

## DNA-Replikation

# Dem Anfang auf der Spur – die Suche nach DNA-Replikationsursprüngen

ELISABETH KRUSE, STEPHAN HAMPERL  
INSTITUT FÜR EPIGENETIK UND STAMMZELLEN, HELMHOLTZ ZENTRUM MÜNCHEN

**Timely and accurate duplication of DNA prior to cell division is a prerequisite for propagation of the genetic material to both daughter cells. DNA synthesis initiates at discrete sites, termed replication origins, and proceeds in a bidirectional manner until all genomic DNA is replicated. Despite the fundamental nature of these events, a uniform method that identifies origins of replication in a comprehensive manner is still missing. Here, we present currently available and discuss new approaches to map replication origins in mammalian genomes.**

DOI: 10.1007/s12268-021-1562-z  
© Die Autoren 2021

■ DNA-Replikation ist ein essenzieller Prozess, der sicherstellt, dass jede Zelle nach der Zellteilung eine exakte und vollständige Kopie des Erbmateri als an die beiden Tochterzellen weitergibt. Zur Verdopplung der DNA muss die Doppelhelix lokal aufgelöst werden. An den beiden Einzelsträngen können dann Replikationsmaschinen – auch

Replisomen genannt – jeweils einen komplementären Tochterstrang synthetisieren. Es entstehen zwei Doppelstränge, die jeweils aus einem alten und einem neuen Strang zusammengesetzt sind. Dieser fundamentale Prozess der semikonservativen Replikation ist in allen Organismen konserviert und streng reguliert, um u. a. die DNA-Synthese

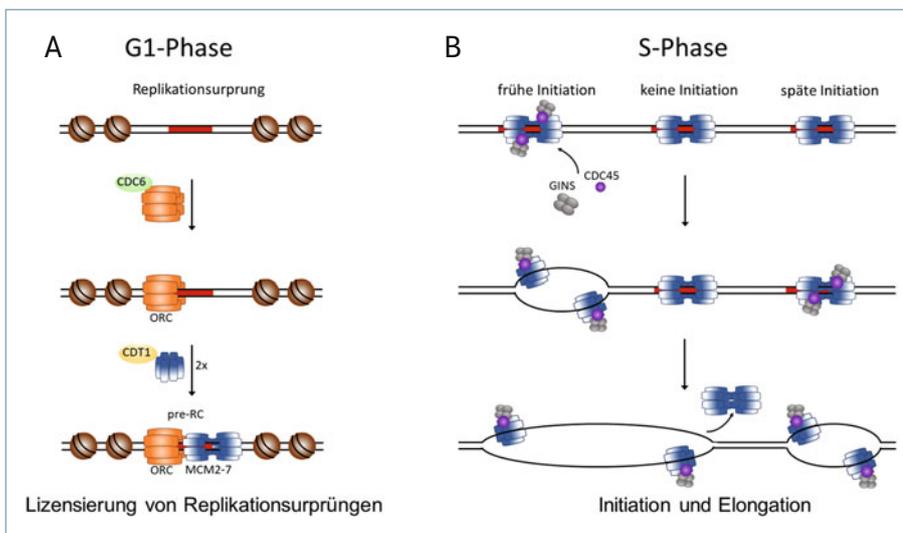
mit dem Zellzyklus und Energiestatus der Zelle zu koordinieren. Eine wichtige Voraussetzung für die DNA-Replikation ist, dass sie eine sehr geringe Fehlerquote bei hoher Effizienz aufweist, um die Anhäufung genetischer Veränderungen und damit potenziell schädliche Folgen für den Organismus zu verhindern.

Solange das gesamte Genom vor der Zellteilung dupliziert wird, könnte man annehmen, dass der genaue Ort der Replikationsstartstellen keine Rolle spielt. Es wurde jedoch gezeigt, dass viele Organismen bevorzugte Genomregionen als Ursprünge verwenden [1, 2, 3]. Dass Zellen diese Ursprungsorte regulieren, ergibt sich aus der Notwendigkeit, die DNA-Replikation mit anderen Prozessen auf der gemeinsamen Chromatinmatrize zu koordinieren, wie z. B. Transkription oder DNA-Reparatur. Dabei verwenden eukaryotische Zellen abhängig von der Größe ihres Genoms mehrere hundert bis zehntausende von Replikationsursprüngen, um die Vervielfältigung ihrer großen linearen Chromosomen in dem kurzen Zeitfenster der S-Phase zu ermöglichen. Die Initiation der DNA-Replikation über diese große Anzahl von Ursprüngen muss daher räumlich und zeitlich präzise koordiniert werden.

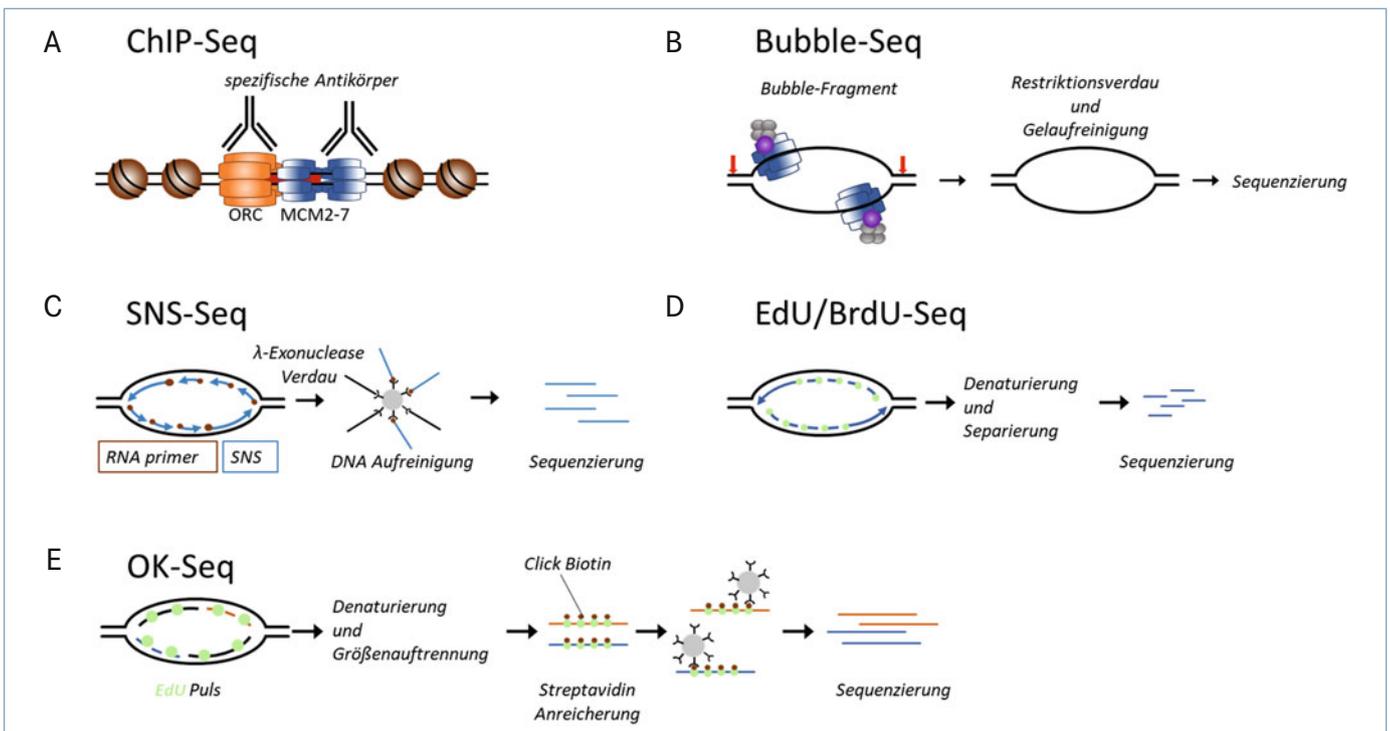
### Replikationsinitiation

Dieser Prozess ist zeitlich auf zwei Zellzyklusstadien verteilt: Während der G1-Phase werden Ursprünge zuerst vom hexameren *origin recognition complex* (ORC) erkannt. Anschließend führt der *helicase-loader* CDC6 zusammen mit dem CDT1-Chaperon zur Rekrutierung von zwei inaktiven *mini-chromosome maintenance* (MCM2-7)-Hexameren an die DNA. Damit ist das *origin licensing* abgeschlossen und alle möglichen Ursprünge als *pre-replication complex* (Pre-RC) für das Feuern in der folgenden S-Phase vorbereitet (**Abb. 1A**).

Beim Eintritt in die S-Phase werden die *cyclin-dependent kinase* (CDK) und die *DBF4-dependent kinase* (DDK) aktiviert. Diese Enzyme vermitteln durch mehrere



▲ **Abb. 1:** Lizenzierung und Aktivierung von DNA-Replikationsursprüngen. **A,** Die Lizenzierung von Replikationsursprüngen findet in der G1-Phase des Zellzyklus statt. Nach der Bindung von ORC durch CDC6 werden über das CDT1-Chaperon zwei inaktive MCM2-7-Hexamere an den Ursprung rekrutiert. **B,** Nach dem Eintritt in die S-Phase binden die Initiationsfaktoren CDC45 und GINS zuerst früh feuernde Replikationsursprünge und erzeugen eine bidirektionale Replikationsblase. Die Aktivierung von Ursprüngen folgt einem zeitlichen Programm, bei dem andere Ursprünge erst spät feuern oder inaktiv sind und nur passiv repliziert werden.



▲ **Abb. 2:** Schematische Darstellung aktueller Methoden zur Charakterisierung von Replikationsursprüngen. **A**, ChIP-seq: Chromatin-Immunopräzipitation mit anschließendem Next Generation Sequencing von ORC und MCM2-7. **B**, Bubble-Seq: Replikationsblasen werden durch verlangsamtes Laufverhalten im Agarosegel aufgetrennt und nach Restriktionsverdau zum Erstellen von Sequenzierbibliotheken genutzt. **C**, SNS-Seq: Die „neu replizierte DNA“ (*short-nascent-strand*, SNS) wird als Initiationszwischenprodukt mit einem 5'-RNA-Primer synthetisiert. Der Verdau mit Lambda-Exonuklease baut kontaminierende DNA-Fragmente ohne einen 5'-RNA-Primer ab. **D**, EdU-/BrdU-Seq: Zellen werden mit den Thymidinanaloga EdU oder BrdU pulsmarkiert und die markierten DNA-Stränge durch anti-BrdU-Antikörper oder EdU-ClickIT-Reaktion immunpräzipitiert und sequenziert. **E**, OK-Seq: kurze Okazaki-Fragmente entstehen spezifisch nur am DNA-Folgestrang. Nach einem EdU-Puls können die kurzen Fragmente aufgetrennt und sequenziert werden.

Phosphorylierungsereignisse von Pre-RC-Komponenten die Rekrutierung von CDC45 und des GINS-Tetramers zum MCM2-7-Hexamere, um den inaktiven Komplex in die aktive CMG-Helikase (CDC45/MCM2-7/GINS) umzuwandeln. Nach dem Schmelzen des DNA-Duplex durch die MCM2-7-Helikase assemblieren DNA-Polymerasen und zusätzliche Replikationsfaktoren, um den Ursprung zu aktivieren und eine bidirektionale Replikationsblase zu erzeugen (**Abb. 1B**).

Interessanterweise wird nur ein kleiner Anteil aller potenziellen Replikationsursprünge (20–30 % in somatischen Säugetierzellen) tatsächlich für die Initiation in einer gegebenen S-Phase ausgewählt. Die überschüssigen Ursprünge werden flexibel und scheinbar zufällig nur in manchen Zellzyklen genutzt oder dienen als Backup – werden also nur aktiviert, wenn nahegelegene Replikationsgabeln verlangsamt oder blockiert werden. Solch ein Replikationsstress kann durch viele Ereignisse in der Zelle verursacht werden, u. a. sekundäre DNA-Strukturen, Kompaktierung des Chromatins, Mangel an essenziellen Replikationsfaktoren oder Kon-

flikte mit anderen Prozessen, wie der Transkription [4]. Eine der großen Herausforderungen ist es herauszufinden, anhand welcher Kriterien ein Ursprung gegenüber den vielen anderen ausgewählt wird, tatsächlich zu feuern. Um dieser Plastizität besser auf die Spur zu kommen, wäre eine umfassende und genaue Kartierung aller benutzten Replikationsursprünge im humanen Genom ein entscheidender Schritt.

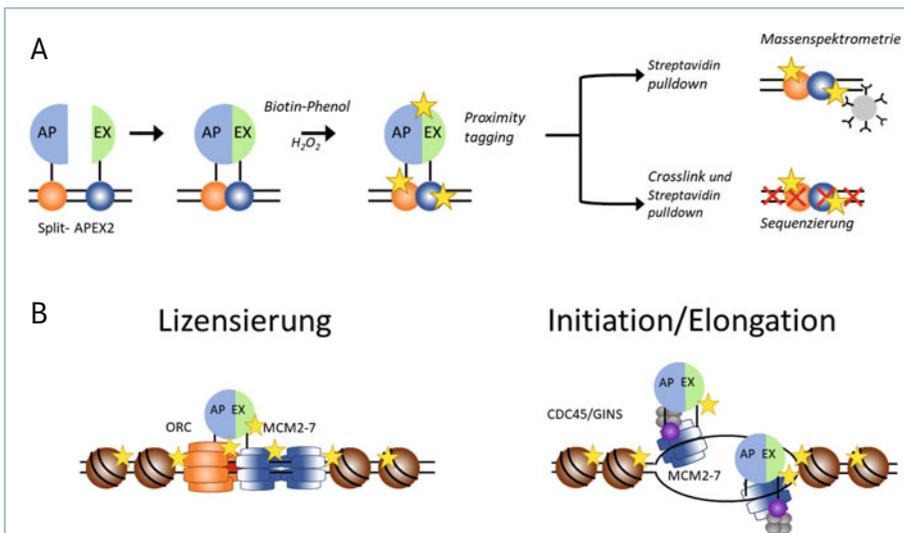
### Methoden zur Charakterisierung von Replikationsursprüngen

Über die letzten Jahre wurden mehrere orthogonale Ansätze entwickelt, um Replikationsursprünge in Säugetiergenomen zu identifizieren (**Abb. 2**). Die Chromatin-Immunopräzipitation, gefolgt von Next Generation Sequencing (ChIP-Seq) von ORC- und MCM2-7-Untereinheiten nutzt die Bindung dieser essenziellen Pre-RC-Komponenten an die Ursprünge [5, 6]. ORC hat jedoch auch andere replikationsunabhängige Funktionen, beispielsweise die Assoziation in Heterochromatin [7, 8]. Zudem identifizieren solche Pre-RC-ChIP-Studien alle lizenzierten Ursprünge im Genom, und nicht spezifisch

die später zur Initiation ausgewählten Ursprünge.

Andere Verfahren zielen daher auf die direkte Identifizierung von Zwischenprodukten der DNA-Replikation ab. Bubble-seq stellt Sequenzierbibliotheken von Restriktionsenzymfragmenten her, die Ursprünge in Form von zirkulären Replikationsblasen enthalten. Diese Sekundärstruktur der DNA verlangsamt die Fragmente in Agarosegelen und kann somit von der linearen DNA abgetrennt und isoliert werden [9].

Eine weitere Technik basiert auf der Anreicherung und Sequenzierung von „neu replizierter DNA“ (*short-nascent-strand*, SNS). SNS werden symmetrisch vom Ursprung in beide Richtungen als Initiationszwischenprodukte mit einem 5'-RNA-Primer synthetisiert und sind daher geeignet für die Identifizierung aktiver Replikationsursprünge. Zur Analyse der SNS wird DNA aus proliferierenden Zellen isoliert, denaturiert und entsprechend der DNA-Größe fraktioniert. Da durch die genomische DNA-Isolierung auch eine Vielzahl an nicht replizierenden DNA-Bruchstücken entsteht, müssen diese kontaminierenden DNA-Fragmente mit Lambda-Exonu-



▲ **Abb. 3:** Schematische Übersicht der Split-APEX-Methode. **A,** Zwei inaktive Fragmente AP und EX rekonstituieren sich zu aktivem APEX2, um Proteine in der Nähe zu biotinylieren. Die biotinylierten Proteine können durch Streptavidin-Pulldown und MS-Analyse identifiziert werden. Alternativ wird Chromatin mit Formaldehyd (FA) vernetzt und nach Streptavidin-Pulldown die DNA-Fragmente sequenziert um die genomischen Ursprungsorte zu identifizieren. **B,** Verschiedene Proteinkombinationen mit AP und EX erlauben die Analyse verschiedener Intermediate bei der Lizenzierung oder Initiierung/Aktivierung.

klase vor dem Sequenzieren verdaut werden. SNS-Seq-Studien haben umfassende Karten von Replikationsursprüngen in mehreren menschlichen und Mauszelllinien generiert, mit einer guten Auflösung im Kilobasen-Bereich [10, 11]. Ein Nachteil von SNS-Seq ist die geringe Reinigungsausbeute der SNS (< 10 %) und somit müssen Ursprünge effizient in einem großen Anteil der Zellpopulation feuern, um über dem Hintergrund detektiert zu werden. Die ineffizienteren flexiblen Ursprünge entziehen sich meist dieser Analyse.

Alternativen zum SNS-Seq sind Bromdesoxyuridin (BrdU)- und 5-ethynyl-2'-deoxyuridin (EdU)-Seq. BrdU und EdU sind Thymidinanaloga, welche in neu synthetisierte DNA inkorporiert werden. Durch anti-BrdU-Immunopräzipitation oder Click-IT-Reaktion von EdU mit Biotin-Azid können diese naszierenden DNA-Fragmente dann spezifisch isoliert und sequenziert werden [12]. Eine Variante dieser Methode nutzt einen EdU-Puls, um dann spezifisch Okazaki-Fragmente zu isolieren und zu sequenzieren (OK-Seq). Okazaki-Fragmente sind kurze Abschnitte (100 bis 200 bp), die spezifisch am DNA-Folgestrang entstehen [13]. Dadurch ist die Bestimmung der Direktionalität, in welcher Richtung eine Replikationsgabel einen bestimmten Bereich im Genom repliziert hat möglich, und erlaubt eine quantitative Analyse der Replikationsrichtung.

All diese Next-Generation-Sequencing-Methoden haben in sich selbst zur Entdeckung von zehntausenden potenzieller Replikationsursprünge in Säugetiergenomen geführt. Überraschenderweise sind die Ergebnisse aber nicht konsistent und nur ein Bruchteil an Ursprüngen stimmt zwischen verschiedenen Methoden überein. Dies kann verschiedenste Ursachen haben, z. B. fehlende Auflösung und Sequenzierungstiefe oder aber auch die Tatsache, dass diese Methoden unterschiedliche Replikationsereignisse detektieren. Daher ist die genaue Menge und Lage der Ursprünge in Säugetiergenomen noch immer eine offene Frage [14]. Zusätzlich können all diese Methoden nur die Orte der Replikationsstartstellen identifizieren, aber nicht die Zusammensetzung der assoziierten Chromatinstruktur, die viele weitere Aufschlüsse über die Regulation dieser funktionellen DNA-Elemente geben könnte.

### Proximity labelling von Replikationsursprüngen

Wir versuchen im Labor eine alternative Methode zur Analyse der Lokalisierung und Verteilung von Replikationsursprüngen im humanen Genom auf der Basis von *proximity labelling* mit APEX2 zu etablieren. APEX2 ist eine künstlich hergestellte Ascorbat-Peroxidase, die nach Zugabe von  $H_2O_2$  Biotin-Phenol als Substrat in ein reaktives Phenoxyra-

dikal umsetzt und dadurch naheliegende Proteine mit einem Biotinmolekül markiert. Im weiteren Verlauf können biotinylierte Proteine mit Streptavidin-Beads aufgereinigt und das resultierende Proxisom durch Massenspektrometrie (MS) analysiert werden [15]. In der Split-APEX2-Version wird das APEX2-Protein in zwei inaktive Unter-einheiten (AP und EX) geteilt, welche an ausgewählte Zielproteine fusioniert werden. Sobald die Zielproteine in der Zelle in räumliche Nähe kommen, können die beiden Fragmente zu einem katalytisch aktiven APEX2-Protein rekonstituiert und Proteine in der Umgebung biotinyliert werden. Für unsere Fragestellung werden die inaktiven Split-APEX2-Fragmente an Proteine der Replikationsmaschinerie fusioniert, die spezifisch an Replikationsursprüngen während der G1- oder S-Phase in räumliche Nähe kommen.

Diese Methode hat mehrere Vorteile gegenüber anderen Mapping-Strategien von Replikationsursprüngen. Zum einen geben uns die Biotinylierungsereignisse nach ChIP-Seq nicht nur Aufschluss über die Position und Verteilung der Replikationsursprünge im humanen Genom, sondern wir erhalten durch MS auf molekularer Ebene auch Aufschluss über die Chromatinumgebung und assoziierten Faktoren (**Abb. 3A**). Zum anderen lassen sich die APEX2-Fragmente flexibel an unterschiedliche Zielproteine fusionieren, was uns erlaubt verschiedenste Fragestellungen im Replikationsablauf zu beantworten. Zum Beispiel eignen sich Proteinkombinationen von ORC-CDC6 oder ORC-MCM2-7 für die Analyse der Lizenzierung von Replikationsursprüngen, ORC-CDC45 für die Bildung des CMG-Komplexes oder MCM2-7-CDC45 oder MCM2-7-GINS für die Aktivierung/Elongation der Replikation (**Abb. 3B**).

Die Kombination von Split-APEX2 mit MS oder ChIP-Seq ist daher ein vielversprechender Ansatz, der eine höhere Auflösung und mehr Information über die beteiligten Faktoren bei der Replikationsinitiierung verspricht. Wir hoffen dadurch ein umfassenderes Bild zu erhalten, wo welche Ursprünge wie oft aktiviert werden und welche molekularen Faktoren hier ausschlaggebend sind, um den Anfang der DNA-Replikation zu bestimmen.

### Danksagung

Wir bedanken uns bei allen Mitgliedern der AG Hamperl für die konstruktiven Kommentare, die das Manuskript erheblich verbessert haben. Großer Dank gilt auch Antonia Kruse

für die hilfreiche Unterstützung bei der farblichen Gestaltung der Abbildungen. ■

## Literatur

- [1] Fragkos M, Ganier O, Coulombe P et al. (2015) DNA replication origin activation in space and time. *Nat Rev Mol Cell Biol* 16: 360–374
- [2] Cayrou C, Coulombe P, Puy A et al. (2012) New insights into replication origin characteristics in metazoans. *Cell Cycle* 11: 658–667
- [3] Lombrana R, Almeida R, Revuelta I et al. (2013) High-resolution analysis of DNA synthesis start sites and nucleosome architecture at efficient mammalian replication origins. *EMBO J* 32: 2631–2644
- [4] Hamperl S, Bocek MJ, Saldivar JC et al. (2017) Transcription-replication conflict orientation modulates R-loop levels and activates distinct DNA damage responses. *Cell* 170: 774–786
- [5] Miotto B, Ji Z, Struhl K (2016) Selectivity of ORC binding sites and the relation to replication timing, fragile sites, and deletions in cancers. *Proc Natl Acad Sci U S A* 113: E4810–E4819
- [6] Dellino GI, Cittaro D, Piccioni R et al. (2013) Genome-wide mapping of human DNA-replication origins: levels of transcription at ORC1 sites regulate origin selection and replication timing. *Genome Res* 23: 1–11
- [7] Chesnokov IN (2007) Multiple functions of the origin recognition complex. *Int Rev Cytol* 256: 69–109
- [8] Hemerly AS, Prasanth SG, Siddiqui K et al. (2009) Orc1 controls centriole and centrosome copy number in human cells. *Science* 323: 789–793
- [9] Mesner LD, Valsakumar V, Cieslik M et al. (2013) Bubble-seq analysis of the human genome reveals distinct chromatin-mediated mechanisms for regulating early- and late-firing origins. *Genome Res* 23: 1774–1788
- [10] Cayrou C, Ballester B, Peiffer I et al. (2015) The chromatin environment shapes DNA replication origin organization and defines origin classes. *Genome Res* 25: 1873–1885
- [11] Besnard E, Babled A, Lapasset L et al. (2012) Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. *Nat Struct Mol Biol* 19: 837–844

- [12] Hyrien O (2015) Peaks cloaked in the mist: the landscape of mammalian replication origins. *J Cell Biol* 208: 147–160
- [13] Petryk N, Kahli M, d'Aubenton-Carafa Y et al. (2016) Replication landscape of the human genome. *Nat Commun* 7: 10208
- [14] Prioleau MN, MacAlpine DM (2016) DNA replication origins-where do we begin? *Genes Dev* 30: 1683–1697
- [15] Ummethum H, Hamperl S (2020) Proximity labeling techniques to study chromatin. *Front Genet* 11: 450

**Funding note:** Open Access funding enabled and organized by Projekt DEAL.  
**Open Access:** Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende

nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

## Korrespondenzadresse:

Dr. Stephan Hamperl  
 Institut für Epigenetik und Stammzellen  
 Helmholtz Zentrum München  
 Ingolstädter Landstraße 1  
 D-85764 Neuherberg  
 stephan.hamperl@helmholtz-muenchen.de  
 www.helmholtz-muenchen.de/ies

## AUTOREN



### Elisabeth Kruse

2007–2010 Bachelor Biologie an der Universität zu Köln. 2011–2013 Master Molecular and Cellular Life Sciences an der Utrecht University, Niederlande. 2013–2018 Technische Assistentin in der Gruppe vom Prof. Dr. E. Wiertz am University Medical Center (UMC) Utrecht. Seit 2018 Labormanagerin in der Gruppe von Dr. S. Hamperl am Helmholtz Zentrum München.



### Stephan Hamperl

2004–2008 Diplom-Biochemiestudium an der Universität Regensburg. 2009–2012 Doktorand am Lehrstuhl Biochemie III Universität Regensburg. 2013–2018 Postdoc im Cimprich Labor am Department für Chemical & Systems Biology, Stanford University, CA, USA. Seit 2018 Juniorgruppenleiter am Helmholtz Zentrum München.