hybrid FROS array resulted in diffuse fluorescence as expected (Fig. 2A). In contrast, clear fluorescence foci were visible when cells contained the OL1/UAS hybrid FROS array showing that the array in fact binds the protein fusion (Fig. 2B).

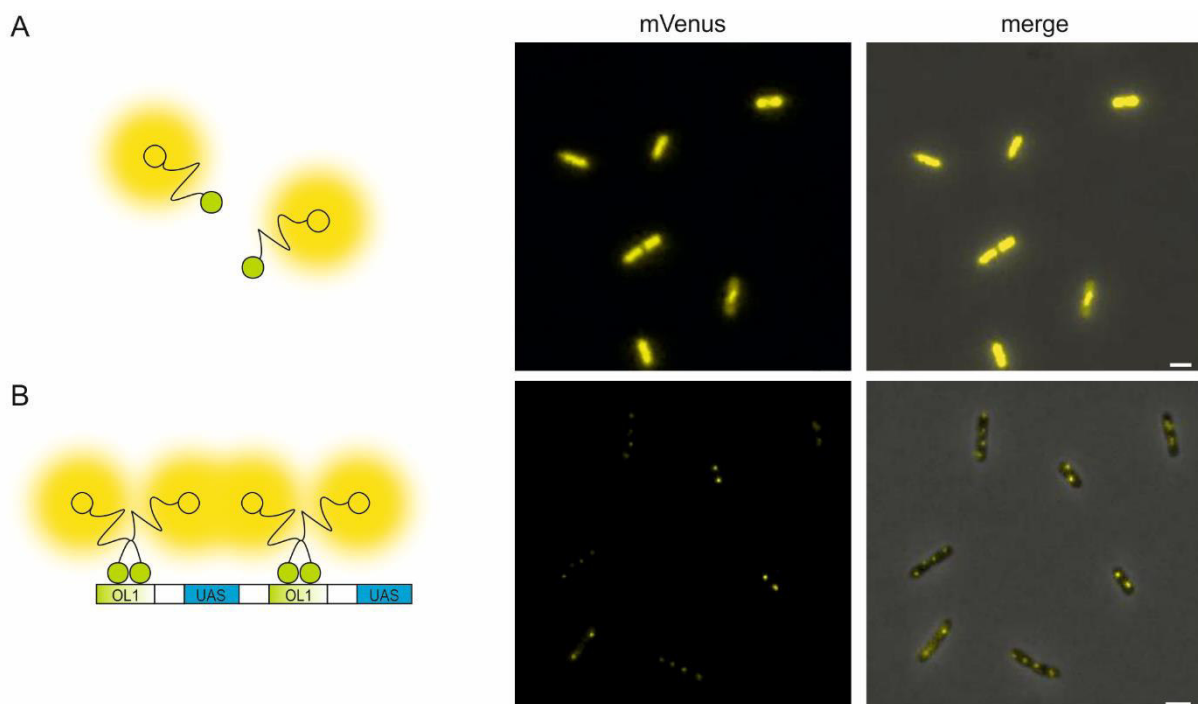


Fig. 2: Specific binding to the OL1/UAS hybrid FROS array. Fluorescence microscopy images of cells without (A, upper panels, strain SM69) and with (B, lower panels, strain SM77) OL1/UAS hybrid FROS array expressing a fusion of the λ cl repressor to full length mVenus. Left panel presents cartoons illustrating fluorescent fusion proteins without OL1/UAS hybrid FROS (A) or bound to the OL1/UAS hybrid FROS array (B). Middle panel shows fluorescence microscopic images of the mVenus channel. Right panel shows merged microscopy images of mVenus and phase contrast channel. Exposure time was set to 100 ms. Scale bar represents 2 μ m.

To test the BiFCROS principle outlined above, we fused the λ cl repressor to the N-terminal half of mVenus and Gal4 to the C-terminal part within an inducible bicistronic operon. *E. coli* cells expressing these fusion proteins showed very low fluorescence and no distinct foci (Fig. 3 A). In contrast, clear fluorescence foci were detected in cells with a chromosomal insertion of the OL1/UAS hybrid FROS array (Fig. 3 B). These results indicate that the hybrid FROS array is in fact able to mediate fluorescence complementation and limits the fluorescence signal to the targeted genetic locus.

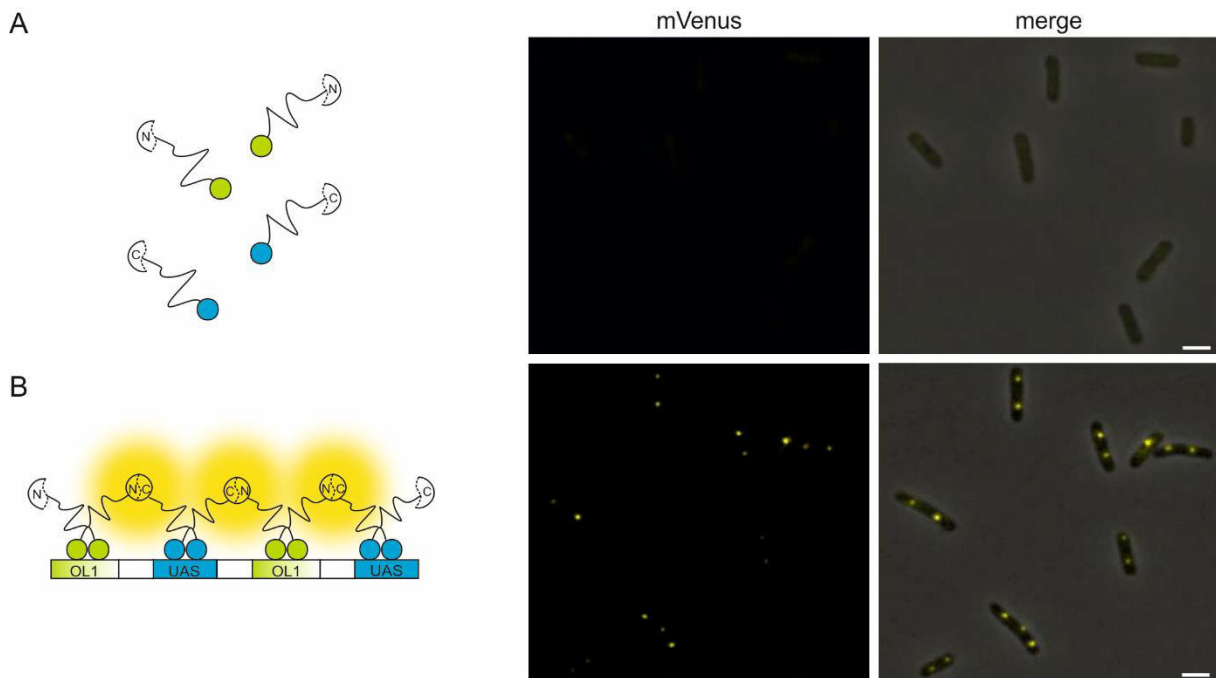


Fig. 3: Bimolecular complementation of split mVenus mediated by binding of fusion proteins to the OL1/UAS hybrid FROS array. Fluorescence microscopy images of cells that express BiFC constructs λ cl fused to the N-terminal fragment of mVenus and Gal4 fused to the C-terminus of mVenus without (A, upper panel, strain SM57) and with OL1/UAS hybrid FROS array (B, lower panel, strain SM65). Left panel presents cartoons illustrating non fluorescent fusion proteins without OL1/UAS hybrid FROS (A) and bound fusion proteins to the OL1/UAS hybrid FROS array (B) resulting in fluorescence signals. Middle panel shows fluorescence microscopic images of the mVenus channel. Right panel shows merged microscopy images of mVenus and phase contrast channel. Exposure time was set to 200 ms. Scale bar represents 2 μ m.

To get a more quantitative insight into background intensities we quantified the fluorescence per cell for *E. coli* strains with (i) the BiFC constructs with λ cl and Gal4 each fused to one part of the split mVenus (strain SM57); (ii) λ cl fused to the full-length mVenus (strain SM69); (iii) Gal4 fused to a full-length mVenus and (iv) an empty vector control (Fig. 4). Cells were grouped according their size and average values calculated. The analysis shows that splitting of mVenus greatly decreases the total cellular fluorescence to almost one order of magnitude compared to the full-length fusion constructs (Fig. 4). The fluorescence of the split fusions was only slightly higher than the negative control (empty vector). More important than the general reduction of fluorescence background by the fluorophore-

splitting approach is an increase of the signal to background ratio. Considering pixels of individual cells one would expect many pixels of very low intensity and a small fraction of pixels with high intensities corresponding to the fluorescence foci. Pixels with intermediate intensities should be rare in a low-background system. To measure these values for BiFCROS, we performed quantitative fluorescence microscopy of cells with the hybrid FROS array and either the split fusions or a fusion of λ cl to mVenus (Fig. 5). Cells were grown in different growth media resulting in either fast (AB-glucose) or slow growth (AB-sodium acetate). Fluorescence intensity values for individual pixel of some hundred cells were calculated and plotted as histogram (Fig. 5 A, B, D, E). The shapes of the histograms are clearly different. In cells with BiFCROS, most pixels belong to the lowest intensity fraction with a sharp drop in pixel numbers belonging to fractions with higher intensities (Fig. 5 B and E).

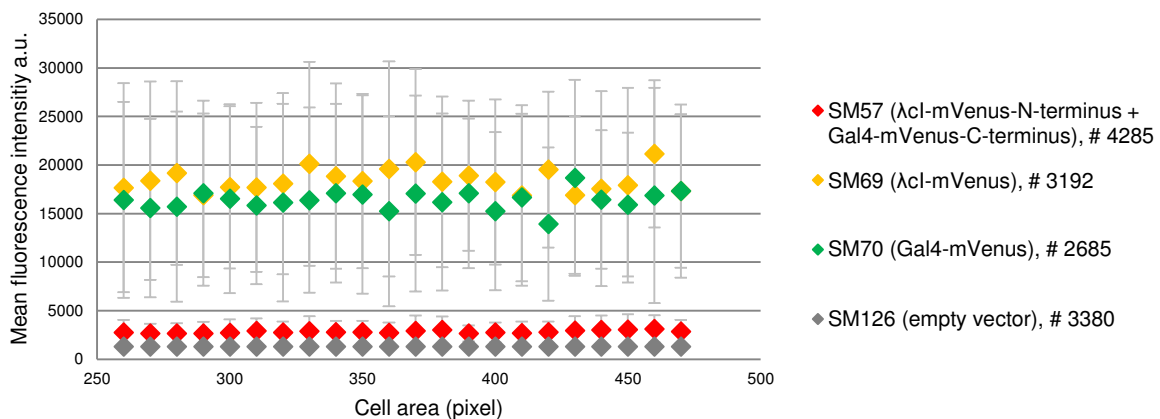


Fig. 4: Fluorescence intensities of cells without OL1/UAS hybrid FROS array with full length mVenus, split mVenus or empty vector. Mean fluorescence intensities of cells versus cell area is plotted as indicated. Strain SM57 contains λ cl fused to the N-terminal fragment of mVenus and Gal4 fused to the C-terminus of mVenus, strain SM69 contains λ cl fused to full length mVenus, strain SM70 contains Gal4 fused to full length mVenus and strain SM126 contains the empty vector. Fluorescence microscopy was performed with cells grown exponentially in AB-glucose. Exposure time was set to 100 ms. Number of recorded cells is indicated by #. Quantification of fluorescence signals was done by MicrobeJ and data analyses by R. Cells were grouped according to their cell area with moving windows of 10 pixels. Groups with ten or more cells were considered for analysis. Average values of two independent experiments are shown.

In cells with the full-length mVenus fusion, comprising the conventional FROS system, this drop is much less pronounced with many pixels comprising an intermediate fluorescence intensity (Fig. 5 A and D). These results indicate that the distribution of background to signal is much clearer with BiFCROS compared to conventional FROS. To analyze the ratio of signal to background for individual cells we defined the signal as 10% of the pixels with the highest fluorescence intensity and the background as the 50% with the lowest signal. Respective mean values of signal and background were plotted versus one another for individual cells (Fig. 5 C and F). Biological replicates of the analysis showed good reproducibility (Fig. S1).

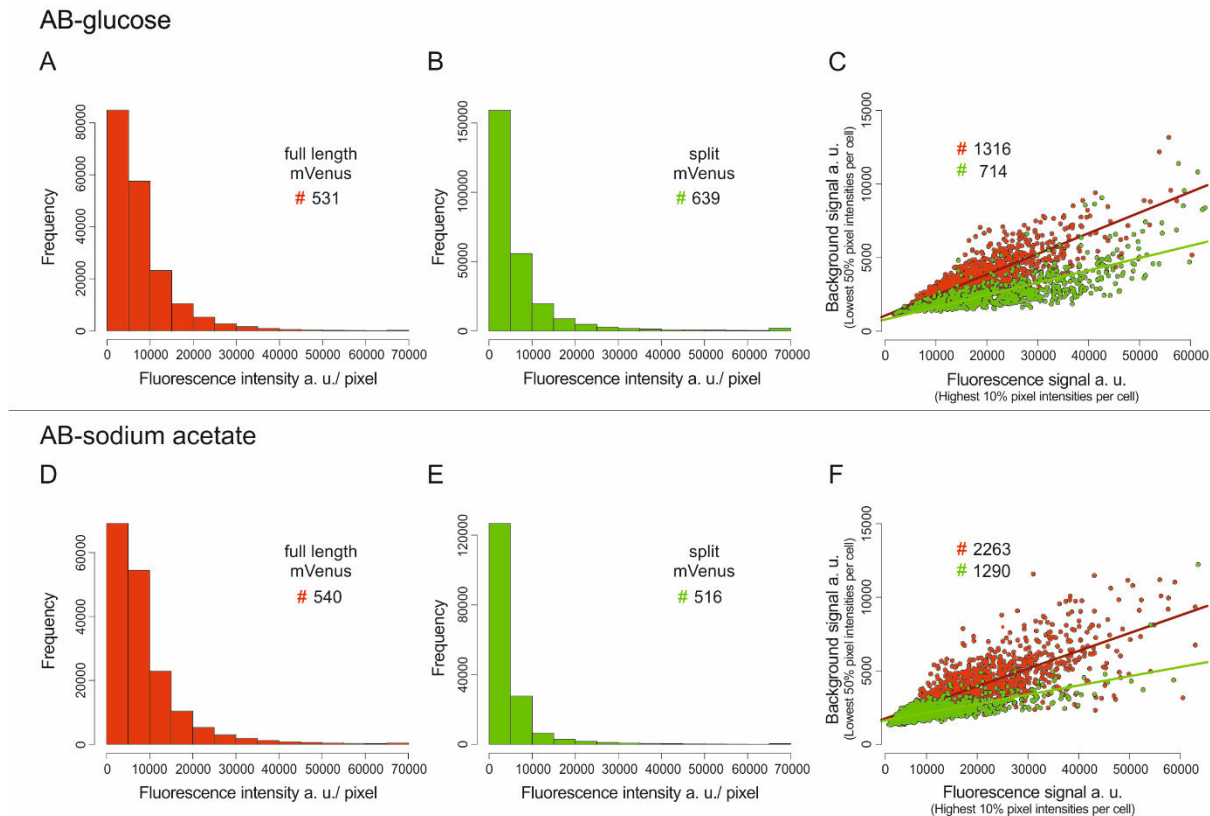


Fig. 5: Increased signal to background with split mVenus compared to full length mVenus. Cells with OL1/UAS hybrid FROS array and λ cl fused to full length mVenus (highlighted in red, strain SM77) or BiFC constructs λ cl fused to the N-terminal fragment of mVenus and Gal4 fused to the C-terminal fragment of mVenus (highlighted in green, strain SM65) were analyzed. **A, B, D, E** Frequency distribution of all pixel fluorescence intensities. Only cells with intensities below 65000 and > 5-fold signal to background ratio were analyzed (mean of 10% pixels with highest (signal)/ mean of 50% pixels with lowest intensity (background)). Left panel presents analysis with OL1/UAS hybrid FROS array and λ cl fused to full length mVenus. Middle panel shows analysis of OL1/UAS hybrid FROS array and λ cl fused to the N-terminal fragment of mVenus and Gal4 fused to the C-terminal fragment of mVenus. **C, F** Scatter plots of mean fluorescence intensity of the highest 10% pixel versus mean fluorescence intensity of the lowest 50% pixel within individual cells. Values above 65000 on x-axis and above 15000 on y-axis were removed. Top panels: fluorescence microscopy was performed in AB-glucose medium with exposure times of 50 ms for full length mVenus and 200 ms for split mVenus. Lower panels: fluorescence microscopy was performed in AB-sodium acetate medium with exposure times of 50 ms for full length mVenus and 600 ms for split mVenus. Quantification of fluorescence signals was done by MicrobeJ and data analyses by R with custom R scripts. Cell numbers of each plot are indicated by color coded #.

Towards copy number determination with BiFCROS

The copy number can provide information about the replication status of a genetic locus. If a locus is present in single copy, it has not yet been replicated. A duplicate copy number indicates the completed DNA replication. The design of BiFCROS should allow simplified copy number determination of respective gene loci. By the use of the split mVenus, only proteins bound to the array fluoresce. Free

proteins give no signal. Thus, the number of copies of the OL1/UAS hybrid FROS array should be proportional to total fluorescence intensity of the cell. Determination of the fluorescence intensity of the whole cell can thus be used to conclude on the number of copies of a genetic locus with inserted BiFCROS array. To verify this assumption, an experimental set up was generated with a strain that contains one copy (SM65) and a strain which contains two copies of the OL1/UAS hybrid FROS array (SM153). The second copy of the array was embedded on an additional synthetic, secondary chromosome pMA301. Fluorescence microscopy measurements of the two strains revealed significantly increased fluorescence intensity for the strain with two copies of the array compared to the intensity of the strain with only one copy (Fig. 6).

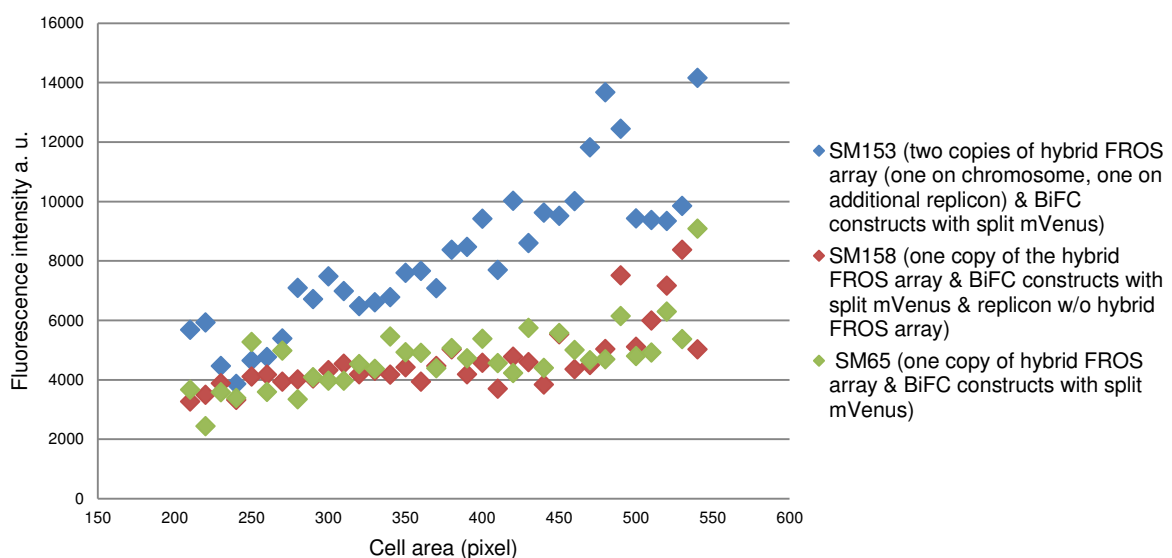


Fig. 6: Fluorescence signals of cells with one or two copies of the OL1/UAS hybrid FROS array. Mean fluorescence intensities of cells versus cell area are plotted as indicated. Strain SM153 contains a replicon (pMA301) with an additional copy of the OL1/UAS hybrid FROS array, strain SM153 and strain SM65 contain only one copy of the OL1/UAS hybrid FROS array. SM158 additionally contains a replicon (pMA310) without an additional copy of the OL1/UAS hybrid FROS array as a control. All strains contain BiFC constructs λ cl fused to the N-terminal fragment of mVenus and Gal4 fused to the C-terminal fragment of mVenus. Fluorescence microscopy was performed with cells grown in AB-sodium acetate medium. Exposure time was set to 600 ms. Approximately 1000 cells per strain were recorded. Quantification of fluorescence signals was done by MicrobeJ and data analyses by R. Mean intensity of cells of strain SM57 that contains BiFC constructs with split mVenus without OL1/UAS hybrid FROS array was determined as background signal and subtracted from cell intensities of strain SM153, SM158 and SM65. Cells were grouped according to their cell area with moving windows of 10 pixels. Groups with ten or more cells were considered for analysis. Average values of three independent experiments are shown. Standard deviation was not plotted for clarity reasons.

To rule out that the synthetic chromosome contributes to the fluorescence in any way, a control strain was generated that contains the array on the native chromosome and additionally carries a synthetic,

secondary chromosome without a BiFCROS array (SM158). The fluorescence intensities of this strain are comparable to the strain which carries no additional replicon (Fig. 6).

Discussion

Observation of spatial and temporal organization of the DNA within the cell is important to gain insights into cellular processes like DNA replication or chromosome segregation. In any measuring system an optimal signal to background ratio is desirable. Although FROS is a commonly used application to highlight DNA *in vivo*, no attempt was made to reduce its background fluorescence. Here, we present BiFCROS, a new single-cell method which has a strongly reduced background signal and thus improves signal to background ratio. It is based on bimolecular fluorescence complementation (BiFC), the ability of a fluorescent protein to be fragmented and complement if in close proximity (Fig. 1). Another method to label DNA, based on BiFC, has been established. Stains and coworker developed a tool for direct targeting of double-stranded DNA sequences, which is called SEER (sequence enabled reassembly of proteins). They used a split GFP fused to two different Zinc finger domains, that reconstitute upon binding to designated DNA *in vitro* (Stains et al., 2005). A method to visualize and track RNA in living cells was first invented in eukaryotes, where RNA is more readily accessible than DNA, which is located in the nucleus. The method is based on fluorescent-protein complementation regulated by the interaction of a split RNA-binding protein with its corresponding RNA aptamer (Valencia-Burton et al., 2007). The system was successfully transferred to *E. coli*, to study temporal and spatial organization of RNA (Valencia-Burton et al., 2009). Both systems also report on less fluorescence background signals compared to analysis with full length fluorescent proteins. mVenus was chosen for BiFCROS, because it has fast and efficient maturation time, is less sensitive to the environment and Venus-based BiFC showed 10-fold higher fluorescence intensity in comparison with signals from eg. EYFP-BiFC (Nagai et al., 2002; Shyu et al., 2006). mVenus fusions were constructed with a monomeric variant to prevent dimer formation and to exclude the resulting fluorescence signal from the latter. For further improvement of the signal to noise ratio we inserted the mutation I152L into the N-terminus of mVenus. The amino acid exchange I152L reduces spontaneous self-assembly of the two non-fluorescent C and N-terminal fragments of mVenus (Kodama and Hu, 2010). It increases the signal to noise ratio 4-fold in BiFC assays in comparison to other mVenus variants (Kodama and Hu, 2010). Experimental data of comparative analyzes of fragmented and full-length fusion proteins in different media reveal that the background in BiFCROS is significantly reduced in contrast to the conventional FROS (Fig. 5, Fig. S1). Thus, BiFCROS can also detect weak signals, which might be covered by the background fluorescence of conventional systems. This is a great advancement because fluorescence-based methods generally carry an intrinsic

background signal due to autofluorescence of the sample or media. Analyses of BiFCROS were performed using fluorescence microscopy. Subsequently data evaluation was rendered using MicrobeJ. In order to increase the attractiveness of BiFCROS, cells could be examined by flow cytometry. Thus, BiFCROS would become a high throughput analysis process. However, we observed less sensitivity of weak fluorescence signals in flow cytometry than in microscopy whereby BiFC signals might have to be amplified. It is known that fluorescence proteins GFP and YFP emit in wavelength regions in which the autofluorescence is very high (Telford et al., 2012). Alternatively, BiFCROS could be performed using another BiFC-compatible fluorescence protein that emits in the far red range like dsRed and the improved derivatives e.g mRFP1 (Knop et al., 2002). mRFP1 has already been applied successfully in BiFC analyzes (Jach et al., 2006). Besides a low background, BiFCROS should in principle allow copy number determination of certain gene loci. Classical methods to determine the copy number variation of genes are microarray-CGH, FISH, qPCR, or sequencing (de Bourcy et al., 2014; Dryselius et al., 2008; Hill et al., 2012; Scherthan and Loidl, 2010). All methods have in common that they are not carried out *in vivo*. BiFCROS could help remedy this limitation because it allows the determination of the number of copies of genetic region in living cells. In order to analyze whether, BiFCROS can be applied to determine copy number differences of genetic loci, an additional copy of the OL1/UAS hybrid FROS array was located on a synthetic, secondary chromosome in *E. coli*. Fluorescence intensities of cells with one copy of the OL1/UAS hybrid FROS array and two copies clearly differ from each other. A linear increase of fluorescence signal intensity was obtained in both strains (Fig. 6). Theoretical expectations for copy number analysis of an operator array that is integrated into the chromosome, suggests a doubling of fluorescence intensity after replication. If two copies of the array are present in the cell, they can also bind and complement twice as many proteins. Thus, at a certain cell mass doubling of the fluorescence signal should be detectable. Our data shows no significant doubling at a certain cell mass. Hence, further analyzes are necessary to investigate why BiFCROS fluorescence intensities increases linearly depending on cell mass.

Acknowledgments

We thank all members of the Waldminghaus lab, especially Nadine Schallopp for excellent technical assistance. We grateful thank W. Margolin, A. Becker, and D. Chatteraj for providing strains and/or plasmids and we would like to thank A. Becker for providing the fluorescence microscope and related devices.

References

- Ammor, M.S., 2007. Recent advances in the use of intrinsic fluorescence for bacterial identification and characterization. *J Fluoresc.* 17, 455-9.
- Axelrod, D., 2001. Total internal reflection fluorescence microscopy in cell biology. *Traffic.* 2, 764-74.
- Bernard, C.S., Sadasivam, M., Shiomi, D. and Margolin, W., 2007. An altered FtsA can compensate for the loss of essential cell division protein FtsN in *Escherichia coli*. *Mol Microbiol.* 64, 1289-305.
- Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y., 1997. The complete genome sequence of *Escherichia coli* K-12. *Science.* 277, 1453-62.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. and Prasher, D.C., 1994. Green fluorescent protein as a marker for gene expression. *Science.* 263, 802-5.
- Datsenko, K.A. and Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A.* 97, 6640-5.
- de Bourcy, C.F., De Vlaminck, I., Kanbar, J.N., Wang, J., Gawad, C. and Quake, S.R., 2014. A quantitative comparison of single-cell whole genome amplification methods. *PLoS One.* 9, e105585.
- DeLong, E.F., Wickham, G.S. and Pace, N.R., 1989. Phylogenetic stains: ribosomal RNA-based probes for the identification of single cells. *Science.* 243, 1360-3.
- Dryselius, R., Izutsu, K., Honda, T. and Iida, T., 2008. Differential replication dynamics for large and small *Vibrio* chromosomes affect gene dosage, expression and location. *BMC Genomics.* 9, 559.
- Fekete, R.A. and Chattoraj, D.K., 2005. A cis-acting sequence involved in chromosome segregation in *Escherichia coli*. *Mol Microbiol.* 55, 175-83.
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd and Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* 6, 343-5.
- Gordon, G.S., Sitnikov, D., Webb, C.D., Teleman, A., Straight, A., Losick, R., Murray, A.W. and Wright, A., 1997. Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell.* 90, 1113-21.
- Hammar, P., Wallden, M., Fange, D., Persson, F., Baltekin, O., Ullman, G., Leroy, P. and Elf, J., 2014. Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation. *Nat Genet.* 46, 405-8.
- Hill, N.S., Kadoya, R., Chattoraj, D.K. and Levin, P.A., 2012. Cell size and the initiation of DNA replication in bacteria. *PLoS Genet.* 8, e1002549.
- Jach, G., Pesch, M., Richter, K., Frings, S. and Uhrig, J.F., 2006. An improved mRFP1 adds red to bimolecular fluorescence complementation. *Nat Methods.* 3, 597-600.
- James, P., Halladay, J. and Craig, E.A., 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics.* 144, 1425-36.
- Jiang, C., Brown, P.J., Ducret, A. and Brun, Y.V., 2014. Sequential evolution of bacterial morphology by co-option of a developmental regulator. *Nature.* 506, 489-93.
- Joglekar, A.P., Salmon, E.D. and Bloom, K.S., 2008. Counting kinetochore protein numbers in budding yeast using genetically encoded fluorescent proteins. *Methods Cell Biol.* 85, 127-51.
- Jonas, K., Chen, Y.E. and Laub, M.T., 2011. Modularity of the bacterial cell cycle enables independent spatial and temporal control of DNA replication. *Curr Biol.* 21, 1092-101.
- Joshi, M.C., Bourniquel, A., Fisher, J., Ho, B.T., Magnan, D., Kleckner, N. and Bates, D., 2011. *Escherichia coli* sister chromosome separation includes an abrupt global transition with concomitant release of late-splitting intersister snaps. *Proc Natl Acad Sci U S A.* 108, 2765-70.
- Kerppola, T.K., 2008. Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu Rev Biophys.* 37, 465-87.
- Knop, M., Barr, F., Riedel, C.G., Heckel, T. and Reichel, C., 2002. Improved version of the red fluorescent protein (drFP583/DsRed/RFP). *Biotechniques.* 33, 592, 594, 596-8 passim.
- Kodama, Y. and Hu, C.D., 2010. An improved bimolecular fluorescence complementation assay with a high signal-to-noise ratio. *Biotechniques.* 49, 793-805.

- Lau, I.F., Filipe, S.R., Soballe, B., Okstad, O.A., Barre, F.X. and Sherratt, D.J., 2003. Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol.* 49, 731-43.
- Lesterlin, C., Gigant, E., Boccard, F. and Espeli, O., 2012. Sister chromatid interactions in bacteria revealed by a site-specific recombination assay. *EMBO J.* 31, 3468-79.
- Matzke, A.J., Huettel, B., van der Winden, J. and Matzke, M., 2005. Use of two-color fluorescence-tagged transgenes to study interphase chromosomes in living plants. *Plant Physiol.* 139, 1586-96.
- Messerschmidt, S.J., Kemter, F.S., Schindler, D. and Waldminghaus, T., 2015. Synthetic secondary chromosomes in *Escherichia coli* based on the replication origin of chromosome II in *Vibrio cholerae*. *Biotechnol J.* 10, 302-14.
- Michaelis, C., Ciosk, R. and Nasmyth, K., 1997. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell.* 91, 35-45.
- Nagai, T., Ibata, K., Park, E.S., Kubota, M., Mikoshiba, K. and Miyawaki, A., 2002. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol.* 20, 87-90.
- Robinett, C.C., Straight, A., Li, G., Wilhelm, C., Sudlow, G., Murray, A. and Belmont, A.S., 1996. *In vivo* localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol.* 135, 1685-700.
- Salic, A. and Mitchison, T.J., 2008. A chemical method for fast and sensitive detection of DNA synthesis *in vivo*. *Proc Natl Acad Sci U S A.* 105, 2415-20.
- Scherthan, H. and Loidl, J., 2010. FISH as a tool to investigate chromosome behavior in budding yeast. *Methods Mol Biol.* 659, 363-77.
- Schnell, S.A., Staines, W.A. and Wessendorf, M.W., 1999. Reduction of lipofuscin-like autofluorescence in fluorescently labeled tissue. *J Histochem Cytochem.* 47, 719-30.
- Sherwood, R., Takahashi, T.S. and Jallepalli, P.V., 2010. Sister acts: coordinating DNA replication and cohesion establishment. *Genes Dev.* 24, 2723-31.
- Shyu, Y.J., Liu, H., Deng, X. and Hu, C.D., 2006. Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions. *Biotechniques.* 40, 61-6.
- Srivastava, P. and Chatteraj, D.K., 2007. Selective chromosome amplification in *Vibrio cholerae*. *Mol Microbiol.* 66, 1016-28.
- Srivastava, P., Fekete, R.A. and Chatteraj, D.K., 2006. Segregation of the replication terminus of the two *Vibrio cholerae* chromosomes. *J Bacteriol.* 188, 1060-70.
- Stains, C.I., Porter, J.R., Ooi, A.T., Segal, D.J. and Ghosh, I., 2005. DNA sequence-enabled reassembly of the green fluorescent protein. *J Am Chem Soc.* 127, 10782-3.
- Straight, A.F., Belmont, A.S., Robinett, C.C. and Murray, A.W., 1996. GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. *Curr Biol.* 6, 1599-608.
- Telford, W.G., Hawley, T., Subach, F., Verkhusha, V. and Hawley, R.G., 2012. Flow cytometry of fluorescent proteins. *Methods.* 57, 318-30.
- Valencia-Burton, M., McCullough, R.M., Cantor, C.R. and Broude, N.E., 2007. RNA visualization in live bacterial cells using fluorescent protein complementation. *Nat Methods.* 4, 421-7.
- Valencia-Burton, M., Shah, A., Sutin, J., Borogovac, A., McCullough, R.M., Cantor, C.R., Meller, A. and Broude, N.E., 2009. Spatiotemporal patterns and transcription kinetics of induced RNA in single bacterial cells. *Proc Natl Acad Sci U S A.* 106, 16399-404.
- Van de Lest, C.H., Versteeg, E.M., Veerkamp, J.H. and Van Kuppevelt, T.H., 1995. Elimination of autofluorescence in immunofluorescence microscopy with digital image processing. *J Histochem Cytochem.* 43, 727-30.
- Waters, J.C., 2009. Accuracy and precision in quantitative fluorescence microscopy. *J Cell Biol.* 185, 1135-48.
- Wei, T. and Dai, H., 2014. Quantification of GFP signals by fluorescent microscopy and flow cytometry. *Methods Mol Biol.* 1163, 23-31.

- Wery, M., Woldringh, C.L. and Rouviere-Yaniv, J., 2001. HU-GFP and DAPI co-localize on the *Escherichia coli* nucleoid. *Biochimie*. 83, 193-200.
- Williams, A., Spilianakis, C.G. and Flavell, R.A., 2010. Interchromosomal association and gene regulation in trans. *Trends Genet*. 26, 188-97.
- Zickler, D. and Kleckner, N., 1999. Meiotic chromosomes: integrating structure and function. *Annu Rev Genet*. 33, 603-754.

Supplements

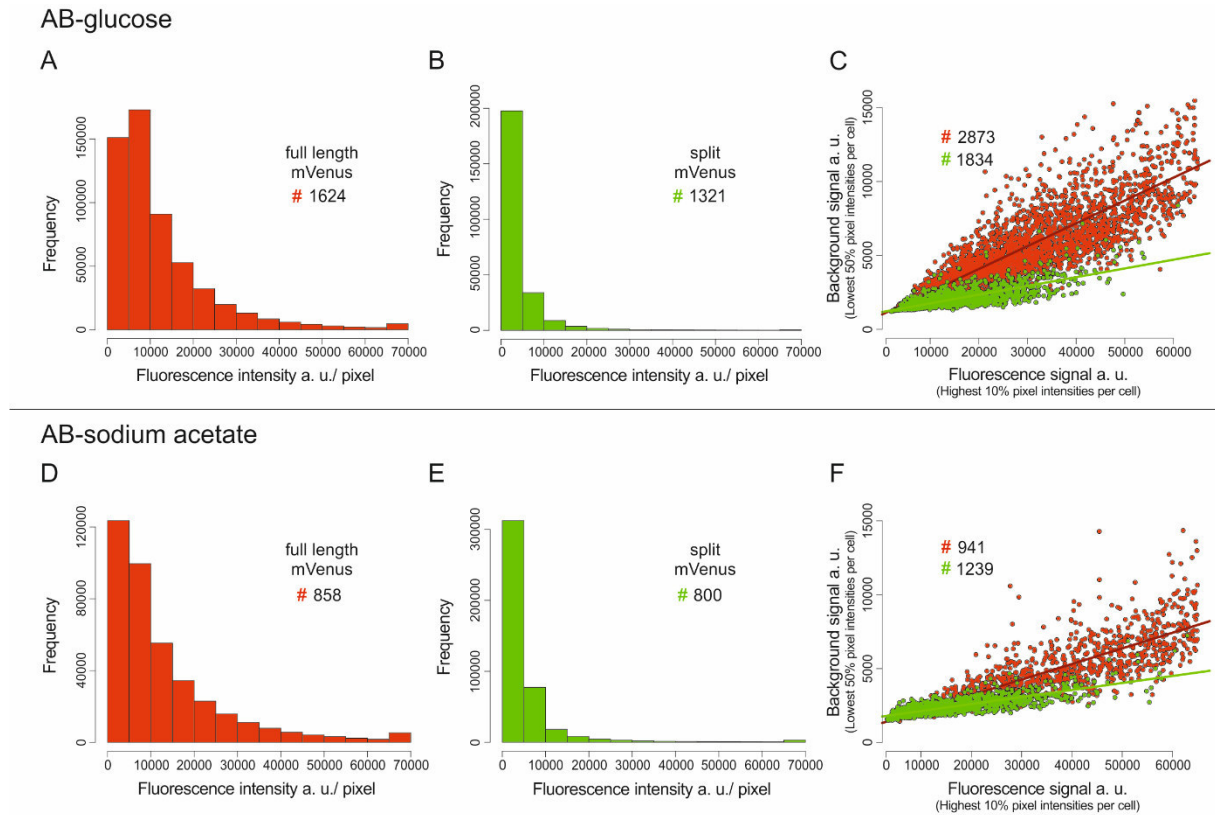


Fig. S1: Increased signal/background ratio with split mVenus compared to full length mVenus can be reproduced. Cells with OL1/UAS hybrid FROS array and λ cl fused to full length mVenus (highlighted in red, strain SM77) or BiFC constructs λ cl fused to the N-terminal fragment of mVenus and Gal4 fused to the C-terminal fragment of mVenus (highlighted in green, strain SM65) were analyzed. **A, B, D, E** Frequency distribution of all pixel fluorescence intensities. Only cells with intensities below 65000 and > 5 -fold signal to background ratio were analyzed (mean of 10% pixels with highest (signal)/ mean of 50% pixels with lowest intensity (background)). Left panel presents analysis with OL1/UAS hybrid FROS array and λ cl fused to full length mVenus. Middle panel shows analysis of OL1/UAS hybrid FROS array and λ cl fused to the N-terminal fragment of mVenus and Gal4 fused to the C-terminal fragment of mVenus. **C, F** Scatter plots of mean fluorescence intensity of the highest 10% pixel versus mean fluorescence intensity of the lowest 50% pixel within individual cells. Values above 65000 on x-axis and above 15000 on y-axis were removed. Top panels: fluorescence microscopy was performed in AB-glucose medium with exposure times of 50 ms for full length mVenus and 200 ms for split mVenus. Lower panels: fluorescence microscopy was performed in AB-Sodium acetate medium with exposure times of 50 ms for full length mVenus and 600 ms for split mVenus. Quantification of fluorescence signals was done by MicrobeJ and data analyses by R with custom R scripts. Cell numbers of each plot are indicated by color coded #.

Table S1: Strains used in this study.

Strain	Relevant genotype and additional characteristics	Resistance	Reference
<i>E. coli</i> XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F proAB lacIqZΔM15 Tn10 (TetR)]</i>	tetracycline, nalidixic acid	Stratagene
<i>E. coli</i> D50	BW25113 + pKD46	ampicillin	(Datsenko and Wanner, 2000)
<i>E. coli</i> MG1655	Wild type	-	(Blattner et al., 1997)
<i>S. cerevisiae</i> PJ69-4a	MATa, <i>trp1-901, leu2-3, leu2-112, ura3-52, his3-200, gal4Δ, gal80Δ, GAL2-ADE2, LYS2::GAL1HIS3, met2::GAL7-lac</i>		(James et al., 1996)
SM29	MG1655, <i>tnaA(74-1319)::OL1/UAS hybrid</i> FROS array, <i>kan^R</i>	kanamycin	This study
SM57	MG1655 + pMA270	ampicillin	This study
SM65	SM29 + pMA270	kanamycin, ampicillin	This study
SM69	MG1655 + pMA275	ampicillin	This study
SM70	MG1655 + pMA276	ampicillin	This study
SM77	SM29 + pMA275	kanamycin, ampicillin	This study
SM126	MG1655 + pWM2060	amp	This study
SM153	SM29 + pMA301 + pMA270	kanamycin, chloramphenicol, ampicillin	This study
SM158	SM29 + pMA310 + pMA270	kanamycin, chloramphenicol, ampicillin	This study

Table S2: Replicons used in this study.

Plasmid	Characteristics	Resistance	Reference
pKD46	<i>bla, ara, oriR101, rep101^{ts}</i>	ampicillin	(Datsenko and Wanner, 2000)
pUC57-kan	Cloning vector	kanamycin	GenScript, Piscataway Township, NJ, USA
pUC19	Cloning vector	ampicillin	Anke Becker
pPS19	<i>pDS132, P_{BAD} lacI-eyfp λcl-ecfp P_{CP18}araE tetR</i>	chloramphenicol, spectinomycin	Dhruba Chatteraj (Srivastava et al., 2006)
pGBKT7	<i>E. coli / S. cerevisiae</i> Shuttle Y2H vector	kanamycin	Clontech
pBad24-LacI-venus	<i>venus</i> fluorophor	ampicillin	(Hammar et al., 2014)
synVicII-0.1	<i>oriR6K, 2μ ori, URA-3</i>	ampicillin	(Messerschmidt et al., 2015)
synVicII-1.3	synVicII-1.0 P _{A1/04/03} -RBSII- <i>gfp</i> (AAV)	ampicillin	(Messerschmidt et al., 2015)

pWM2060	pTrc99A , promoter down mutations in -35 and -10	ampicillin	Provided by W. Margolin (Bernard et al., 2007)
pMA164	pUC57 kan + OL1/UAS hybrid FROS array	kanamycin	This study
pMA165	pMA164 + tnaA-OL1/UAS hybrid FROS array-tnaA	kanamycin	This study
pMA182	synVicII-1.3, amp::cat	chloramphenicol	This study
pMA252	pMA164, <i>ColE1::oriR6K</i>	kanamycin	This study
pMA270	pWM2060 with λ cl fused to N-terminal fragment of mVenus, Gal4 fused to C-terminal fragment of mVenus	ampicillin	This study
pMA275	pWM2060, λ cl fused to full length mVenus	ampicillin	This study
pMA276	pWM2060, Gal4 fused to full length mVenus	ampicillin	This study
pMA301	pMA182, <i>gfp::OL1/UAS</i> hybrid FROS array	chloramphenicol	This study
pMA310	pMA182 Δ <i>gfp</i>	chloramphenicol	This study

Table S3: Oligonucleotides used in this study.

Name	Sequence
46	GTCTATCTGACCGCCGACAAGCAGAAGAAC
47	GTTCTTCTGCTTGTCGGCGGTGATAGAC
95	ACGCCAAGCTTGCATGCAGGCCTCTGCAGTCGACGGGCCGGGAC CGGCAAGATCAACAGGTA
96	TCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTCCCGGTAGCCATCACCAGAGCCAAA
97	GGTTAATTGGTTGTAACATTATTAGATTGGGCTTGATTTCCGCGAGTGGTACGTTTTACT
98	CAGTCATAACACTGGCGGAGACCTCTAGATATCGGATCCCACATGTGAAAGAGAACGCGG
102	GCGAGGATAAGTGCATTATG
103	ACAATAACCCGGAATGAAGC
257	AATTCGAAGACCTGGCGGTGATANNGNNCNNCGGAGTACTGTCCTCCGNNCNNNGNNTAC CACTGGCGGGAGACCT
258	CTAGAGGTCTCCCGCCAGTGGTANNNCNNNGNCGGAGGACAGTACTCCGNNNGNNCNNTAT CACCGCCAGGTCTTCG
267	ACAGCTCCTCGCCCTTGCTCACCATGGTCGCAATGCTGCGGCCAAACGTCTCTCAGGCC
268	CGCAGCATTGCGACCATGGTGAGCAAGGGCGAGGAGCTGTTAC
270	CAACTACAACAGCCACAACGTCTATCTGACCGCCTAACCTAGCAGGAGGAATTCACCATGAAGCT ACTGTCTTCTATC
389	GTATGTTTGTCTTACCTGTTGATCTTGCCGGTCCCGGGAACCTGTTGATAGTACGTACTAAGC
390	GAATCTAATCGGTTTGGCTCTGGTGATGGCTACCCGGGCATGTCAGCCGTTAAGTGTCC
409	CTACAACAGGGCAAAGCGCAAC
410	CAAATCAAGGGCGGTGATCGAC
646	ATGGAATTCGAGCTCGGTACCCGCGTTCTAGCAGGAAACAGCTATG
655	TATATGAGTAACTTGGTCTGACAGTCATCGCAGTACTGTTGTATTCATT

656	CTGCAGGTCGACTCTAGAGGATCCCCTTACTTGTACAGCTCGTCCATGCC
657	CCTGACTCAGTGCTTTCTATCCCGCTTTTGGCGAAAATGAGACG
712	GGTGAAAACCTCTGACACATGCAGCTTTACTTGTACAGCTCGTCCATGCCGAGA
714	GTGAGCAAGGGCGAGGAGCTGTTC
757	GACAAGCAGAAGAACGGCATCAAG
758	TTGATGCCGTTCTTCTGCTTGTCCGATACAGTCAACTGTCTTTG
769	GTTTCGATAGAAGACAGTAGCTTCATGGTGAATTCCTCCTGCTAGG
809	ATGGAATTCGAGCTCGGTACCCGCGTTCTAGCAGGAAACAGCTATGAAGCTACTGTCTTCTATC
1045	GTGAGTTTTCGTTCCACTAGGGATAACAGGGTAATGGGATCCGATATCTAGAGGTC
1046	GTACGTACTATCAACAGGTTCCAAGCTAGCGTTGTA AACGACGGCCAG
1259	TTTTCGTTCCACTAGGGATAACAGGGTAATCCAAGCTAGC
1260	GTACTATCAACAGGTTCCAAGCTAGCTTGGATTACCCTGT

4. Diskussion

Die Mehrheit aller Bakterien verfügt über ein zirkuläres Chromosom, welches von einem Replikationsursprung repliziert wird (Gao and Zhang, 2007; Stanier and Van Niel, 1962). Das eukaryotische Genom umfasst hingegen zahlreiche lineare Chromosomen, die mehr als einen Replikationsursprung enthalten (Bell and Dutta, 2002). Das gramnegative Bakterium *E. coli* toleriert einen zweiten Replikationsursprung auf dem Chromosom (Wang et al., 2011). Dieser kann zusätzlich zum nativen Replikationsursprung replizieren und ist zudem in der Lage, als alleiniger Replikationsursprung auf dem Chromosom zu fungieren (Ivanova et al., 2015; Wang et al., 2011). Des Weiteren ist *E. coli* in der Lage ein lineares Chromosom stabil zu vererben (Cui et al., 2007; Liang et al., 2013). Ob es möglich ist, das Genom von *E. coli* weiter umzustrukturieren, um es so dem eukaryotischen Genom weiter anzunähern und wie die DNA-Replikation dieser Rekonstruktionen, welche essentiell für die stabile Vererbung des Erbguts ist, abläuft, ist deshalb eine spannende Frage. In der vorliegenden Arbeit wurden *E. coli* Stämme mit multiplen Replikationsursprüngen und multi-chromosomale *E. coli* Stämme bezüglich ihrer DNA-Replikation untersucht (Milbredt et al., 2016). Zudem sollte eine Methode etabliert werden, mit der es möglich ist die Kopienzahl von Genloki zu bestimmen (Schindler et al., 2016)(Kapitel 3.3). In Bezug auf multi-chromosomale Bakterien könnte diese Methode die Möglichkeit bieten, zwei oder mehrere Chromosomen innerhalb einer Zelle unabhängig voneinander bezüglich der DNA-Replikation zu analysieren. In den folgenden Kapiteln sollen die wichtigsten Erkenntnisse dieser Arbeit und ihre Bedeutung diskutiert werden.

4.1 Fortschritt auf Einzelzellniveau: Entwicklung von Methoden zur Kopienzahlbestimmung

Analysen auf Einzelzellebene bieten den Vorteil, dass die Heterogenität einer Zellpopulation aufgelöst wird (Lidstrom and Konopka, 2010). Jede Zelle ist individuell und verfügt über charakteristische Merkmale. Auf Einzelzellniveau können seltene Merkmale wie Kopienzahlvariationen, die im Populationsmaßstab nicht auffallen, erkannt und analysiert werden. Variationen in der Kopienzahl von Genloki können dazu genutzt werden, genetische Prädispositionen zu erkennen oder Individuen voneinander zu unterscheiden (Nari, 1995; Redon et al., 2006). Die Anzahl von Kopien eines genetischen Locus innerhalb einer Zelle kann aber auch dazu genutzt werden, den Status der DNA-Replikation zu bestimmen. So liegt ein Genloki vor der DNA-Replikation in einfacher und nach der DNA-Replikation in doppelter Kopienzahl vor. Anhand der Kopienzahl kann so Rückschluss auf den Replikationsstatus der Zelle gezogen werden. Aufgrund des breiten Anwendungsspektrums von Kopienzahlbestimmungen in Eukaryoten und Prokaryoten sollten in der vorliegenden Arbeit neue Methoden etabliert werden, die auf Einzelzellebene eine einfache und zuverlässige Kopienzahlanalyse

ermöglichen. Ansätze für solche Methoden können auf verschiedenen Grundlagen basieren.

4.1.1 Entwicklung eines fluoreszenzbasierten Reportersystems für die Analyse von Genloki

Die DNA-Replikation, Chromosomensegregation und Chromosomenfaltung stellen elementare Prozesse aller lebenden Zellen dar. Um einen genaueren Einblick in die räumliche und zeitliche Organisation der DNA innerhalb der Zelle zu erhalten, kann die DNA markiert werden. Eine Methode um die DNA zu markieren ist das Fluorescence Repressor Operator System (FROS). FROS basiert auf Fusionen aus einem Repressor und einem Fluoreszenzprotein, die an ein Operator-Array, welches die Repressorbindestelle enthält und in die DNA integriert ist, binden (Lau et al., 2003; Robinett et al., 1996). Durch das Binden der Fusionsproteine an das Operator-Array wird die DNA markiert. Mit Hilfe von FROS kann die räumliche und zeitliche Anordnung von Genloki visuell verfolgt und die Kopienzahl auf Grund der resultierenden Foci quantifiziert werden (Gordon et al., 1997; Thanbichler and Shapiro, 2006; Wang et al., 2014). Die Differenzierung von Foci kann jedoch durch Faktoren wie „Sister chromosome cohesion“ erschwert werden (Joshi et al., 2011; Nielsen et al., 2006). Zudem wurde für das konventionelle FROS eine Protein-vermittelte Chromosomenpaarung zwischen den Kopien der Operator-Arrays beschrieben (Mirkin et al., 2014). Auf Grund dieser Faktoren können zwei Foci möglicherweise nicht getrennt wahrgenommen werden und so zu Missinterpretationen von zellulären Vorgängen, wie beispielsweise der Initiation der DNA-Replikation, führen. Eine automatisierte Foci-Detektion ist mit Programmen wie Oufiti oder MicrobeJ möglich (Ducret et al., 2016; Paintdakhi et al., 2016). Diese ist jedoch zeitintensiv und muss für jede Probe individuell angepasst werden. Die angeführten Nachteile limitieren das Anwendungsspektrum von FROS. Deshalb wurde in der vorliegenden Arbeit eine neue fluoreszenzbasierte Methode entwickelt, die FROS optimiert und zur Kopienzahlbestimmung genetischer Loki eingesetzt werden könnte. BiFCROS (Bimolecular Fluorescence complementation and Repressor Operator System) (Kap. 3.3) ist eine neue Methode, die auf der bimolekularen Komplementation von zwei nicht-fluoreszierenden Fragmenten eines Fluoreszenzproteins basiert, welche an ein hybrid FROS-Array binden (Kap. 3.3). BiFCROS umgeht aufgrund der Quantifizierung von Fluoreszenzsignalen ganzer Zellen die aufgeführten Probleme von FROS. Durch den Einsatz des hybrid FROS-Arrays und den Fusionen mit N- bzw. C-terminalen Fragmenten des Fluoreszenzproteins, ergibt sich nur ein Fluoreszenzsignal, wenn die Fusionsproteine an das hybrid FROS-Array gebunden sind (Kap. 3.3). Freie Fusionsproteine fluoreszieren nicht, wodurch die Hintergrundfluoreszenz sehr gering ist (Kap. 3.3). Die Fluoreszenzintensität der Zelle sollte somit proportional mit der Anzahl an Kopien des hybrid FROS-Arrays ansteigen, sodass ein oder zwei Foci anhand der Intensitätsunterschiede differenziert werden könnten. Mit Hilfe des konventionellen FROS kann diese Messung nicht durchgeführt werden, weil Vollängen Fluoreszenzproteine unabhängig von ihrer Bindung an das FROS-Array fluoreszieren und so das Hintergrundsignal deutlich erhöht ist. Für

die Auswertung der fluoreszenzmikroskopischen Daten von BiFCROS müssen keine Foci gezählt werden, sondern Zellumriss determiniert und die Fluoreszenzintensität innerhalb dieser Fläche bestimmt werden (Kap. 3.3). Alternativ könnten die Fluoreszenzintensitäten der Zellen mittels Durchflusszytometrie bestimmt werden, wodurch BiFCROS als Hochdurchsatz-Analyseverfahren genutzt werden könnte. Für BiFCROS konnte gezeigt werden, dass Zellen mit einer oder zwei Kopien des hybrid FROS-Arrays unterschiedliche Fluoreszenzintensitäten aufweisen (Kap. 3.3). Jedoch steigt die Fluoreszenzintensität mit der Zellgröße linear an. Wäre die Fluoreszenzintensität abhängig von der Kopienzahl müsste diese sprunghaft ansteigen, nachdem das hybrid FROS-Array repliziert wurde. Die Ergebnisse könnten darauf hinweisen, dass das Operator-Array als Plattform dient, an die die Fusionsproteine binden, sich komplementieren und nach einer gewissen Zeit wieder lösen, sodass ein nächstes Pärchen binden kann. BiFC ist sowohl *in vitro* als auch *in vivo* ein irreversibler Prozess (Shyu and Hu, 2008). Haben sich der N- und C-Terminus eines fragmentierten Fluoreszenzproteins komplementiert, wird diese Assoziation beibehalten (Kerppola, 2008). Um die „signal to noise ratio“ von BiFCROS zu verbessern, könnte ein zweites, anderes hybrid FROS-Array auf dem Chromosom integriert werden (Abb. 5 A). Die entsprechenden Proteine werden mit einem andersfarbigen Fluoreszenzprotein fusioniert, sodass die Fluoreszenzintensitäten beider Systeme parallel gemessen werden könnten. Aufgrund der Position und der daraus resultierenden Kopienzahl im Verlauf der Replikation, sollte sich eine unterschiedliche Verteilung der Intensitäten beider hybrid FROS-Arrays ergeben (Abb. 5 B + C). Im Rahmen der vorliegenden Dissertation wurde bereits ein zweites hybrid FROS-Array basierend auf den Repressoren LacI und TetR und ihren jeweiligen Bindestellen generiert (Schindler et al., 2016). Um weitere hybrid FROS-Arrays zu bauen, um mehr als zwei Loki zu analysieren, werden auch weitere Repressor-Operator Paare benötigt. Generell ist die Auswahl an diesen Paaren jedoch limitiert. Voigt und Kollegen haben deshalb in einer Studie gezielt neue Repressor-Operator Paare der TetR-Familie gesucht und getestet (Stanton et al., 2014). Diese könnten für den Aufbau von neuen hybrid FROS-Arrays verwendet werden. In dieser Arbeit wurden zwei hybrid FROS-Arrays generiert (Schindler et al., 2016)(Kapitel 3.3). In beiden Klonierungsprozessen konnten deutliche Abbauprodukte der hybrid FROS-Arrays detektiert werden. Dieses Problem kann auf die initiale Oligonukleotidsynthese zurückgeführt werden, die neben den korrekten Sequenzen auch verkürzte Sequenzen enthält. Verkürzte Oligonukleotide können während der Synthese entstehen, wenn die Schutzgruppe am 5'-Ende der Oligonukleotidkette nicht abgelöst wird. Deshalb ist die regelmäßige Verifizierung für den Aufbau eines hybrid FROS-Arrays essentiell. Für das LacO/TetO hybrid FROS-Array konnte mittels fluoreszenzmikroskopischer Analysen nachgewiesen werden, dass beide Repressoren an das Operator-Array binden (Schindler et al., 2016).

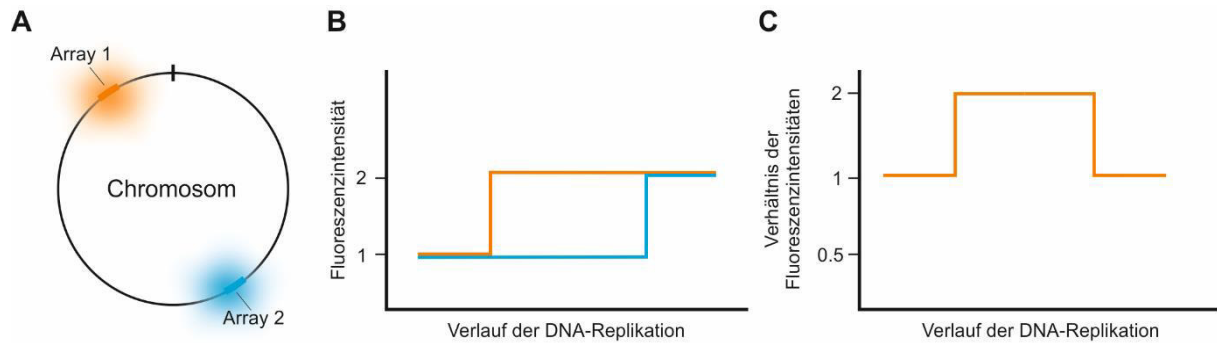


Abbildung 5: Verbesserung der „signal to noise ratio“. **A** Schematische Darstellung des *E. coli* Chromosoms mit zwei verschiedenen hybrid FROS-Arrays (orange und blau markiert). Der Replikationsursprung ist indiziert. **B** Verlauf der Fluoreszenzintensität in Abhängigkeit zur DNA-Replikation. Wird ein Operator-Array repliziert, steigt die Fluoreszenzintensität um den Faktor 2 an. Die Fluoreszenzintensität von Array 1 wird in orange und Array 2 in blau dargestellt. **C** Das Verhältnis der Fluoreszenzintensität von Array 1/Array 2 wird in orange dargestellt.

Fluoreszenzmikroskopische Analysen mit den Fusionen aus dem N- bzw. C-Terminus von mVenus und LacI und TetR ergaben jedoch kein distinktes Fluoreszenzsignal (vergl. Kap. 3.3). Die Komplementation von mVenus durch das Binden der Repressoren an das LacO/TetO-Array konnte somit nicht bestätigt werden. Das diffuse Fluoreszenzsignal könnte auf das Expressionslevel der Fusionen mit LacI und TetR zurückgeführt werden. Diese werden Plasmid-basiert exprimiert. Um das Expressionslevel zu verändern, wurden unter anderem der Replikationsursprung und der Promotor ausgetauscht. Beide Ansätze lieferten schwache distinkte Signale gepaart mit einem starken diffusen Signal, welches sich über die ganze Zelle erstreckte. Für die Induktion der Fusionsproteine mit λ cl und Gal4 wurde ein IPTG induzierbarer, synthetischer Promotor verwendet (Kap. 3.3) (Bernard et al., 2007). Für die Induktion von LacI konnte dieser Promotor nicht verwendet werden, weil IPTG die Konformation des Repressors verändert und dieser somit nicht mehr an das Operator-Array binden würde (Wilson et al., 2007).

4.1.2 Charakterisierung der DNA-Replikation mit Hilfe von Click-Chemistry

Mit Hilfe von Click-Chemistry können sowohl der arbeitsintensive Aufbau und die genomische Integration eines hybrid FROS-Arrays als auch die Expression der Fusionsproteine des BiFCROS Verfahrens umgangen werden. Click-Chemistry basiert auf dem Thymidin-Analogon EdU (5-Ethin-2-deoxyuridin), welches in die DNA eingebaut wird und einem Fluorochrom-markierten Alkin, welches in einer Cyclo-Addition an das Analogon gebunden wird (Salic and Mitchison, 2008). Die neu synthetisierte DNA wird somit an jeder Replikationsgabel markiert und die DNA-Synthese kann anhand der eingebauten Fluoreszenz nachgewiesen werden (Salic and Mitchison, 2008). Diese Methode konnte bereits in Eukaryoten und Prokaryoten angewandt werden, um replizierende Chromosomen zu visualisieren und die DNA-Replikation nachzuweisen (Buck et al., 2008; Darzynkiewicz et al., 2011; Dimude et al., 2015; Ferullo et al., 2009). Um die DNA-Replikation genauer zu charakterisieren könnte

das durch Click-Chemistry entstehende Fluoreszenzsignal quantifiziert werden. Theoretisch könnte das Fluoreszenzsignal Aufschluss darüber geben, wie viel EdU in die DNA integriert wurde um daraus abzuleiten, wie viele Replikationsgabeln in der Zelle vorliegen. Mit Hilfe der Quantifizierung der Fluoreszenzsignale von Click-Chemistry könnte es möglich sein, die DNA-Replikation anhand der Replikationsgabeln zu charakterisieren, jedoch ist es nicht möglich spezifischen Genloki zu analysieren. Die DNA-Replikation kann mittels Click-Chemistry auf globalem Level jedoch nicht auf lokalem Level, das heißt auf Genniveau, analysiert werden, wodurch das Anwendungsspektrum von Click-Chemistry limitiert wird. Click-Chemistry und BiFCROS erfordern zunächst eine chemische Behandlung der Zelle oder eine Genommanipulation, die die DNA-Replikation beeinflussen könnten. Eine Methode die dieses Problem umgeht ist Linker-counting.

4.2 Ein neuer Ansatz um die Kopienzahl von chromosomalen Elementen anhand von kurzen, individuellen Sequenzen zu bestimmen

Anhand von Einzelzellanalysen können individuelle Merkmale genau untersucht werden (Lidstrom and Konopka, 2010). Seltene Merkmale, die bei Analysen von mehreren Zellen nicht auffallen, werden so erkannt. Allerdings beträgt der DNA- und RNA-Gehalt einer einzelnen Zelle nur wenige Femto-Pikogramm (Raghunathan et al., 2005). Diese Menge ist bei weitem nicht genug für Anwendungen wie Sequenzierungen oder Microarray-basierte komparative genomische Hybridisierung (CGH) (Huang et al., 2015). Die vorliegende Gesamt-DNA/RNA einer Zelle kann aber mittels Whole Genome Amplification (WGA) angereichert werden. WGA-Ansätze können in PCR-basierte und isothermale Methoden unterteilt werden. Zu den PCR-basierten Methoden zählt die DOP-PCR (Degenerate Oligonucleotide-Primed Polymerase Chain Reaction), die degenerierte Primer für die Amplifizierung verwendet (Telenius et al., 1992). Auf diese Weise konnten bereits erfolgreich Nanogramm von einem humanen Genom amplifiziert werden (Cheung and Nelson, 1996). MDA (Multi Displacement Amplification) ist eine isothermale WGA-Methode (Dean et al., 2001). Es basiert auf zufälligen Hexameren, die als Primer eingesetzt werden und einer ϕ 29 DNA-Polymerase, die eine starke Strangverdrängungsaktivität aufweist, um in einer isothermalen Reaktion die DNA anzureichern. Im Gegensatz zur DOP-PCR liefert MDA eine höhere Leseabdeckung des Genoms, für beide Ansätze wird jedoch eine ungleiche Anreicherung der DNA-Fragmente beschrieben (Huang et al., 2015). Das heißt, einige DNA-Abschnitte werden häufiger amplifiziert als andere. Ausgehend von einer Einzelzelle, in der eine einzige Kopie des Genoms vorliegt, kann diese ungleichmäßige Amplifizierung das Ergebnis von Kopienzahlanalysen. Deshalb wurde in der vorliegenden Arbeit ein alternativer Ansatz entwickelt, um die ungleiche Anreicherung durch eine indirekte Quantifizierung von Genloki anhand definierter DNA-Sequenzen zu umgehen. Die zu untersuchenden DNA-Fragmente könnten mit kurzen, individuellen

Sequenzen, sogenannten Linkern, versehen und die DNA-Linker-Konstrukte gezielt angereichert werden. Anhand der Anzahl von verschiedenen Linkern könnte so die Anzahl an Genloki in der vorliegenden Probe bestimmt werden. Eine Gegenüberstellung der Kopienzahlbestimmung mittels MDA und der Linker-basierten Methode ist in Abbildung 6 aufgezeigt. Mit der Linker-basierten Methode sollte es theoretisch möglich sein, die verzerrte Kopienzahlbestimmung der WGA-Methoden zu umgehen und die Kopienzahl mehrerer, genetischer Loci parallel in einer Reaktion zu bestimmen.

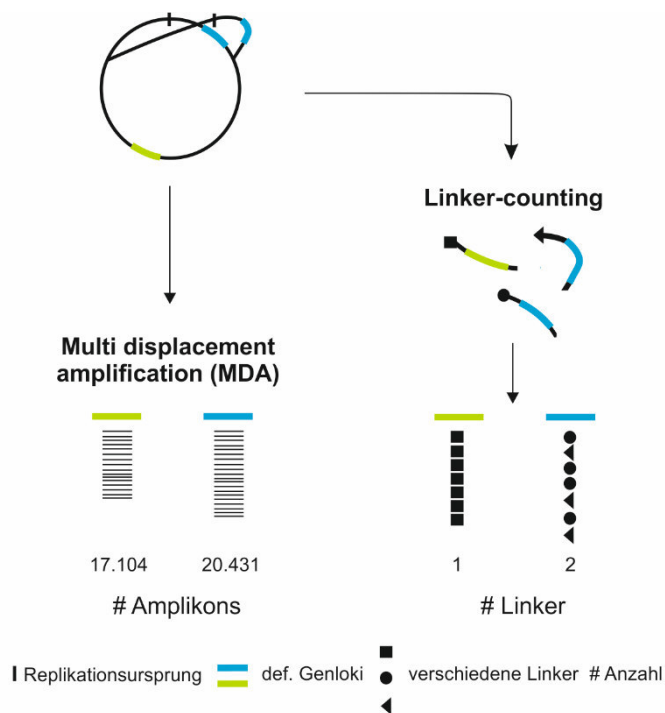


Abbildung 6: Vergleich von MDA und der Linker-basierten Methode zur Bestimmung von Kopienzahlen. Zu untersuchende Genloki sind blau und grün markiert. MDA bestimmt die Kopienzahl anhand der Anzahl an Amplicons, die Linker-basierte Methode bestimmt die Anzahl anhand der verschiedenen Linker-Sequenzen.

Diese Grundidee wurde im Rahmen der vorliegenden Dissertation unter dem Arbeitstitel Linker-counting bearbeitet. Die Methode basiert auf einem Set aus zwölf Linkern mit einer Länge von 21 bp, deren DNA-Sequenzen sich komplett voneinander unterscheiden (Abb. 7). Wird ein bestimmter Locus mittels Linker-counting analysiert, wird die chromosomale DNA einer einzigen Zelle mit dem Restriktionsenzym *MseI* verdaut, wodurch DNA-Fragmente zwischen 500 und 1000 bp entstehen (Abb. 7, 1.). Der entstehende Überhang der DNA-Fragmente ist kompatibel mit den Linkern, sodass diese gezielt an die DNA-Fragmente ligiert werden können (Abb. 7, 2.). Liegt der zu untersuchende Locus in zwei Kopien in der Zelle vor, so steigt die Wahrscheinlichkeit, dass zwei verschiedene Kombinationen von Linkern an die DNA-Fragmente ligiert werden. Zur Vervielfältigung der DNA-Fragmente erfolgt eine PCR mit Linker-spezifischen Primern (PCR1) (Abb. 7, 3.). Im nächsten Schritt wird die angereicherte Bibliothek von DNA-Fragmenten in ein Plasmid kloniert, um eine unspezifische Anreicherung in folgenden PCRs zu verhindern (Abb. 7, 4.). Die Plasmid-Bibliothek ist dann Template für eine zweite PCR mit einem Plasmid-spezifischen und einem Genlokus-spezifischen Oligonukleotid (PCR2) (Abb. 7, 5.). In einer dritten PCR werden die DNA-Fragmente von Interesse mit einem Plasmid-spezifischen und

einem Lokus-spezifischen Oligonukleotid angereichert (PCR3) (Abb. 7, 6.). PCR3 ergibt ein kürzeres Amplikon als PCR2, um unspezifische Anreicherungen aus der vorherigen PCR zu vermeiden.

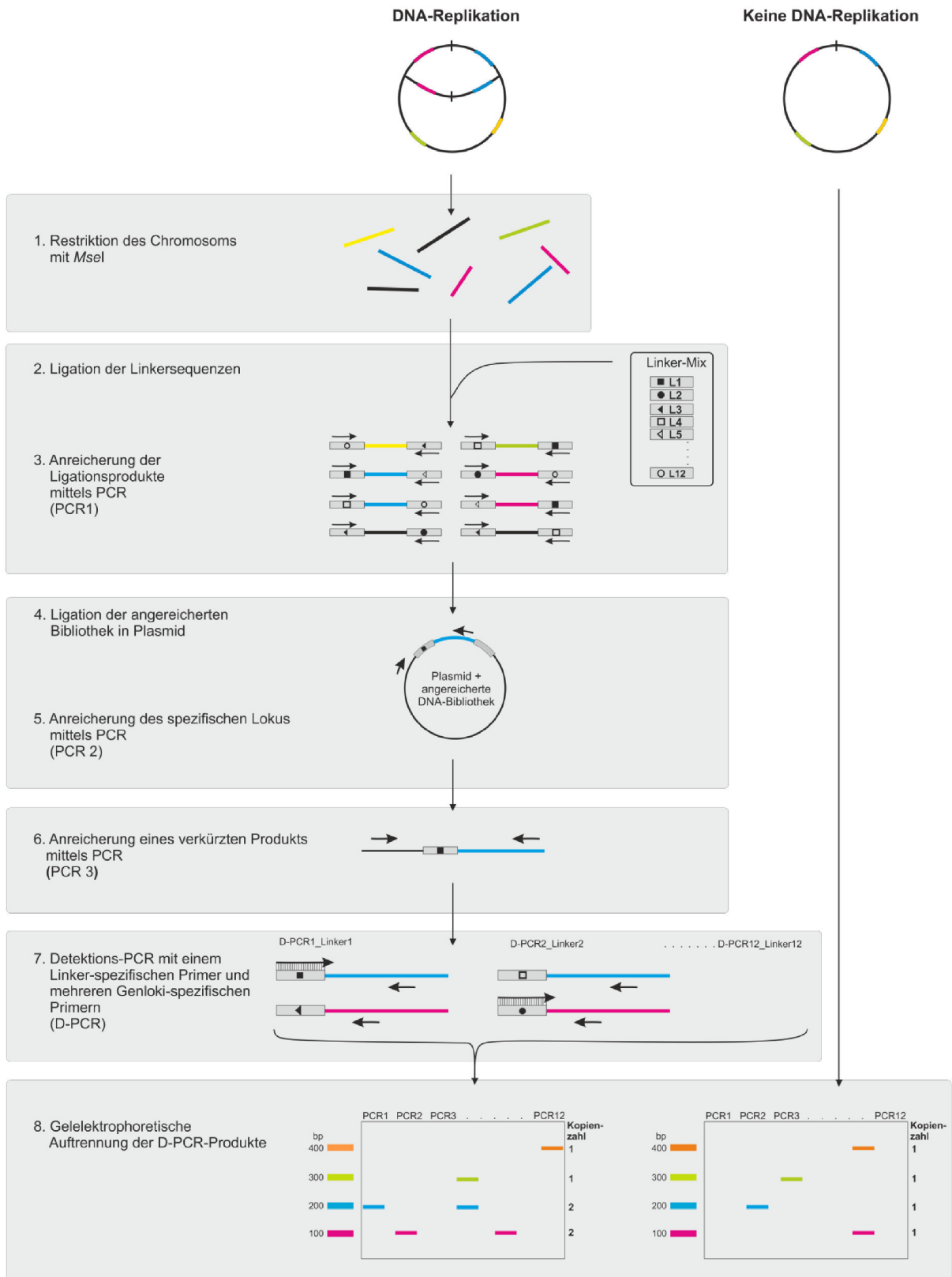


Abbildung 7: Prinzip von Linker-counting. Linker-counting wird am Beispiel eines replizierenden Chromosoms dargestellt. Mit einem nicht replizierenden Chromosom wird analog verfahren. Die einzelnen Schritte von Linker-counting sind grau hinterlegt und werden im Text ausführlich beschrieben. Zu untersuchende Genloki sind farblich markiert. Die Linker1-12 werden durch unterschiedliche Symbole (2.) dargestellt. Der blaue und magenta-farbene Genloki liegen zweimal in der Zelle vor. Durch die Ligation mit zwei unterschiedlichen Linkern sind nach der gelelektrophoretischen Auftrennung der zwölf D-PCR-Produkte zwei Fragmente mit einer Größe von 100 bzw. 200 bp zu erkennen. Der grüne und orange-farbene Lokus liegen mit einer Kopie in der Zelle vor und können nur einmal in der D-PCR nachgewiesen werden. Ein Chromosom, das nicht repliziert, enthält alle Genloki in einfacher Kopie. Mit den 3' angefügten Linkern wird analog verfahren.

Die letzte PCR, D-PCR, erfolgt mit einem Oligonukleotid, das an den Linker bindet und mit Oligonukleotiden, die spezifisch an die zu untersuchenden Genloki binden (Abb. 7, 7.). Mehrere Genloki können so in einer Reaktion nachgewiesen werden. Insgesamt werden zwölf D-PCRs durchgeführt, da eingangs auch zwölf verschiedene Linker-Sequenzen an die DNA-Fragmente ligiert wurden. Ist das Linker-spezifische Oligonukleotid kompatibel mit dem vorliegenden DNA-Fragment in der Probe, entsteht ein Amplikon. Nach der gelelektrophoretischen Auftrennung der D-PCR-Produkte können die Genloki anhand der eingesetzten Kombination aus Linker-spezifischen Oligonukleotiden und der Anzahl an spezifischen Amplikons quantifiziert werden (Abb. 7, 8.). Liegt ein Lokus in mehreren Kopien vor, sollten pro Kopie unterschiedliche Linker-Kombinationen nachgewiesen werden. Dies sollte anhand von spezifischen Amplikons in verschiedenen PCRs (PCR1-12) zu beobachten sein (Abb. 7, 8.). Theoretisch sollte es möglich sein zwischen Chromosomen zu unterscheiden, die replizieren bzw. nicht replizieren (Abb. 7). Neben der Gelelektrophorese könnte die Kopienzahl der Genloki auch durch Illumina-Sequenzierung bestimmt werden (Quail et al., 2009). Um eine Illumina-kompatible DNA-Bibliothek zu erhalten, könnte durch die Verwendung der Taq-Polymerase oder dem Klenow-Fragment ein A-Nukleotid an die 3'-Enden der DNA-Fragmente (Abb. 7, 3.) angefügt werden. Illumina-Adaptoren mit einem überhängenden 3'-T-Nukleotid könnten an die DNA-Bibliothek ligieren, sodass nachfolgende Schritte für die Illumina-Sequenzierung ausgeführt werden könnten (Illumina® Technology). Das Auslesen der Kopienzahlen mit Hilfe der Illumina-Sequenzierung oder der PCR-basierten Methode würde Linker-counting zu einem Hochdurchsatz-Analyseverfahren machen. Linker-counting könnte generell für die Analyse von Kopienzahlen in allen Domänen des Lebens eingesetzt werden. Kopienzahlvariationen (CNVs) sind ubiquitär verbreitet und oft an Krankheitsbildern beteiligt (Redon et al., 2006). Eine effiziente und genaue Analyse ist deshalb unerlässlich. Linker-counting stellt einen Ansatz dar, mit dem es möglich sein könnte, die Kopienzahl von Genloki zu bestimmen. Für die Durchführung der einzelnen Schritte (Abb. 7) werden grundlegende Techniken der Molekularbiologie kombiniert, sodass keine besonderen Gerätschaften benötigt werden würden. Demnach würde kein großer Kostenaufwand entstehen, sodass die Methode von jedem Laboratorium genutzt werden könnte. Im Hinblick auf die Analyse synthetischer Chromosomen und multi-chromosomaler Bakterien

könnten mittels Linker-counting zwei oder mehr Replikons innerhalb einer Zelle unabhängig voneinander auf Einzelzellniveau analysiert werden. Alle acht Schritte des Linker-counting Protokolls konnten unabhängig voneinander erfolgreich ausgeführt werden. Dafür wurden jedoch keine Einzelzellen verwendet, sondern die aufgearbeitete, genomische DNA einer *E. coli* Kultur. Um zu verifizieren, dass alle Schritte auch miteinander funktionieren, sollte ein geeignetes Testsystem verwendet werden. Dieses Testsystem sollte gut zugänglich sein, eine definierte Anzahl an Chromosomen aufweisen und keine Substanzen enthalten, die anschließende Prozesse, wie Restriktionen, Ligation oder PCR in irgendeiner Form inhibieren. Ein Testsystem, welches über die genannten Charakteristika verfügt, wird im Folgenden beschrieben.

4.2.1 Nukleotide als Testsystem für Linker-counting

2011 verknüpfte die Arbeitsgruppe um Michael Wigler „single cell amplification“ mit Hochdurchsatz-Sequenzierung und konnte so erfolgreich CNVs von Brustkrebszellen zweier Individuen determinieren (Navin et al., 2011). Dabei verwendeten sie als Ausgangsmaterial keine intakten Zellen, sondern nur Zellkerne, die sie mittels Fluorescence-Activated Cell Sorting (FACS) detektierten und sortierten. Auch andere Gruppen verwendeten Nuclei als Template für Einzelzellanalysen (Navin et al., 2011). Die Verwendung von Nuclei statt ganzer Zellen bietet den Vorteil, dass inhibierende Wirkungen von Zellbestandteilen oder Puffer, die zur Zellyse genutzt werden, vermieden wird. Prokaryoten enthalten im Gegensatz zu eukaryotischen Zellen keinen Zellkern, jedoch verfügen sie über ein kern-ähnliches Organell, das Nukleoid, welches die Region innerhalb der bakteriellen Zellen beschreibt, in der das Chromosom kompaktiert vorliegt (Robinow and Kellenberger, 1994). Das Nukleoid konnte bereits erfolgreich aus intakten *E. coli* Zellen isoliert werden (Cunha et al., 2001; Zimmerman and Murphy, 2001). Dazu wurde die bakterielle Zelle zunächst lysiert, wodurch Sphäroplasten entstanden. Durch einen anschließenden osmotischen Schock wurde das Nukleoid aus den Sphäroplasten gelöst (Cunha et al., 2001; Zimmerman and Murphy, 2001). Um ein geeignetes Testsystem für Linker-counting zu generieren, wurde im Rahmen der vorliegenden Dissertation die Idee entwickelt Nukleotide aus *E. coli* zu isolieren und diese mittels FACS zu vereinzeln. Um zunächst die Effizienz der Nukleotidisolierung aus *E. coli* Zellen zu testen, wurden die verschiedenen Fraktionen, intakte Zellen, Sphäroplasten und Nukleotide, mit dem DNA-Farbstoff LDS751 gefärbt und im Durchflusszytometer analysiert. Anhand individueller Verteilungsmuster konnte deutlich zwischen intakten Zellen, Sphäroplasten und Nukleotiden differenziert werden (Abb. 8). Die identifizierten Nukleotide wurden mittels FACS in 96-well-Platten sortiert, sodass pro „well“ ein Nukleoid vorlag. Diese vereinzelt Nukleotide sollten genutzt werden um Linker-counting durchzuführen. Es gelang jedoch nicht die zuvor beschriebenen Schritte dieser Methode im Einzelnen zu etablieren.

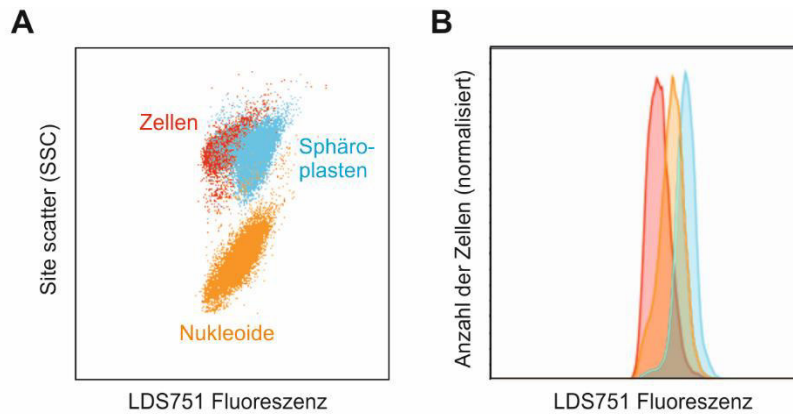


Abbildung 8: Differenzierung zwischen Zellen (rot), Sphäroplasten (blau) und Nukleoiden (orange) anhand durchflusszytometrischer Muster. Die Proben wurden mit LDS751 gefärbt. **A** Dot plot der verschiedenen Populationen **B** Histogramm der verschiedenen Populationen.

Alle durchgeführten PCRs lieferten keine oder unspezifische Ergebnisse. Auch zwei aufeinander folgende PCRs - um die Anreicherung zu erhöhen - lieferten ähnliche Ergebnisse. Unspezifische Amplikons sowie eine fehlende Reproduzierbarkeit der Ergebnisse können auf Kontaminationen zurückgeführt werden, da die Gewinnung, Sortierung und Analyse der Nukleoiden nicht unter sterilen Bedingungen durchgeführt wurden. Es wird deutlich, dass Einzelzellanalysen aufgrund des geringen Ausgangsmaterials fehleranfällig sein können.

4.3 Die „Eukaryotisierung“ von *Escherichia coli* – ein Gedankenexperiment

Eukaryoten und Prokaryoten unterscheiden sich in vielerlei Hinsicht voneinander. Die genomische DNA der Eukaryoten ist in linearen Chromosomen angeordnet, die mitochondriale DNA ist hingegen zirkulär (Taanman, 1999). Es wird postuliert, dass die Linearität essentiell für die Meiose ist, in der Homologe gepaart und akkurat verteilt werden (Ishikawa and Naito, 1999). Zirkuläre Chromosomen würden wahllos verteilt, weil diesen die homologen Bereiche, die Centrosomen, fehlen (Ishikawa and Naito, 1999). Die Mehrheit der Bakterien prozessiert ein einziges zirkuläres Chromosom (Bendich and Drlica, 2000). Die DNA-Replikation startet auf eukaryotischen Chromosomen an mehreren Positionen, wohingegen bakterielle Chromosomen nur einen Replikationsursprung besitzen (Robinson and Bell, 2005). Warum die Natur diese distinkte Organisation der Genome in Pro- und Eukaryoten generiert hat, ist eine spannende Frage. Aufschluss könnten Experimente liefern, die die „Eukaryotisierung“ von Prokaryoten oder *vice versa* fokussieren. Studien konnten zeigen, dass zirkuläre Chromosomen in Eukaryoten instabil sind (Ishikawa and Naito, 1999). Die zirkuläre Organisation von Chromosom III der Hefe *Saccharomyces cerevisiae* konnte nur in einem diploiden Stamm aufrechterhalten werden, ein haploider Stamm zeigte starke Wachstumsdefekte (Haber et al., 1984). Die geringere Größe und Komplexität und die resultierende bessere genetische Manipulierbarkeit des bakteriellen Genoms, spricht hingegen für eine höhere Toleranz prokaryotischer Systeme gegenüber der „Eukaryotisierung“.

Ein Gedankenexperiment um die Toleranz von Bakterien hinsichtlich der „Eukaryotisierung“ zu testen soll im Folgenden näher ausgeführt werden. Abgeleitet von dem Genom der Bäckerhefe *S. cerevisiae* ist es denkbar, einen *E. coli* Stamm herzustellen, der 16 lineare Chromosomen statt einem einzigen, zirkulären Chromosom enthält (Abb. 9). Im Durchschnitt würde ein Chromosom ca. 290 kbp umfassen. Kleine Untereinheiten bieten den Vorteil, dass diese molekularbiologisch besser zugänglich und somit genetisch besser manipulierbar sind (Liang et al., 2013; Schindler and Waldminghaus, 2015). Modifikationen könnten leichter eingefügt werden. Welche Voraussetzungen für ein lineares Chromosom in Prokaryoten geschaffen werden müssten, könnte von natürlichen Vertretern abgeleitet werden.

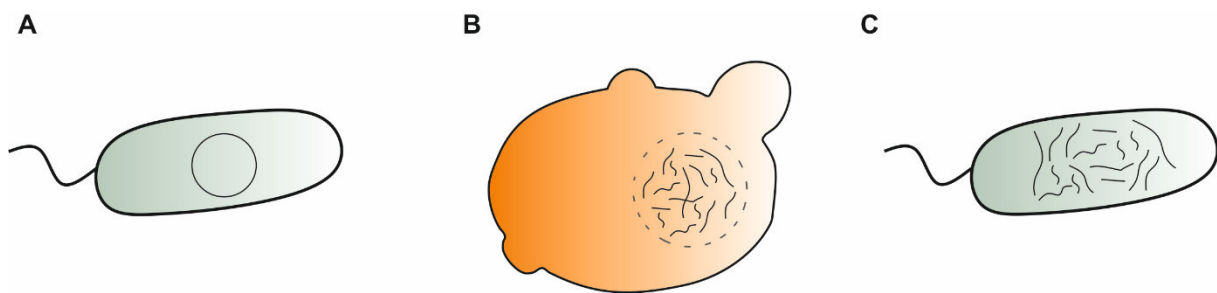


Abbildung 9: „Eukaryotisierung“ von *E. coli*. **A** *E. coli* Zelle mit einem zirkulären Chromosom. **B** Hefezelle mit 16 linearen Chromosomen im Zellkern. **C** „Eukaryotisierte“ *E. coli* Zelle, die 16 lineare Chromosomen enthält.

90% aller sequenzierten Prokaryoten enthalten ihre genetische Information auf einem einzigen, zirkulär angeordneten Chromosom (Ochman, 2002). Dazu gehören die Modellorganismen *E. coli*, *Bacillus subtilis* und *Caulobacter crescentus* (Blattner et al., 1997; Itaya and Tanaka, 1991; Nierman et al., 2001). Zu den übrigen 10% zählt *V. cholerae*, ein Modellorganismus für die Gruppe der multi-chromosomalen Bakterien (Heidelberg et al., 2000). *V. cholerae* gehört zur Familie der *Vibrionaceae*, deren Angehörigen das Genom auf zwei zirkuläre Chromosomen verteilt enthalten (Okada et al., 2005). Andere Spezies wie *Rhodobacter sphaeroides* und *Leptospira interrogans* enthalten ebenfalls zwei zirkuläre Chromosomen (Nascimento et al., 2004; Suwanto and Kaplan, 1989). Das Pflanzenpathogen *Agrobacterium tumefaciens* ist ein „Individualist“ in der Gruppe der multi-chromosomalen Bakterien: Es verfügt über ein zirkuläres und ein lineares Chromosom (Allardet-Servent et al., 1993). *Borellia burgdorferi* hingegen war das erste charakterisierte Bakterium mit einem einzigen, linearen Chromosom (Ferdows and Barbour, 1989). In der Gattung der *Borrelia* ist diese Anordnung des Genoms hochkonserviert (Casjens et al., 1995). Auch Vertreter der *Streptomyceten* wie *Streptomyces coelicolor* und *Streptomyces avermitilis* verfügen über ein lineares Chromosom (Galperin, 2007). Allen Chromosomen, zirkuläre oder linear ist gemein, dass für deren Vererbung die Duplikation extrem wichtig ist (Robinson and Bell, 2005). Ein Replikationsursprung ist deshalb essentiell für ein Chromosom. Dieser ist sowohl auf zirkulären als auch linearen Bakterienchromosomen so angeordnet,

dass beide Replichors, ausgehend von der bidirektionalen DNA-Replikation, ungefähr gleich lang sind (Musialowski et al., 1994; Picardeau et al., 1999; Prescott and Kuempel, 1972). Für das Gedankenexperiment zur „Eukaryotisierung“ von *E. coli* sollen 16 lineare Chromosomen prozessiert werden. Dafür müssten zunächst 16 Replikationsursprünge gefunden werden, die alle untereinander kompatibel sind. Das heißt, dass die 16 Replikons stabil in einer Zelle koexistieren können und dass die Replikationsursprünge nicht um Replikationsfaktoren konkurrieren (Stuitje and Meijer, 1983; Yamaguchi et al., 1982). Alle 16 Chromosomen könnten mit einer Kopie des *oriC* ausgestattet werden. Es konnte bereits gezeigt werden, dass zwei *oriC*-Kopien auf dem gleichen Chromosom toleriert werden und diese zur gleichen Zeit, also synchron mit der DNA-Replikation beginnen (Ivanova et al., 2015; Milbredt et al., 2016; Wang et al., 2011). Mehrere Kopien des gleichen Replikationsursprungs können jedoch auch zu Problemen führen. Zum einen werden diese Replikationsursprünge immer um die gleichen Replikationsfaktoren wie DnaA konkurrieren und zum anderen ist es aufgrund der homologen DNA-Sequenz von *oriC* möglich, dass ein Chromosom in das andere integriert (Lobner-Olesen, 1999; Messerschmidt et al., 2016 submitted; Skarstad and Lobner-Olesen, 2003). Es wäre möglich die *oriC*-Sequenz derart zu verändern, dass außer den wichtigen Sequenzmotiven wie DnaA-Boxen, GATCs und der 13-mer AT-reichen Region, die restliche Sequenz variabel gestaltet wird. Dazu könnte das Programm MARSeG (Motif Avoiding Randomized Sequence Generator) verwendet werden (Schindler et al., 2016). Mit Hilfe dieses Programms könnten die Sequenzen mit hoher Diversität designt und gleichzeitig zusätzliche DnaA-Boxen oder GATC-Motive, die die Funktionalität des *oriC* verändern könnten, ausgeschlossen werden. Die unterschiedlichen *oriC* könnten dann als Replikationsursprünge für die 16 Chromosomen fungieren. Ob die variablen *oriC* ebenfalls zur Integration von einem Chromosom führen, könnte durch Evolutionsexperimente analysiert werden. Als alternativer Replikationsursprung könnte *orill* von *V. cholerae* fungieren, der bereits als geeigneter Replikationsursprung für sekundäre Chromosomen in *E. coli* identifiziert wurde (Egan and Waldor, 2003; Liang et al., 2013; Messerschmidt et al., 2015). Es wäre denkbar die *orill*-Sequenzen von anderen Vertretern der *Vibrionaceae* wie *Vibrio parahaemolyticus* oder *Vibrio tubiashii* für weitere Chromosomen zu verwenden (Okada et al., 2005; Richards et al., 2014). Es konnte jedoch bereits durch Konjugationsassays gezeigt werden, dass diese Replikationsursprünge nicht stabil in einer Zelle koexistieren können (Messerschmidt, pers. Kommunikation). Das α -Proteobakterium *Sinorhizobium meliloti* enthält neben dem Hauptchromosom zwei Megaplasmide, pSymA und pSymB, deren Replikationsursprünge charakterisiert wurden und zeigen, dass sie, genauso wie das Hauptchromosom, nur einmal pro Zellzyklus mit der DNA-Replikation beginnen (Frage et al., 2016). Theoretisch könnten die Replikationsursprünge der Megaplasmide für zwei der 16 Chromosomen verwendet werden. Es konnte jedoch bereits gezeigt werden, dass die repABC-basierten Replikationsursprünge der Megaplasmide von *S. meliloti* nicht in der Lage sind in *E. coli* zu replizieren

(Bartosik, 2001; Bartosik et al., 2001; Lipps, 2008). Es ist also wichtig das Wirtsspektrum von Replikationsursprüngen zu berücksichtigen. Für das Wirtsspektrum von Plasmiden gibt es einige Studien (del Solar et al., 1996; Jain and Srivastava, 2013). So können Plasmide bezüglich ihrer Wirtsspezifität in zwei Klassen, „narrow host range“ (NHR) und „broad host range“ (BHR) unterteilt werden (Datta and Hedges, 1972; Krishnapillai, 1988; Mazodier and Davies, 1991; Tardif and Grant, 1983). Ein Plasmid kann auf eine bestimmte Bakterienart beschränkt sein und repliziert auch in verwandten Arten nicht, oder das Plasmid weist ein breites Wirtsspektrum auf (De Gelder et al., 2007; Pinkney and Thomas, 1987). Eine weitere Möglichkeit, um 16 kompatible Replikationsursprünge zu identifizieren, wäre die Replikationsursprünge von verschiedenen multi-chromosomalen Bakterien untereinander bezüglich ihrer Kompatibilität zu testen. Ein Sequenzvergleich der Replikationsursprünge könnte bereits einen Hinweis über die Kompatibilität untereinander geben. Dabei könnte es möglich sein, dass ähnliche Sequenzen nicht miteinander kompatibel sind, weniger ähnliche hingegen schon. Im Gegensatz zu Prokaryoten verfügen Eukaryoten nicht über einen definierten Replikationsursprung. Die Bäckerhefe *S. cerevisiae* enthält beispielsweise AT-reiche Regionen auf den Chromosomen, sogenannte ARS (Autonomously Replicating Sequence), an denen der ORC (Origin Recognition Complex) analog zu DnaA in Bakterien bindet und die DNA-Replikation einleitet (Robinson and Bell, 2005). Solche DNA-„Hot spots“ könnte es laut einer Studie auch in Bakterien geben (Gao, 2015). Aufgrund Computer-basierter Analysen wurde postuliert, dass Bakterien mehr als nur einen Replikationsursprung auf dem Chromosom tragen und dass an diesen Orten Replisomen assemblieren und die DNA-Replikation starten könnten (Gao, 2015). Dass *E. coli* in der Lage ist, generell mehr als einen Replikationsursprung pro Chromosom zu tolerieren und dieser auch aktiv repliziert, konnte bereits gezeigt werden (Liang et al., 2013; Milbredt et al., 2016; Wang et al., 2011). Es ist bekannt, dass die DNA-Replikation auch unabhängig von einem Replikationsursprung starten kann (Hawkins et al., 2013; Kogoma, 1997; Kreuzer and Brister, 2010). Für das Archaeon *Haloferax vulcanii* konnte beispielsweise ein deutlicher Wachstumsvorteil gezeigt werden, wenn kein Replikationsursprung auf dem Chromosom vorhanden war und die DNA-Replikation Replikationsursprung-unabhängig stattfand (Hawkins et al., 2013). RNase HI defiziente *E. coli* Stämme mit einem deletierten Replikationsursprung waren ebenfalls in der Lage, die DNA-Replikation an sogenannten R-Loops, einem dreisträngigen Hybridmolekül aus DNA und RNA, zu starten (Dimude et al., 2015; Kogoma, 1997; Maduiké et al., 2014). Möglicherweise könnte dieses Erkenntnis auf den „eukaryotisierten“ *E. coli* Stamm übertragen werden. Statt 16 verschiedene Replikationsursprünge zu verwenden, könnten alle Chromosomen in einem RNase HI-defizienten Stamm über R-Loops mit der DNA-Replikation starten. Um alle genetischen Informationen stabil zu vererben, müsste sichergestellt werden, dass alle Chromosomen einen R-Loop ausbilden. Es konnte gezeigt werden, dass sich R-Loops häufig in hochtranskribierten Regionen ausbilden (Maduiké et al., 2014). So wäre es denkbar einen

künstlichen Replikationsursprung auf dieser Grundlage zu entwerfen. Dieser könnte aus dem Promotor einer hochtranskribierten Region und einer variablen, „nonsense“ DNA bestehen. Durch die variable Sequenz wird die Homologie unter den künstlichen Replikationsursprüngen reduziert. Hinter dieses Konstrukt sollte ein starker Terminator platziert werden, um die Transkription nachfolgender Gene nicht zu beeinflussen. Um zu testen ob dieser künstliche Replikationsursprung überhaupt in *E. coli* funktioniert, könnte er in ein Replikon inseriert werden, welches in der Lage ist, in RNase HI-defizienten Zellen zu replizieren und keinen funktionellen Replikationsursprung für *E. coli* enthält. Zudem sollte das Replikon nach Möglichkeit wenige Gene enthalten, damit ausgeschlossen werden kann, dass die Initiation der DNA-Replikation möglicherweise von dem Transkriptionsstart an diesen gestartet wird. Wird das Replikon stabil vererbt, sollte auch die Initiation der DNA-Replikation durch den künstlichen Replikationsursprung funktionieren. Die regulierte Initiation erlaubt eine geordnete DNA-Replikation und Segregation. Diese ist sehr wichtig für Organismen wie beispielsweise *E. coli* und die meisten Eukaryoten, deren Ploidie streng geregelt ist (Lindas and Bernander, 2013). Ob die Chromosomen mit dem künstlichen Replikationsursprung einmal pro Zellzyklus replizieren, sollte deshalb berücksichtigt werden.

Neben dem Replikationsursprung gibt es weitere Faktoren, sogenannte „Chromosome maintenance systems“, die für die Aufrechterhaltung und stabile Vererbung eines bakteriellen Chromosoms von Bedeutung sind.

4.3.1 Welche „Chromosome maintenance systems“ werden auf linearen Chromosomen benötigt?

Für die „Eukaryotisierung“ sollen die linearen Chromosomen stabil in *E. coli* aufrechterhalten werden. Im Folgenden werden einige Systeme und ihre Notwendigkeit für lineare Chromosomen diskutiert, die an der Erhaltung und stabilen Vererbung des nativen *E. coli* Chromosoms beteiligt sind. Für die Termination der DNA-Replikation verfügt das zirkuläre Chromosom von *E. coli* über eine Terminationsdomäne (*ter*) (Duggin et al., 2008). Aufgrund des Protein-DNA Komplexes Tus-*ter* werden die Replikationsgabeln in dieser Region aufgehalten und die Replisomen disassemblieren (Neylon et al., 2005). Lineare Chromosomen benötigen hingegen keine Terminationsdomäne, weil die Enden des Chromosoms durch Telomere, welche aus repetitiven Sequenzen bestehen, vorgegeben sind (Blackburn, 1991). Für die DNA-Replikation ergibt sich durch die lineare Organisation des Chromosoms ein Problem. Die DNA-Polymerase kann keine linearen DNA-Fragmente replizieren, weil sie RNA-Primer als Anfangspunkt für die Synthese benötigt (Hinnebusch and Tilly, 1993). Bei einem linearen Chromosom, bei dem die Enden zum Schluss repliziert werden, ergibt sich ein 3' Überhang, der noch aufgefüllt werden muss. Bakterien mit linearen Chromosomen haben zur Bewältigung dieses Problems verschiedene Strategien entwickelt (Hinnebusch and Tilly, 1993). Die Gattung der *Streptomyces*

verfügt über „invertron“ Telomere, die invertierte Sequenzwiederholungen und ein 5'-kovalent gebundenes Terminalprotein enthalten, welches an einem noch nicht aufgeklärten Mechanismus beteiligt ist, der das rückständige Ende des Folgestrangs nach der DNA-Replikation auffüllt (Chen, 1996; Ishikawa and Naito, 1999; Lezhava et al., 1995). *Borrelia*-Spezies verfügen über kovalent geschlossene Hairpin-Strukturen an den Enden ihrer linearen Chromosomen, die nach der Replikation durch ein entsprechendes Enzym aufgelöst werden, sodass zwei lineare DNA-Moleküle entstehen (Casjens et al., 1997; Kitten and Barbour, 1990). Der *E. coli* Phage N15 nutzt ein ähnliches System, *TelN/tos*, bestehend aus der pallindromischen DNA-Sequenz *tos* und einer bakteriellen Telomerase *TelN*, zur Prozessierung seines linearen Genoms (Ravin et al., 2003). Mit Hilfe dieses Systems konnte das zirkuläre *E. coli* Chromosom erfolgreich in ein lineares umstrukturiert werden (Cui et al., 2007; Liang et al., 2013). Für die Prozessierung des linearen Chromosoms ist die Expression von *TelN* essentiell (Cui et al., 2007; Ravin et al., 2001). Dieser zusätzliche Faktor kostet die Zelle im Gegensatz zu einem zirkulären Chromosom mehr Energie. Jedoch bietet das lineare Design eines Chromosoms auch Vorteile, denn Systeme, die ein zirkuläres Chromosom benötigt, sind möglicherweise für lineare Chromosomen nicht essentiell. Durch die DNA-Replikation können Multimere der DNA-Moleküle entstehen, die für eine erfolgreiche Vererbung der DNA zunächst getrennt werden müssen (Lesterlin et al., 2004). Nach der Replikation von zirkulären Chromosomen können Catenane entstehen, die an der *dif*-site, welche sich in der Terminationsdomäne befindet, durch den Proteinkomplex der Rekombinasen *XerCD* aufgelöst werden (Blakely et al., 1993; Crozat et al., 2014). Cui *et al.* konnten zeigen, dass ein *E. coli* Stamm mit einem linearen Chromosom im Gegensatz zum Wildtyp-Stamm keine Wachstumsdefizite aufwies, wenn keine bzw. eine mutierte *dif*-site auf dem Chromosom vorhanden war (Cui et al., 2007). Am Beispiel von linearen Chromosomen von *B. burgdorferi* konnte gezeigt werden, dass während der DNA-Replikation zirkuläre Zwischenprodukte entstehen, die durch die Telomerase aufgelöst werden (Chaconas and Kobryn, 2010). Eine *dif*-site wird auf linearen Chromosomen demnach nicht benötigt. Dies ist konsistent mit Sequenzdaten von natürlich vorkommenden Bakterien mit linearen Genomen (Bentley et al., 2002; Fraser et al., 1997). Ein lineares Chromosom sollte demnach auch keine Faktoren benötigen, die dafür verantwortlich sind, die *dif*-site auf dem zirkulären Chromosom zu finden. Die Rekombinasen *XerCD* werden von der Translokase *FtsK* auf dem zirkulären Chromosom von *E. coli* geführt. *FtsK* bindet an Sequenzmotive, sogenannte KOPS (*FtsK* Orienting Polar Sequence), die vom Replikationsursprung zur Terminusregion orientiert angeordnet sind, sodass *FtsK* die Rekombinasen *XerCD* gezielt zum Terminus leiten kann (Bigot et al., 2006). Ein Fehlen oder eine fehlerhafte Anordnung der KOPS kann dazu führen, dass *XerCD* nicht rekrutiert wird und Dimere nicht mehr gelöst werden. Dies kann zu einer ungleichen Verteilung der zirkulären Chromosomen führen (Bigot et al., 2005; Liu et al., 1998). Sowohl eine *dif*-site als auch das System um die *dif*-site auf dem Chromosom zu identifizieren, werden bei einem linearen Chromosom nicht benötigt. Somit wird deutlich, dass die

Anforderungen an lineare und zirkuläre Chromosomen durchaus unterschiedlich sein können. Ein System, welches von linearen und zirkulären Chromosomen verwendet werden sollte, ist das „Nucleoid occlusion“-System, welches die Zellteilung unterdrückt, solange die DNA-Replikation noch nicht abgeschlossen ist (Bernhardt and de Boer, 2005; Mannik et al., 2012). Das „Nucleoid occlusion“-System besteht aus dem Protein SImA (Selective lethal with mutation in Min system) und der zugehörigen Bindestelle (Bernhardt and de Boer, 2005). SImA bindet überall auf dem Chromosom, jedoch nicht in der Terminationsregion (Cho et al., 2011; Tonthat et al., 2011). Erst wenn sich die Terminationsregion, die nicht von SImA gebunden wird, in der Zellmitte befindet, kann die Einschnürung der Membran und somit die Zellteilung eingeleitet werden (Bernhardt and de Boer, 2005; Tonthat et al., 2011). Für lineare und zirkuläre Chromosomen ist dieses System notwendig um Tochterzellen mit komplett replizierten Chromosomen zu erhalten. Gleiches gilt für Systeme, wie beispielsweise Chi-sites die von dem Helikase-Nuklease Komplex RecBCD erkannt werden und von hoher Bedeutung für die Doppelstrangbruch-Reparatur und homologe Rekombination sind (Chaudhury and Smith, 1984; Malone et al., 1978; Smith, 2012). Nach der DNA-Replikation müssen die Chromosomen auf die Tochterzellen verteilt werden, was als Segregation bezeichnet wird. In viele Bakterien wie auch in *V. cholerae* und *C. crescentus* werden die Chromosomen durch ein par-System, ein Partitionierungssystem, segregiert (Heidelberg et al., 2000; Mohl and Gober, 1997). In *E. coli* konnte kein par-System identifiziert werden. Wie die Segregation des *E. coli* Chromosoms erfolgt, ist noch nicht verstanden. Jedoch gibt es verschiedenen Modelle, die versuchen den Segregationsmechanismus unabhängig von einem Partitionierungssystem zu erklären (Di Ventura et al., 2013; Fisher et al., 2013; Jacob, 1963; Lemon and Grossman, 2001; Wang and Sherratt, 2010). Auch wenn der Segregationsmechanismus in *E. coli* noch nicht aufgeklärt ist, konnte bereits gezeigt werden, dass auch lineare Chromosomen in *E. coli* stabil vererbt und somit erfolgreich segregiert werden können (Cui et al., 2007; Liang et al., 2013). Theoretisch sollte es deshalb ebenfalls möglich sein 16 lineare Chromosomen in *E. coli* zu segregieren. Es muss jedoch beachtet werden, dass ein deutlicher Unterschied darin besteht, ein oder 16 Chromosomen zu replizieren und zu segregieren. Bisherige Studien konnten zeigen, dass ein zweites Replikon in *E. coli* stabil weitervererbt wird (Messerschmidt et al., 2015). Dieses war bei weitem keine triviale Aufgabe. Einen *E. coli* Stamm zu generieren, der die achtfache Anzahl an Chromosomen enthält, wäre demnach sehr ambitioniert.

Aufgrund der sehr geringen Menge an bakteriellen Spezies, die lineare Chromosomen tragen, liegt es nah, dass ein evolutionärer Druck vorliegt, der die zirkuläre Organisation bakterieller Chromosomen begünstigt. Bakterien durchlaufen keine Meiose und führen horizontalen Gentransfer durch, zudem sind die Generationszeiten im Vergleich zu Eukaryoten sehr gering. Diese Faktoren könnten begründen, warum Eukaryoten lineare und Prokaryoten zirkuläre Chromosomen prozessieren. Gründe

lineare Chromosomen in Bakterien zu limitieren, könnten mit den aufgezeigten Ansätzen der Synthetischen Biologie ganz nach dem Leitgedanken von Richard Feynman

“What I cannot create, I do not understand“

ergründet werden.

5. Zusammenfassung

Eukaryoten und Prokaryoten unterscheiden sich in der Organisation ihrer Genome. Prokaryoten tragen mehrheitlich ein zirkuläres Chromosom, welches von einem Replikationsursprung repliziert wird. Das eukaryotische Genom hingegen umfasst mehrere lineare Chromosomen, die mehr als einen Replikationsursprung enthalten. In wie weit die prokaryotische Genomorganisation der eukaryotischen angepasst werden kann, ist eine spannende Frage, von deren Antwort grundlegende Konstruktionsregeln für die Entwicklung von synthetischen Chromosomen abgeleitet werden könnten. In der vorliegenden Arbeit wurden *Escherichia coli* Stämme mit multiplen Replikationsursprüngen und Chromosomen bezüglich ihrer DNA-Replikation untersucht und es wurde eine fluoreszenzbasierte Einzelzell-Methode designt und etabliert, die eine verbesserte „signal to background ratio“ aufweist. Die Ergebnisse der vorliegenden Arbeit können wie folgt zusammengefasst werden:

I. *E. coli* toleriert einen zusätzlichen Replikationsursprung auf dem Chromosom, der aktiv repliziert. Dieser Ansatz funktioniert sowohl mit einer zusätzlichen Kopie von *oriC*, als auch mit *orill*, dem Replikationsursprung des zweiten Chromosoms von *V. cholerae*. Es konnte gezeigt werden, dass eine verkürzte *oriC*-Sequenz, die keine DnaA-Box R4 enthält, nicht ausreichend für die DNA-Replikation auf dem nativen Chromosom ist. Wird *orill* als extra Replikationsursprung eingesetzt, verläuft die DNA-Replikation dysreguliert. Diese Dysregulation wird aufgehoben, sobald sich *oriC* und *orill* nicht mehr auf dem gleichen Chromosom, sondern auf zwei unterschiedlichen Replikons befinden.

II. Für das Design und die Assemblierung von DNA-Sequenzbibliotheken zur chromosomalen Integration in Bakterien wurde das Programm MARSeG (Motif Avoiding Randomized Sequences Generator) geschrieben und das MoClo-System optimiert. Mit dem Sequenzgenerator MARSeG, ist es möglich Sequenzen zu designen, die eine hohe Variabilität aufweisen und gleichzeitig bestimmte DNA-Motive ausschließen. Das MoClo-System wurde optimiert um die Klonierungseffizienz zu erhöhen, Assemblierungen im Bibliotheken-Maßstab zu vereinfachen und eine Genomintegration assemblierter DNA-Fragmente einfacher zu gestalten. Mit Hilfe von MARSeG und den modifizierten MoClo-Vektoren konnte das Design und der Aufbau eines LacO/TetO hybrid FROS-Arrays effizient und schnell vollzogen werden.

III. Es wurde eine neue Einzelzell-Methode etabliert mit der das Hintergrundsignal konventioneller FROS (Fluorescence Repressor Operator System) stark reduziert wird. Die neue Methode BiFCROS (Bimolecular Fluorescence Complementation and Repressor Operator System) basiert auf Fusionen aus Repressoren und fragmentierten Fluoreszenzproteinen, die an ein hybrid FROS-Array binden, wodurch es durch Bimolekulare Fluoreszenzkomplementation zu einem Fluoreszenzsignal kommt. BiFCROS

enthält im Vergleich zum konventionellen FROS eine deutlich verminderte Hintergrundfluoreszenz, weil nur Proteine fluoreszieren, die an das Operator-Array gebunden sind. Um zu analysieren, ob BiFCROS für die Kopienzahlbestimmung von Genloki eingesetzt werden könnte, wurde es in diese Richtung weiterentwickelt. Anhand der Fluoreszenzintensität der gesamten Zelle könnte Rückschluss auf die Kopienzahl des hybrid FROS-Arrays gezogen werden, weil diese mit Anstieg der Kopienzahl proportional zunehmen sollte.

5. Summary

Eukaryotes and prokaryotes differ in the organization of their genomes. Most prokaryotes contain a single, circular chromosome that is replicated by one single replication origin. Contrary, eukaryotes possess multiple, linear chromosomes, which are replicated by more than one replication origin. Whether, and to what extent, *Escherichia coli* tolerates an eukaryotic genome organization is an exciting question whose answer could help to develop construction rules for synthetic chromosomes. This thesis presents DNA replication studies of *E. coli* strains with multiple-origins on the chromosome and an *E. coli* strain with two linear chromosomes. Additionally, a fluorescence-based single-cell method has been designed and established, which has an improved "signal to background ratio". The main results can be summarized as follows:

I. *E. coli* tolerates additional replication origins on the native chromosome, which are able to initiate DNA replication. DNA replication activity of extra copies of *oriC* or *oriII*, the replication origin of the second chromosome of *V. cholerae*, could be verified. A shortened sequence of *oriC*, which does not contain DnaA box R4, is not sufficient to initiate DNA replication on the native chromosome. If *oriII* is used as additional replication origin, DNA replication shows dysregulated patterns, which are abolished if *oriC* and *oriII* locate on two different replicons within the cell.

II. For design and assembly of DNA sequence libraries for chromosomal integration into bacteria the program MARSeG (Motif Avoiding Randomized Sequence Generator) was written and the modular cloning system MoClo was modified. MARSeG allows the design of sequences with a high variability and the simultaneous exclusion of certain motifs. MoClo has been optimized to increase cloning efficiency, to simplify assembly operations in library scale and to facilitate genomic integration of assembled DNA fragments. MARSeG and the modified MoClo vector were used to design and construct a LacO/TetO hybrid FROS array.

III. A low background FROS (Fluorescence Repressor Operator System) was designed and established. The new method BiFCROS (Bimolecular Fluorescence complementation and Repressor Operator System) is based on fusions of repressor proteins with a split fluorescent protein, which bind to a hybrid FROS array resulting in fluorescence signals due to bimolecular fluorescence complementation. Background fluorescence is reduced, because only proteins that are bound to the hybrid FROS array fluoresce. In order to analyze whether BiFCROS could be used for copy number determination of gene loci, it was further developed in this direction. Thus, total cell fluorescence intensity should reveal the copy number of the hybrid FROS array, because it should increase proportionally with increasing copy number.

6. Literaturverzeichnis

- Allardet-Servent, A., Michaux-Charachon, S., Jumas-Bilak, E., Karayan, L. and Ramuz, M., 1993. Presence of one linear and one circular chromosome in the *Agrobacterium tumefaciens* C58 genome. *J Bacteriol.* 175, 7869-74.
- Annaluru, N., Muller, H., Mitchell, L.A., Ramalingam, S., Stracquadiano, G., Richardson, S.M., Dymond, J.S., Kuang, Z., Scheifele, L.Z., Cooper, E.M., Cai, Y., Zeller, K., Agmon, N., Han, J.S., Hadjithomas, M., Tullman, J., Caravelli, K., Cirelli, K., Guo, Z., London, V., Yeluru, A., Murugan, S., Kandavelou, K., Agier, N., Fischer, G., Yang, K., Martin, J.A., Bilgel, M., Bohutski, P., Boulier, K.M., Capaldo, B.J., Chang, J., Charoen, K., Choi, W.J., Deng, P., DiCarlo, J.E., Doong, J., Dunn, J., Feinberg, J.I., Fernandez, C., Floria, C.E., Gladowski, D., Hadidi, P., Ishizuka, I., Jabbari, J., Lau, C.Y., Lee, P.A., Li, S., Lin, D., Linder, M.E., Ling, J., Liu, J., London, M., Ma, H., Mao, J., McDade, J.E., McMillan, A., Moore, A.M., Oh, W.C., Ouyang, Y., Patel, R., Paul, M., Paulsen, L.C., Qiu, J., Rhee, A., Rubashkin, M.G., Soh, I.Y., Sotuyo, N.E., Srinivas, V., Suarez, A., Wong, A., Wong, R., Xie, W.R., Xu, Y., Yu, A.T., Koszul, R., Bader, J.S., Boeke, J.D. and Chandrasegaran, S., 2014. Total synthesis of a functional designer eukaryotic chromosome. *Science.* 344, 55-8.
- Ansaruzzaman, M., Bhuiyan, N.A., Nair, B.G., Sack, D.A., Lucas, M., Deen, J.L., Ampuero, J. and Chaignat, C.L., 2004. Cholera in Mozambique, variant of *Vibrio cholerae*. *Emerg Infect Dis.* 10, 2057-9.
- Asai, T., Bates, D.B., Boye, E. and Kogoma, T., 1998. Are minichromosomes valid model systems for DNA replication control? Lessons learned from *Escherichia coli*. *Mol Microbiol.* 29, 671-5.
- Baker, M., 2012. Digital PCR hits its stride. *Nat Methods.* 9, 541-544.
- Bartosik, D., 2001. Bacterial plasmid stability. *Postepy Biochem.* 47, 138-45.
- Bartosik, D., Szymanik, M. and Wysocka, E., 2001. Identification of the partitioning site within the repABC-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. *J Bacteriol.* 183, 6234-43.
- Bates, D.B., Asai, T., Cao, Y., Chambers, M.W., Cadwell, G.W., Boye, E. and Kogoma, T., 1995. The DnaA box R4 in the minimal *oriC* is dispensable for initiation of *Escherichia coli* chromosome replication. *Nucleic Acids Res.* 23, 3119-25.
- Bates, D.B., Boye, E., Asai, T. and Kogoma, T., 1997. The absence of effect of *gid* or *mioC* transcription on the initiation of chromosomal replication in *Escherichia coli*. *Proc Natl Acad Sci U S A.* 94, 12497-502.
- Beckmann, J.S., Estivill, X. and Antonarakis, S.E., 2007. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat Rev Genet.* 8, 639-46.
- Bell, S.P. and Dutta, A., 2002. DNA replication in eukaryotic cells. *Annu Rev Biochem.* 71, 333-74.
- Bendich, A.J. and Drlica, K., 2000. Prokaryotic and eukaryotic chromosomes: what's the difference? *Bioessays.* 22, 481-6.
- Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J. and Hopwood, D.A., 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature.* 417, 141-7.
- Bernard, C.S., Sadasivam, M., Shiomi, D. and Margolin, W., 2007. An altered FtsA can compensate for the loss of essential cell division protein FtsN in *Escherichia coli*. *Mol Microbiol.* 64, 1289-305.
- Bernhardt, T.G. and de Boer, P.A., 2005. SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over chromosomes in *E. coli*. *Mol Cell.* 18, 555-64.
- Bigot, S., Saleh, O.A., Cornet, F., Allemand, J.F. and Barre, F.X., 2006. Oriented loading of FtsK on KOPS. *Nat Struct Mol Biol.* 13, 1026-8.

- Bigot, S., Saleh, O.A., Lesterlin, C., Pages, C., El Karoui, M., Dennis, C., Grigoriev, M., Allemand, J.F., Barre, F.X. and Cornet, F., 2005. KOPS: DNA motifs that control *E. coli* chromosome segregation by orienting the FtsK translocase. *EMBO J.* 24, 3770-80.
- Bird, R.E., Louarn, J., Martuscelli, J. and Caro, L., 1972. Origin and sequence of chromosome replication in *Escherichia coli*. *J Mol Biol.* 70, 549-66.
- Blackburn, E.H., 1991. Structure and function of telomeres. *Nature.* 350, 569-73.
- Blainey, P.C., 2013. The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol Rev.* 37, 407-27.
- Blakely, G., May, G., McCulloch, R., Arciszewska, L.K., Burke, M., Lovett, S.T. and Sherratt, D.J., 1993. Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. *Cell.* 75, 351-61.
- Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y., 1997. The complete genome sequence of *Escherichia coli* K-12. *Science.* 277, 1453-62.
- Boye, E., Lobner-Olesen, A. and Skarstad, K., 2000. Limiting DNA replication to once and only once. *EMBO Rep.* 1, 479-83.
- Bregeon, D., Colot, V., Radman, M. and Taddei, F., 2001. Translational misreading: a tRNA modification counteracts a +2 ribosomal frameshift. *Genes Dev.* 15, 2295-306.
- Bremer, H. and Churchward, G., 1977. An examination of the Cooper-Helmstetter theory of DNA replication in bacteria and its underlying assumptions. *J Theor Biol.* 69, 645-54.
- Buck, S.B., Bradford, J., Gee, K.R., Agnew, B.J., Clarke, S.T. and Salic, A., 2008. Detection of S-phase cell cycle progression using 5-ethynyl-2'-deoxyuridine incorporation with click chemistry, an alternative to using 5-bromo-2'-deoxyuridine antibodies. *Biotechniques.* 44, 927-9.
- Casjens, S., Delange, M., Ley, H.L., 3rd, Rosa, P. and Huang, W.M., 1995. Linear chromosomes of Lyme disease agent *spirochetes*: genetic diversity and conservation of gene order. *J Bacteriol.* 177, 2769-80.
- Casjens, S., Murphy, M., DeLange, M., Sampson, L., van Vugt, R. and Huang, W.M., 1997. Telomeres of the linear chromosomes of Lyme disease spirochaetes: nucleotide sequence and possible exchange with linear plasmid telomeres. *Mol Microbiol.* 26, 581-96.
- Chaconas, G. and Kobryn, K., 2010. Structure, function, and evolution of linear replicons in *Borrelia*. *Annu Rev Microbiol.* 64, 185-202.
- Chao, R., Yuan, Y. and Zhao, H., 2014. Recent advances in DNA assembly technologies. *FEMS Yeast Res.* DOI: 10.1111/1567-1364.12171.
- Chaudhury, A.M. and Smith, G.R., 1984. A new class of *Escherichia coli* recBC mutants: implications for the role of RecBC enzyme in homologous recombination. *Proc Natl Acad Sci U S A.* 81, 7850-4.
- Chen, C.W., 1996. Complications and implications of linear bacterial chromosomes. *Trends Genet.* 12, 192-6.
- Cheung, V.G. and Nelson, S.F., 1996. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. *Proc Natl Acad Sci U S A.* 93, 14676-9.
- Cho, H., McManus, H.R., Dove, S.L. and Bernhardt, T.G., 2011. Nucleoid occlusion factor SlmA is a DNA-activated FtsZ polymerization antagonist. *Proc Natl Acad Sci U S A.* 108, 3773-8.
- Choi, Y.J. and Lee, S.Y., 2013. Microbial production of short-chain alkanes. *Nature.* 502, 571-4.
- Cobb, R.E., Ning, J.C. and Zhao, H., 2014. DNA assembly techniques for next-generation combinatorial biosynthesis of natural products. *J Ind Microbiol Biotechnol.* 41, 469-77.
- Cohen, S.N., Chang, A.C., Boyer, H.W. and Helling, R.B., 1973. Construction of biologically functional bacterial plasmids *in vitro*. *Proc Natl Acad Sci U S A.* DOI: 10.1128/microbiolspec.PLAS-0025-2014. 70, 3240-4.
- Crozat, E., Fournes, F., Cornet, F., Hallet, B. and Rousseau, P., 2014. Resolution of multimeric forms of circular plasmids and chromosomes. *Microbiol Spectr.* 2.

- Cui, T., Moro-oka, N., Ohsumi, K., Kodama, K., Ohshima, T., Ogasawara, N., Mori, H., Wanner, B., Niki, H. and Horiuchi, T., 2007. *Escherichia coli* with a linear genome. *EMBO Rep.* 8, 181-7.
- Cunha, S., Odijk, T., Suleymanoglu, E. and Woldringh, C.L., 2001. Isolation of the *Escherichia coli* nucleoid. *Biochimie.* 83, 149-54.
- Darzynkiewicz, Z., Traganos, F., Zhao, H., Halicka, H.D. and Li, J., 2011. Cytometry of DNA replication and RNA synthesis: Historical perspective and recent advances based on "click chemistry". *Cytometry A.* 79, 328-37.
- Datta, N. and Hedges, R.W., 1972. Host ranges of R factors. *J Gen Microbiol.* 70, 453-60.
- de Bourcy, C.F., De Vlaminck, I., Kanbar, J.N., Wang, J., Gawad, C. and Quake, S.R., 2014. A quantitative comparison of single-cell whole genome amplification methods. *PLoS One.* 9, e105585.
- De Gelder, L., Ponciano, J.M., Joyce, P. and Top, E.M., 2007. Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology.* 153, 452-63.
- de Kok, S., Stanton, L.H., Slaby, T., Durot, M., Holmes, V.F., Patel, K.G., Platt, D., Shapland, E.B., Serber, Z., Dean, J., Newman, J.D. and Chandran, S.S., 2014. Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synth Biol.* 3, 97-106.
- Dean, F.B., Nelson, J.R., Giesler, T.L. and Lasken, R.S., 2001. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* 11, 1095-9.
- del Solar, G., Alonso, J.C., Espinosa, M. and Diaz-Orejas, R., 1996. Broad-host-range plasmid replication: an open question. *Mol Microbiol.* 21, 661-6.
- del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M. and Diaz-Orejas, R., 1998. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev.* 62, 434-64.
- Di Ventura, B., Knecht, B., Andreas, H., Godinez, W.J., Fritsche, M., Rohr, K., Nickel, W., Heermann, D.W. and Sourjik, V., 2013. Chromosome segregation by the *Escherichia coli* Min system. *Mol Syst Biol.* 9, 686.
- Dimude, J.U., Stockum, A., Midgley-Smith, S.L., Upton, A.L., Foster, H.A., Khan, A., Saunders, N.J., Retkute, R. and Rudolph, C.J., 2015. The consequences of replicating in the wrong orientation: Bacterial chromosome duplication without an active replication origin. *MBio.* 6, e01294-15.
- Ducret, A., Quardokus, E.M. and Brun, Y.V., 2016. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nat Microbiol.* 1, 16077.
- Duggin, I.G., Wake, R.G., Bell, S.D. and Hill, T.M., 2008. The replication fork trap and termination of chromosome replication. *Mol Microbiol.* 70, 1323-33.
- Duigou, S., Knudsen, K.G., Skovgaard, O., Egan, E.S., Lobner-Olesen, A. and Waldor, M.K., 2006. Independent control of replication initiation of the two *Vibrio cholerae* chromosomes by DnaA and RctB. *J Bacteriol.* 188, 6419-24.
- Egan, E.S., Fogel, M.A. and Waldor, M.K., 2005. Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol Microbiol.* 56, 1129-38.
- Egan, E.S., Lobner-Olesen, A. and Waldor, M.K., 2004. Synchronous replication initiation of the two *Vibrio cholerae* chromosomes. *Curr Biol.* 14, R501-2.
- Egan, E.S. and Waldor, M.K., 2003. Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell.* 114, 521-30.
- Ellis, T., Adie, T. and Baldwin, G.S., 2011. DNA assembly for synthetic biology: from parts to pathways and beyond. *Integr Biol (Camb).* 3, 109-18.
- Engler, C., Gruetzner, R., Kandzia, R. and Marillonnet, S., 2009. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One.* 4, e5553.
- Engler, C., Kandzia, R. and Marillonnet, S., 2008. A one pot, one step, precision cloning method with high throughput capability. *PLoS One.* 3, e3647.
- Ferdows, M.S. and Barbour, A.G., 1989. Megabase-sized linear DNA in the bacterium *Borrelia burgdorferi*, the Lyme disease agent. *Proc Natl Acad Sci U S A.* 86, 5969-73.
- Ferullo, D.J., Cooper, D.L., Moore, H.R. and Lovett, S.T., 2009. Cell cycle synchronization of *Escherichia coli* using the stringent response, with fluorescence labeling assays for DNA content and replication. *Methods.* 48, 8-13.

- Feuk, L., Carson, A.R. and Scherer, S.W., 2006. Structural variation in the human genome. *Nat Rev Genet.* 7, 85-97.
- Filutowicz, M., Ross, W., Wild, J. and Gourse, R.L., 1992. Involvement of Fis protein in replication of the *Escherichia coli* chromosome. *J Bacteriol.* 174, 398-407.
- Fisher, J.K., Bourniquel, A., Witz, G., Weiner, B., Prentiss, M. and Kleckner, N., 2013. Four-dimensional imaging of *E. coli* nucleoid organization and dynamics in living cells. *Cell.* 153, 882-95.
- Frage, B., Dohlemann, J., Robledo, M., Lucena, D., Sobetzko, P., Graumann, P.L. and Becker, A., 2016. Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle. *Mol Microbiol.* 100, 808-23.
- Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R., Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K., Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S., Hanson, M., van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J., Weidman, J., Utterback, T., Watthey, L., McDonald, L., Artiach, P., Bowman, C., Garland, S., Fuji, C., Cotton, M.D., Horst, K., Roberts, K., Hatch, B., Smith, H.O. and Venter, J.C., 1997. Genomic sequence of a Lyme disease *spirochaete*, *Borrelia burgdorferi*. *Nature.* 390, 580-6.
- Fuller, R.S., Funnell, B.E. and Kornberg, A., 1984. The dnaA protein complex with the *E. coli* chromosomal replication origin (*oriC*) and other DNA sites. *Cell.* 38, 889-900.
- Galperin, M.Y., 2007. Linear chromosomes in bacteria: no straight edge advantage? *Environ Microbiol.* 9, 1357-62.
- Gao, F., 2015. Bacteria may have multiple replication origins. *Front Microbiol.* 6, 324.
- Gao, F. and Zhang, C.T., 2007. DoriC: a database of *oriC* regions in bacterial genomes. *Bioinformatics.* 23, 1866-7.
- Gawad, C., Koh, W. and Quake, S.R., 2016. Single-cell genome sequencing: current state of the science. *Nat Rev Genet.* 17, 175-88.
- Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., Stockwell, T.B., Brownley, A., Thomas, D.W., Algire, M.A., Merryman, C., Young, L., Noskov, V.N., Glass, J.I., Venter, J.C., Hutchison, C.A., 3rd and Smith, H.O., 2008a. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science.* 319, 1215-20.
- Gibson, D.G., Benders, G.A., Axelrod, K.C., Zaveri, J., Algire, M.A., Moodie, M., Montague, M.G., Venter, J.C., Smith, H.O. and Hutchison, C.A., 3rd, 2008b. One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc Natl Acad Sci U S A.* 105, 20404-9.
- Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.Y., Algire, M.A., Benders, G.A., Montague, M.G., Ma, L., Moodie, M.M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E.A., Young, L., Qi, Z.Q., Segall-Shapiro, T.H., Calvey, C.H., Parmar, P.P., Hutchison, C.A., 3rd, Smith, H.O. and Venter, J.C., 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science.* 329, 52-6.
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd and Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* 6, 343-5.
- Gille, H., Egan, J.B., Roth, A. and Messer, W., 1991. The FIS protein binds and bends the origin of chromosomal DNA replication, *oriC*, of *Escherichia coli*. *Nucleic Acids Res.* 19, 4167-72.
- Gonzalez-Roca, E., Garcia-Albeniz, X., Rodriguez-Mulero, S., Gomis, R.R., Kornacker, K. and Auer, H., 2010. Accurate expression profiling of very small cell populations. *PLoS One.* 5, e14418.
- Gordon, G.S., Sitnikov, D., Webb, C.D., Teleman, A., Straight, A., Losick, R., Murray, A.W. and Wright, A., 1997. Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell.* 90, 1113-21.
- Govindarajan, R., Duraiyan, J., Kaliyappan, K. and Palanisamy, M., 2012. Microarray and its applications. *J Pharm Bioallied Sci.* 4, S310-2.
- Grimwade, J.E., Ryan, V.T. and Leonard, A.C., 2000. IHF redistributes bound initiator protein, DnaA, on supercoiled *oriC* of *Escherichia coli*. *Mol Microbiol.* 35, 835-44.

- Grindberg, R.V., Ishoey, T., Brinza, D., Esquenazi, E., Coates, R.C., Liu, W.T., Gerwick, L., Dorrestein, P.C., Pevzner, P., Lasken, R. and Gerwick, W.H., 2011. Single cell genome amplification accelerates identification of the apratoxin biosynthetic pathway from a complex microbial assemblage. *PLoS One*. 6, e18565.
- Grunewald, J., Hunt, G.S., Dong, L., Niessen, F., Wen, B.G., Tsao, M.L., Perera, R., Kang, M., Laffitte, B.A., Azarian, S., Ruf, W., Nasoff, M., Lerner, R.A., Schultz, P.G. and Smider, V.V., 2009. Mechanistic studies of the immunochemical termination of self-tolerance with unnatural amino acids. *Proc Natl Acad Sci U S A*. 106, 4337-42.
- Haber, J.E., Thorburn, P.C. and Rogers, D., 1984. Meiotic and mitotic behavior of dicentric chromosomes in *Saccharomyces cerevisiae*. *Genetics*. 106, 185-205.
- Hansen, F.G., Christensen, B.B. and Atlung, T., 2007. Sequence characteristics required for cooperative binding and efficient *in vivo* titration of the replication initiator protein DnaA in *E. coli*. *Journal of Molecular Biology*. 367, 942-52.
- Hansen, F.G., Christensen, B.B., Nielsen, C.B. and Atlung, T., 2006. Insights into the quality of DnaA boxes and their cooperativity. *Journal of Molecular Biology*. 355, 85-95.
- Harrison, P.W., Lower, R.P., Kim, N.K. and Young, J.P., 2010. Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol*. 18, 141-8.
- Hawkins, M., Malla, S., Blythe, M.J., Nieduszynski, C.A. and Allers, T., 2013. Accelerated growth in the absence of DNA replication origins. *Nature*. 503, 544-7.
- Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R.D., Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R., Mekalanos, J.J., Venter, J.C. and Fraser, C.M., 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*. 406, 477-83.
- Hill, N.S., Kadoya, R., Chatteraj, D.K. and Levin, P.A., 2012. Cell size and the initiation of DNA replication in bacteria. *PLoS Genet*. 8, e1002549.
- Hill, T.M., Henson, J.M. and Kuempel, P.L., 1987. The terminus region of the *Escherichia coli* chromosome contains two separate loci that exhibit polar inhibition of replication. *Proc Natl Acad Sci U S A*. 84, 1754-8.
- Hinnebusch, J. and Tilly, K., 1993. Linear plasmids and chromosomes in bacteria. *Mol Microbiol*. 10, 917-22.
- Hiraga, S., 1976. Novel F prime factors able to replicate in *Escherichia coli* Hfr strains. *Proc Natl Acad Sci U S A*. 73, 198-202.
- Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K. and Pease, L.R., 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*. 77, 61-8.
- Huang, L., Ma, F., Chapman, A., Lu, S. and Xie, X.S., 2015. Single-cell Whole-Genome Amplification and sequencing: Methodology and applications. *Annu Rev Genomics Hum Genet*. 16, 79-102.
- Hutchison, C.A., 3rd, Chuang, R.Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L., Pelletier, J.F., Qi, Z.Q., Richter, R.A., Strychalski, E.A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K.S., Smith, H.O., Glass, J.I., Merryman, C., Gibson, D.G. and Venter, J.C., 2016. Design and synthesis of a minimal bacterial genome. *Science*. 351, aad6253.
- Ishikawa, F. and Naito, T., 1999. Why do we have linear chromosomes? A matter of Adam and Eve. *Mutat Res*. 434, 99-107.
- Itaya, M., Fujita, K., Kuroki, A. and Tsuge, K., 2008. Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat Methods*. 5, 41-3.
- Itaya, M. and Tanaka, T., 1991. Complete physical map of the *Bacillus subtilis* 168 chromosome constructed by a gene-directed mutagenesis method. *J Mol Biol*. 220, 631-48.
- Itaya, M. and Tanaka, T., 1997. Experimental surgery to create subgenomes of *Bacillus subtilis* 168. *Proc Natl Acad Sci U S A*. 94, 5378-82.
- Ivanova, D., Taylor, T., Smith, S.L., Dimude, J.U., Upton, A.L., Mehrjouy, M.M., Skovgaard, O., Sherratt, D.J., Retkute, R. and Rudolph, C.J., 2015. Shaping the landscape of the *Escherichia coli*

- chromosome: replication-transcription encounters in cells with an ectopic replication origin. *Nucleic Acids Res.* 43, 7865-77.
- Jacob, F., Brenner S., Cuzin F., 1963. On the regulation of DNA replication in bacteria. *Cold Spring Harb Symp Quant Biol.* 28, 329-348.
- Jain, A. and Srivastava, P., 2013. Broad host range plasmids. *FEMS Microbiol Lett.* 348, 87-96.
- Jha, J.K., Baek, J.H., Venkova-Canova, T. and Chatteraj, D.K., 2012. Chromosome dynamics in multichromosome bacteria. *Biochim Biophys Acta.* 1819, 826-9.
- Joshi, M.C., Bourniquel, A., Fisher, J., Ho, B.T., Magnan, D., Kleckner, N. and Bates, D., 2011. *Escherichia coli* sister chromosome separation includes an abrupt global transition with concomitant release of late-splitting intersister snaps. *Proc Natl Acad Sci U S A.* 108, 2765-70.
- Kaguni, J.M., 1997. *Escherichia coli* DnaA protein: the replication initiator. *Mol Cells.* 7, 145-57.
- Kalisky, T., Blainey, P. and Quake, S.R., 2011. Genomic analysis at the single-cell level. *Annu Rev Genet.* 45, 431-45.
- Kano, Y., Ogawa, T., Ogura, T., Hiraga, S., Okazaki, T. and Imamoto, F., 1991. Participation of the histone-like protein HU and of IHF in minichromosomal maintenance in *Escherichia coli*. *Gene.* 103, 25-30.
- Kawakami, H. and Katayama, T., 2010. DnaA, ORC, and Cdc6: similarity beyond the domains of life and diversity. *Biochem Cell Biol.* 88, 49-62.
- Kerppola, T.K., 2008. Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu Rev Biophys.* 37, 465-87.
- Kitten, T. and Barbour, A.G., 1990. Juxtaposition of expressed variable antigen genes with a conserved telomere in the bacterium *Borrelia hermsii*. *Proc Natl Acad Sci U S A.* 87, 6077-81.
- Klein, D., 2002. Quantification using real-time PCR technology: applications and limitations. *Trends Mol Med.* 8, 257-60.
- Kogoma, T., 1997. Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol Mol Biol Rev.* 61, 212-38.
- Kolling, R., Gielow, A., Seufert, W., Kucherer, C. and Messer, W., 1988. AsnC, a multifunctional regulator of genes located around the replication origin of *Escherichia coli*, *oriC*. *Mol Gen Genet.* 212, 99-104.
- Kortmann, H., Blank, L.M. and Schmid, A., 2011. Single cell analytics: an overview. *Adv Biochem Eng Biotechnol.* 124, 99-122.
- Kosuri, S. and Church, G.M., 2014. Large-scale de novo DNA synthesis: technologies and applications. *Nat Methods.* 11, 499-507.
- Kreuzer, K.N. and Brister, J.R., 2010. Initiation of bacteriophage T4 DNA replication and replication fork dynamics: a review in the *Virology Journal* series on bacteriophage T4 and its relatives. *Virology J.* 7, 358.
- Krishnapillai, V., 1988. Molecular genetic analysis of bacterial plasmid promiscuity. *FEMS Microbiol Rev.* 4, 223-37.
- Langer, U., Richter, S., Roth, A., Weigel, C. and Messer, W., 1996. A comprehensive set of DnaA-box mutations in the replication origin, *oriC*, of *Escherichia coli*. *Mol Microbiol.* 21, 301-11.
- Lau, I.F., Filipe, S.R., Soballe, B., Okstad, O.A., Barre, F.X. and Sherratt, D.J., 2003. Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol.* 49, 731-43.
- Lecault, V., White, A.K., Singhal, A. and Hansen, C.L., 2012. Microfluidic single cell analysis: from promise to practice. *Curr Opin Chem Biol.* 16, 381-90.
- Lemon, K.P. and Grossman, A.D., 2001. The extrusion-capture model for chromosome partitioning in bacteria. *Genes Dev.* 15, 2031-41.
- Leonard, A.C. and Helmstetter, C.E., 1986. Cell cycle-specific replication of *Escherichia coli* minichromosomes. *Proc Natl Acad Sci U S A.* 83, 5101-5.
- Lesterlin, C., Barre, F.X. and Cornet, F., 2004. Genetic recombination and the cell cycle: what we have learned from chromosome dimers. *Mol Microbiol.* 54, 1151-60.

- Lezhava, A., Mizukami, T., Kajitani, T., Kameoka, D., Redenbach, M., Shinkawa, H., Nimi, O. and Kinashi, H., 1995. Physical map of the linear chromosome of *Streptomyces griseus*. *J Bacteriol.* 177, 6492-8.
- Liang, X., Baek, C.H. and Katzen, F., 2013. *Escherichia coli* with two linear chromosomes. *ACS Synth Biol.* 2, 734-40.
- Lidstrom, M.E. and Konopka, M.C., 2010. The role of physiological heterogeneity in microbial population behavior. *Nat Chem Biol.* 6, 705-12.
- Lies, M., Visser, B.J., Joshi, M.C., Magnan, D. and Bates, D., 2015. MioC and GidA proteins promote cell division in *E. coli*. *Front Microbiol.* 6, 516.
- Lindas, A.C. and Bernander, R., 2013. The cell cycle of archaea. *Nat Rev Microbiol.* 11, 627-38.
- Lipps, G., 2008. Plasmids - Current Research and future Trends. Caister Academic Press.
- Liu, C.C. and Schultz, P.G., 2006. Recombinant expression of selectively sulfated proteins in *Escherichia coli*. *Nat Biotechnol.* 24, 1436-40.
- Liu, C.C. and Schultz, P.G., 2010. Adding new chemistries to the genetic code. *Annu Rev Biochem.* 79, 413-44.
- Liu, G., Draper, G.C. and Donachie, W.D., 1998. FtsK is a bifunctional protein involved in cell division and chromosome localization in *Escherichia coli*. *Mol Microbiol.* 29, 893-903.
- Lobner-Olesen, A., 1999. Distribution of minichromosomes in individual *Escherichia coli* cells: implications for replication control. *EMBO J.* 18, 1712-21.
- Lobner-Olesen, A., Atlung, T. and Rasmussen, K.V., 1987. Stability and replication control of *Escherichia coli* minichromosomes. *J Bacteriol.* 169, 2835-42.
- Lobner-Olesen, A. and Boye, E., 1992. Different effects of *mioC* transcription on initiation of chromosomal and minichromosomal replication in *Escherichia coli*. *Nucleic Acids Res.* 20, 3029-36.
- Maduiké, N.Z., Tehranchi, A.K., Wang, J.D. and Kreuzer, K.N., 2014. Replication of the *Escherichia coli* chromosome in RNase HI-deficient cells: multiple initiation regions and fork dynamics. *Mol Microbiol.* 91, 39-56.
- Malone, R.E., Chatteraj, D.K., Faulds, D.H., Stahl, M.M. and Stahl, F.W., 1978. Hotspots for generalized recombination in the *Escherichia coli* chromosome. *J Mol Biol.* 121, 473-91.
- Mannik, J., Wu, F., Hol, F.J., Bisicchia, P., Sherratt, D.J., Keymer, J.E. and Dekker, C., 2012. Robustness and accuracy of cell division in *Escherichia coli* in diverse cell shapes. *Proc Natl Acad Sci U S A.* 109, 6957-62.
- Marcus, J.S., Anderson, W.F. and Quake, S.R., 2006. Parallel picoliter rt-PCR assays using microfluidics. *Anal Chem.* 78, 956-8.
- Marsh, R.C. and Worcel, A., 1977. A DNA fragment containing the origin of replication of the *Escherichia coli* chromosome. *Proc Natl Acad Sci U S A.* 74, 2720-4.
- Mazodier, P. and Davies, J., 1991. Gene transfer between distantly related bacteria. *Annu Rev Genet.* 25, 147-71.
- McGarry, K.C., Ryan, V.T., Grimwade, J.E. and Leonard, A.C., 2004. Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP. *Proc Natl Acad Sci U S A.* 101, 2811-6.
- McHenry, C.S., 2011. Bacterial replicases and related polymerases. *Curr Opin Chem Biol.* 15, 587-94.
- Mechali, M., 2010. Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat Rev Mol Cell Biol.* 11, 728-38.
- Messer, W., 2002. The bacterial replication initiator DnaA. DnaA and *oriC*, the bacterial mode to initiate DNA replication. *FEMS Microbiol Rev.* 26, 355-74.
- Messer, W., Bergmans, H.E., Meijer, M., Womack, J.E., Hansen, F.G. and von Meyenburg, K., 1978. Mini-chromosomes: plasmids which carry the *E. coli* replication origin. *Mol Gen Genet.* 162, 269-75.
- Messer, W. and Weigel, C., 1996. Initiation of chromosome replication. In Neidhardt, F.C. (ed.), *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. ASM Press, Washington, DC, pp. 1579-1601.

- Messer, W. and Weigel, C., 1997. DnaA initiator--also a transcription factor. *Mol Microbiol.* 24, 1-6.
- Messerschmidt, S.J., Kemter, F.S., Schindler, D. and Waldminghaus, T., 2015. Synthetic secondary chromosomes in *Escherichia coli* based on the replication origin of chromosome II in *Vibrio cholerae*. *Biotechnol J.* 10, 302-14.
- Messerschmidt, S.J., Schindler, D., Zumkeller, C.M., Kemter, F.S., Schallopp, N. and Waldminghaus, T., 2016. Optimization and characterization of the synthetic secondary chromosome synVicII in *Escherichia coli*. submitted.
- Milbredt, S., Farmani, N., Sobetzko, P. and Waldminghaus, T., 2016. DNA replication in engineered *Escherichia coli* genomes with extra replication origins. *ACS Synth Biol.* 5, 1167-1176.
- Mills, R.E., Walter, K., Stewart, C., Handsaker, R.E., Chen, K., Alkan, C., Abyzov, A., Yoon, S.C., Ye, K., Cheetham, R.K., Chinwalla, A., Conrad, D.F., Fu, Y., Grubert, F., Hajirasouliha, I., Hormozdiari, F., Iakoucheva, L.M., Iqbal, Z., Kang, S., Kidd, J.M., Konkel, M.K., Korn, J., Khurana, E., Kural, D., Lam, H.Y., Leng, J., Li, R., Li, Y., Lin, C.Y., Luo, R., Mu, X.J., Nemesh, J., Peckham, H.E., Rausch, T., Scally, A., Shi, X., Stromberg, M.P., Stutz, A.M., Urban, A.E., Walker, J.A., Wu, J., Zhang, Y., Zhang, Z.D., Batzer, M.A., Ding, L., Marth, G.T., McVean, G., Sebat, J., Snyder, M., Wang, J., Eichler, E.E., Gerstein, M.B., Hurler, M.E., Lee, C., McCarroll, S.A. and Korbel, J.O., 2011. Mapping copy number variation by population-scale genome sequencing. *Nature.* 470, 59-65.
- Mirkin, E.V., Chang, F.S. and Kleckner, N., 2014. Protein-mediated chromosome pairing of repetitive arrays. *J Mol Biol.* 426, 550-7.
- Mohapatra, S.S., Ramachandran, D., Mantri, C.K. and Singh, D.V., 2007. Characterization of the genetic background of *Vibrio cholerae* O1 biotype El Tor serotype Inaba strains isolated in Trivandrum, southern India. *J Med Microbiol.* 56, 260-5.
- Mohl, D.A. and Gober, J.W., 1997. Cell cycle-dependent polar localization of chromosome partitioning proteins in *Caulobacter crescentus*. *Cell.* 88, 675-84.
- Moukadiri, I., Prado, S., Pira, J., Velazquez-Campoy, A., Bjork, G.R. and Armengod, M.E., 2009. Evolutionarily conserved proteins MnmE and GidA catalyze the formation of two methyluridine derivatives at tRNA wobble positions. *Nucleic Acids Res.* 37, 7177-93.
- Musialowski, M.S., Flett, F., Scott, G.B., Hobbs, G., Smith, C.P. and Oliver, S.G., 1994. Functional evidence that the principal DNA replication origin of the *Streptomyces coelicolor* chromosome is close to the *dnaA-gyrB* region. *J Bacteriol.* 176, 5123-5.
- Nari, A., 1995. Strategies for the control of one-host ticks and relationship with tick-borne diseases in South America. *Vet Parasitol.* 57, 153-65.
- Nascimento, A.L., Verjovski-Almeida, S., Van Sluys, M.A., Monteiro-Vitorello, C.B., Camargo, L.E., Digiampietri, L.A., Harstkeerl, R.A., Ho, P.L., Marques, M.V., Oliveira, M.C., Setubal, J.C., Haake, D.A. and Martins, E.A., 2004. Genome features of *Leptospira interrogans* serovar Copenhageni. *Braz J Med Biol Res.* 37, 459-77.
- Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., Muthuswamy, L., Krasnitz, A., McCombie, W.R., Hicks, J. and Wigler, M., 2011. Tumour evolution inferred by single-cell sequencing. *Nature.* 472, 90-4.
- Neylon, C., Kralicek, A.V., Hill, T.M. and Dixon, N.E., 2005. Replication termination in *Escherichia coli*: structure and antihelicase activity of the Tus-Ter complex. *Microbiol Mol Biol Rev.* 69, 501-26.
- Nielsen, H.J., Li, Y., Youngren, B., Hansen, F.G. and Austin, S., 2006. Progressive segregation of the *Escherichia coli* chromosome. *Mol Microbiol.* 61, 383-93.
- Nierman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J.A., Heidelberg, J.F., Alley, M.R., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Venter, J.C., Shapiro, L. and Fraser, C.M., 2001. Complete genome sequence of *Caulobacter crescentus*. *Proc Natl Acad Sci U S A.* 98, 4136-41.
- Nordstrom, K. and Dasgupta, S., 2006. Copy-number control of the *Escherichia coli* chromosome: a plasmidologist's view. *EMBO Rep.* 7, 484-9.

- O'Donnell, M., 2006. Replisome architecture and dynamics in *Escherichia coli*. *J Biol Chem.* 281, 10653-6.
- Ochman, H., 2002. Bacterial evolution: chromosome arithmetic and geometry. *Curr Biol.* 12, R427-8.
- Okada, K., Iida, T., Kita-Tsukamoto, K. and Honda, T., 2005. Vibrios commonly possess two chromosomes. *J Bacteriol.* 187, 752-7.
- Ozaki, S. and Katayama, T., 2009. DnaA structure, function, and dynamics in the initiation at the chromosomal origin. *Plasmid.* 62, 71-82.
- Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M.D., Tai, A., Main, A., Eng, D., Polichuk, D.R., Teoh, K.H., Reed, D.W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K.W., Fickes, S., Galazzo, J., Gaucher, S.P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L.F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievens, J., Covello, P.S., Keasling, J.D., Reiling, K.K., Renninger, N.S. and Newman, J.D., 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature.* 496, 528-32.
- Paintdakhi, A., Parry, B., Campos, M., Irnov, I., Elf, J., Surovtsev, I. and Jacobs-Wagner, C., 2016. Oufiti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Mol Microbiol.* 99, 767-77.
- Pan, X., Urban, A.E., Palejev, D., Schulz, V., Grubert, F., Hu, Y., Snyder, M. and Weissman, S.M., 2008. A procedure for highly specific, sensitive, and unbiased whole-genome amplification. *Proc Natl Acad Sci U S A.* 105, 15499-504.
- Picardeau, M., Lobry, J.R. and Hinnebusch, B.J., 1999. Physical mapping of an origin of bidirectional replication at the centre of the *Borrelia burgdorferi* linear chromosome. *Mol Microbiol.* 32, 437-45.
- Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W.L., Chen, C., Zhai, Y., Dairkee, S.H., Ljung, B.M., Gray, J.W. and Albertson, D.G., 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet.* 20, 207-11.
- Pinkney, M. and Thomas, C.M., 1987. Replication and maintenance of promiscuous plasmids of gram-negative bacteria. *Microbiol Sci.* 4, 186-91.
- Polaczek, P., 1990. Bending of the origin of replication of *E. coli* by binding of IHF at a specific site. *New Biol.* 2, 265-71.
- Pollack, J.R., Perou, C.M., Alizadeh, A.A., Eisen, M.B., Pergamenschikov, A., Williams, C.F., Jeffrey, S.S., Botstein, D. and Brown, P.O., 1999. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet.* 23, 41-6.
- Prescott, D.M. and Kuempel, P.L., 1972. Bidirectional replication of the chromosome in *Escherichia coli*. *Proc Natl Acad Sci U S A.* 69, 2842-5.
- Quail, M.A., Swerdlow, H. and Turner, D.J., 2009. Improved protocols for the illumina genome analyzer sequencing system. *Curr Protoc Hum Genet.* Chapter 18, Unit 18 2.
- Raghunathan, A., Ferguson, H.R., Jr., Bornarth, C.J., Song, W., Driscoll, M. and Lasken, R.S., 2005. Genomic DNA amplification from a single bacterium. *Appl Environ Microbiol.* 71, 3342-7.
- Rasmussen, T., Jensen, R.B. and Skovgaard, O., 2007. The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. *Embo J.* 26, 3124-31.
- Ravin, N.V., Kuprianov, V.V., Gilcrease, E.B. and Casjens, S.R., 2003. Bidirectional replication from an internal *ori* site of the linear N15 plasmid prophage. *Nucleic Acids Res.* 31, 6552-60.
- Ravin, N.V., Strakhova, T.S. and Kuprianov, V.V., 2001. The protelomerase of the phage-plasmid N15 is responsible for its maintenance in linear form. *J Mol Biol.* 312, 899-6.
- Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., Cho, E.K., Dallaire, S., Freeman, J.L., Gonzalez, J.R., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J.R., Marshall, C.R., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M.J., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Armengol, L., Conrad, D.F., Estivill, X., Tyler-Smith, C., Carter, N.P.,

- Aburatani, H., Lee, C., Jones, K.W., Scherer, S.W. and Hurles, M.E., 2006. Global variation in copy number in the human genome. *Nature*. 444, 444-54.
- Richards, G.P., Needleman, D.S., Watson, M.A. and Bono, J.L., 2014. Complete genome sequence of the larval shellfish pathogen *Vibrio tubiashii* Type Strain ATCC 19109. *Genome Announc.* 2.
- Robinett, C.C., Straight, A., Li, G., Willhelm, C., Sudlow, G., Murray, A. and Belmont, A.S., 1996. *In vivo* localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol.* 135, 1685-700.
- Robinow, C. and Kellenberger, E., 1994. The bacterial nucleoid revisited. *Microbiol Rev.* 58, 211-32.
- Robinson, N.P. and Bell, S.D., 2005. Origins of DNA replication in the three domains of life. *FEBS J.* 272, 3757-66.
- Roth, A., Urmoneit, B. and Messer, W., 1994. Functions of histone-like proteins in the initiation of DNA replication at *oriC* of *Escherichia coli*. *Biochimie.* 76, 917-23.
- Rozgaja, T.A., Grimwade, J.E., Iqbal, M., Czerwonka, C., Vora, M. and Leonard, A.C., 2011. Two oppositely oriented arrays of low-affinity recognition sites in *oriC* guide progressive binding of DnaA during *Escherichia coli* pre-RC assembly. *Mol Microbiol.* 82, 475-88.
- Salic, A. and Mitchison, T.J., 2008. A chemical method for fast and sensitive detection of DNA synthesis *in vivo*. *Proc Natl Acad Sci U S A.* 105, 2415-20.
- Sarrion-Perdigones, A., Falconi, E.E., Zandalinas, S.I., Juarez, P., Fernandez-del-Carmen, A., Granell, A. and Orzaez, D., 2011. GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. *PLoS One.* 6, e21622.
- Schaper, S. and Messer, W., 1995. Interaction of the initiator protein DnaA of *Escherichia coli* with its DNA target. *J Biol Chem.* 270, 17622-6.
- Schindler, D., Milbredt, S., Sperlea, T. and Waldminghaus, T., 2016. Design and assembly of DNA sequence libraries for chromosomal insertion in bacteria based on a set of modified MoClo vectors. *ACS Synth Biol.* DOI: 10.1021/acssynbio.6b00089.
- Schindler, D. and Waldminghaus, T., 2015. Synthetic chromosomes. *FEMS Microbiol Rev.* 39, 871-91.
- Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M., Chi, M., Navin, N., Lucito, R., Healy, J., Hicks, J., Ye, K., Reiner, A., Gilliam, T.C., Trask, B., Patterson, N., Zetterberg, A. and Wigler, M., 2004. Large-scale copy number polymorphism in the human genome. *Science.* 305, 525-8.
- Sekimizu, K., Bramhill, D. and Kornberg, A., 1987. ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell.* 50, 259-65.
- Shetty, R.P., Endy, D. and Knight, T.F., Jr., 2008. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng.* 2, 5.
- Shlien, A. and Malkin, D., 2009. Copy number variations and cancer. *Genome Med.* 1, 62.
- Shyu, Y.J. and Hu, C.D., 2008. Fluorescence complementation: an emerging tool for biological research. *Trends Biotechnol.* 26, 622-30.
- Skarstad, K. and Boye, E., 1994. The initiator protein DnaA: evolution, properties and function. *Biochim Biophys Acta.* 1217, 111-30.
- Skarstad, K. and Katayama, T., 2013. Regulating DNA replication in bacteria. *Cold Spring Harb Perspect Biol.* 5, a012922.
- Skarstad, K. and Lobner-Olesen, A., 2003. Stable co-existence of separate replicons in *Escherichia coli* is dependent on once-per-cell-cycle initiation. *EMBO J.* 22, 140-50.
- Smith, C.J. and Osborn, A.M., 2009. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol Ecol.* 67, 6-20.
- Smith, G.R., 2012. How RecBCD enzyme and Chi promote DNA break repair and recombination: a molecular biologist's view. *Microbiol Mol Biol Rev.* 76, 217-28.
- Smolke, C.D., 2009. Building outside of the box: iGEM and the BioBricks Foundation. *Nat Biotechnol.* 27, 1099-102.
- Sonntag Brown, M., Zanders, S. and Alani, E., 2011. Sustained and rapid chromosome movements are critical for chromosome pairing and meiotic progression in budding yeast. *Genetics.* 188, 21-32.

- Stanier, R.Y. and Van Niel, C.B., 1962. The concept of a bacterium. *Arch Mikrobiol.* 42, 17-35.
- Stanton, B.C., Nielsen, A.A., Tamsir, A., Clancy, K., Peterson, T. and Voigt, C.A., 2014. Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat Chem Biol.* 10, 99-105.
- Steen, E.J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., Del Cardayre, S.B. and Keasling, J.D., 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature.* 463, 559-62.
- Stokke, C., Waldminghaus, T. and Skarstad, K., 2011. Replication patterns and organization of replication forks in *Vibrio cholerae*. *Microbiology.* 157, 695-708.
- Straight, A.F., Belmont, A.S., Robinett, C.C. and Murray, A.W., 1996. GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. *Curr Biol.* 6, 1599-608.
- Stuitje, A.R. and Meijer, M., 1983. Maintenance and incompatibility of plasmids carrying the replication origin of the *Escherichia coli* chromosome: evidence for a control region of replication between *oriC* and *asnA*. *Nucleic Acids Res.* 11, 5775-91.
- Sugimoto, K., Oka, A., Sugisaki, H., Takanami, M., Nishimura, A., Yasuda, Y. and Hirota, Y., 1979. Nucleotide sequence of *Escherichia coli* K-12 replication origin. *Proc Natl Acad Sci U S A.* 76, 575-9.
- Suwanto, A. and Kaplan, S., 1989. Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: presence of two unique circular chromosomes. *J Bacteriol.* 171, 5850-9.
- Sykes, P.J., Neoh, S.H., Brisco, M.J., Hughes, E., Condon, J. and Morley, A.A., 1992. Quantitation of targets for PCR by use of limiting dilution. *Biotechniques.* 13, 444-9.
- Taanman, J.W., 1999. The mitochondrial genome: structure, transcription, translation and replication. *Biochim Biophys Acta.* 1410, 103-23.
- Tabata, S., Oka, A., Sugimoto, K., Takanami, M., Yasuda, S. and Hirota, Y., 1983. The 245 base-pair *oriC* sequence of the *E. coli* chromosome directs bidirectional replication at an adjacent region. *Nucleic Acids Res.* 11, 2617-26.
- Tanaka, M. and Hiraga, S., 1985. Negative control of *oriC* plasmid replication by transcription of the *oriC* region. *Mol Gen Genet.* 200, 21-6.
- Tang, F., Barbacioru, C., Nordman, E., Li, B., Xu, N., Bashkirov, V.I., Lao, K. and Surani, M.A., 2010. RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat Protoc.* 5, 516-35.
- Tardif, G. and Grant, R.B., 1983. Transfer of plasmids from *Escherichia coli* to *Pseudomonas aeruginosa*: characterization of a *Pseudomonas aeruginosa* mutant with enhanced recipient ability for enterobacterial plasmids. *Antimicrob Agents Chemother.* 24, 201-8.
- Telenius, H., Carter, N.P., Bebb, C.E., Nordenskjold, M., Ponder, B.A. and Tunnacliffe, A., 1992. Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics.* 13, 718-25.
- Thanbichler, M. and Shapiro, L., 2006. MipZ, a spatial regulator coordinating chromosome segregation with cell division in *Caulobacter*. *Cell.* 126, 147-62.
- Thompson, S.L. and Compton, D.A., 2010. Proliferation of aneuploid human cells is limited by a p53-dependent mechanism. *J Cell Biol.* 188, 369-81.
- Tonthat, N.K., Arold, S.T., Pickering, B.F., Van Dyke, M.W., Liang, S., Lu, Y., Beuria, T.K., Margolin, W. and Schumacher, M.A., 2011. Molecular mechanism by which the nucleoid occlusion factor, SlmA, keeps cytokinesis in check. *EMBO J.* 30, 154-64.
- Trucksis, M., Michalski, J., Deng, Y.K. and Kaper, J.B., 1998. The *Vibrio cholerae* genome contains two unique circular chromosomes. *Proc Natl Acad Sci U S A.* 95, 14464-9.
- Val, M.E., Marbouty, M., de Lemos Martins, F., Kennedy, S.P., Kemble, H., Bland, M.J., Possoz, C., Koszul, R., Skovgaard, O. and Mazel, D., 2016. A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci Adv.* 2, e1501914.
- Venkova-Canova, T. and Chattoraj, D.K., 2011. Transition from a plasmid to a chromosomal mode of replication entails additional regulators. *Proc Natl Acad Sci U S A.* 108, 6199-204.
- Vogelstein, B. and Kinzler, K.W., 1999. Digital PCR. *Proc Natl Acad Sci U S A.* 96, 9236-41.

- von Meyenburg, K., Hansen, F.G., Riise, E., Bergmans, H.E., Meijer, M. and Messer, W., 1979. Origin of replication, *oriC*, of the *Escherichia coli* K12 chromosome: genetic mapping and minichromosome replication. Cold Spring Harb Symp Quant Biol. 43 Pt 1, 121-8.
- Waldminghaus, T., Weigel, C. and Skarstad, K., 2012. Replication fork movement and methylation govern SeqA binding to the *Escherichia coli* chromosome. Nucleic Acids Res. 40, 5465-76.
- Wang, S.C. and Shapiro, L., 2004. The topoisomerase IV ParC subunit colocalizes with the *Caulobacter* replisome and is required for polar localization of replication origins. Proc Natl Acad Sci U S A. 101, 9251-6.
- Wang, X., Lesterlin, C., Reyes-Lamothe, R., Ball, G. and Sherratt, D.J., 2011. Replication and segregation of an *Escherichia coli* chromosome with two replication origins. Proc Natl Acad Sci U S A. 108, E243-50.
- Wang, X., Montero Llopis, P. and Rudner, D.Z., 2013. Organization and segregation of bacterial chromosomes. Nat Rev Genet. 14, 191-203.
- Wang, X., Montero Llopis, P. and Rudner, D.Z., 2014. *Bacillus subtilis* chromosome organization oscillates between two distinct patterns. Proc Natl Acad Sci U S A. 111, 12877-82.
- Wang, X. and Sherratt, D.J., 2010. Independent segregation of the two arms of the *Escherichia coli* *ori* region requires neither RNA synthesis nor MreB dynamics. J Bacteriol. 192, 6143-53.
- Weber, E., Engler, C., Gruetzner, R., Werner, S. and Marillonnet, S., 2011. A modular cloning system for standardized assembly of multigene constructs. PLoS One. 6, e16765.
- Wilson, C.J., Zhan, H., Swint-Kruse, L. and Matthews, K.S., 2007. The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. Cell Mol Life Sci. 64, 3-16.
- Woelker, B. and Messer, W., 1993. The structure of the initiation complex at the replication origin, *oriC*, of *Escherichia coli*. Nucleic Acids Res. 21, 5025-33.
- Yamaguchi, K., Yamaguchi, M. and Tomizawa, J., 1982. Incompatibility of plasmids containing the replication origin of the *Escherichia coli* chromosome. Proc Natl Acad Sci U S A. 79, 5347-51.
- Yamaichi, Y., Fogel, M.A., McLeod, S.M., Hui, M.P. and Waldor, M.K., 2007a. Distinct centromere-like *parS* sites on the two chromosomes of *Vibrio* spp. J Bacteriol. 189, 5314-24.
- Yamaichi, Y., Fogel, M.A. and Waldor, M.K., 2007b. *par* genes and the pathology of chromosome loss in *Vibrio cholerae*. Proc Natl Acad Sci U S A. 104, 630-5.
- Yasuda, S. and Hirota, Y., 1977. Cloning and mapping of the replication origin of *Escherichia coli*. Proc Natl Acad Sci U S A. 74, 5458-62.
- Yilmaz, S. and Singh, A.K., 2012. Single cell genome sequencing. Curr Opin Biotechnol. 23, 437-43.
- Yin, H. and Marshall, D., 2012. Microfluidics for single cell analysis. Curr Opin Biotechnol. 23, 110-9.
- Zare, R.N. and Kim, S., 2010. Microfluidic platforms for single-cell analysis. Annu Rev Biomed Eng. 12, 187-201.
- Zhang, K., Martiny, A.C., Reppas, N.B., Barry, K.W., Malek, J., Chisholm, S.W. and Church, G.M., 2006. Sequencing genomes from single cells by polymerase cloning. Nat Biotechnol. 24, 680-6.
- Zhao, L., Kroenke, C.D., Song, J., Piwnica-Worms, D., Ackerman, J.J. and Neil, J.J., 2008. Intracellular water-specific MR of microbead-adherent cells: the HeLa cell intracellular water exchange lifetime. NMR Biomed. 21, 159-64.
- Zimmerman, S.B. and Murphy, L.D., 2001. Release of compact nucleoids with characteristic shapes from *Escherichia coli*. J Bacteriol. 183, 5041-9.

Erklärung der selbständigen Erarbeitung der Dissertation

Ich versichere, dass ich meine Dissertation

„Charakterisierung der DNA-Replikation in genetisch modifizierten *Escherichia coli* Stämmen durch neue Methoden der Synthetischen Biologie“

selbstständig und ohne unerlaubte Hilfsmittel angefertigt habe. Ich habe mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient. Die Dissertation wurde in der jetzigen oder in einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinem sonstigen Prüfungszweck gedient.

Marburg, den 20.12.2016

S. Milbredt

Sarah Milbredt