

das Nukleosom – Zugangskontrolle zum Genom?

Computersimulationen und Hochdurchsatzsequenzierungen entschlüsseln die Zusammenhänge

von Karsten Rippe und Gero Wedemann

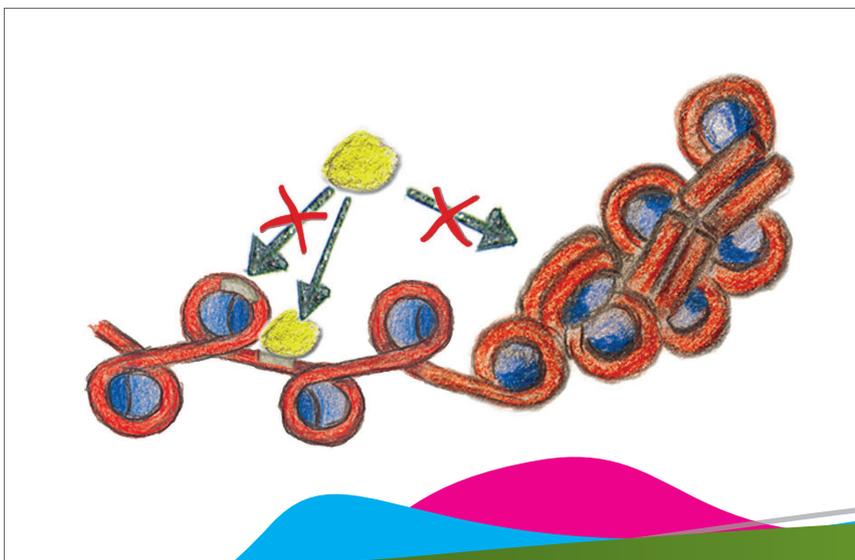
Das menschliche Genom bildet eine Kette von Nukleosomen, in denen die DNA um einen Proteinkern gewickelt ist. Zahlreiche Transkriptionsfaktoren binden besonders gut an die Verbindungs-DNA zwischen zwei Nukleosomen. Die Positionierung von Nukleosomen kann dadurch die Zugänglichkeit einer Genomsequenz direkt bestimmen und diese zusätzlich auch indirekt beeinflussen: Regelmäßige Abstände begünstigen eine Kompaktierung der Nukleosomenkette in übergeordnete Faserstrukturen, in denen auch die Verbindungs-DNA unzugänglich sein kann. Eine Vorhersage der Genexpression anhand der Bindung von Transkriptionsfaktoren muss deshalb nicht nur die DNA-Sequenz und die Proteinkonzentration berücksichtigen, sondern auch die räumliche Anordnung der Nukleosomen im Genom.

Welche Bereiche der DNA-Sequenz werden ausgelesen?

Alle Zellen des menschlichen Körpers enthalten im Wesentlichen die gleiche DNA-Sequenz und damit dieselbe genetische

Information. Die DNA-Sequenz kodiert den Bau- und Funktionsplan des Organismus. Allerdings „liest“ jede Zelle nur bestimmte Kapitel dieses „DNA-Buches“ und verwendet so lediglich einige der Pläne und Rezepte der DNA, um spezifische Funktionen umzusetzen. Wie aber wählt eine undifferenzierte Stammzelle während der Entwicklung das entsprechende DNA-Programm aus, das sie beispielsweise zu einer Muskel-, Leber- oder Hautzelle werden lässt? Einer der an diesem Prozess beteiligten Faktoren ist die Verpackung des DNA-Genoms im Zellkern zum sogenannten *Chromatin* (der Begriff stammt von dem griechischen Wort *chroma*, d. h. Farbe, denn es lässt sich einfach für Lichtmikroskopie-Aufnahmen einfärben). Das Chromatin in einem menschlichen Zellkern enthält DNA-Ketten, die aneinandergereiht eine Länge von annähernd zwei Metern ergeben. Etwa $\frac{1}{4}$ dieser DNA sind Bestandteil der rund 30 Millionen Nukleosomen. In jedem Nukleosom ist DNA in ungefähr zwei Windungen um einen Histonproteinkern gewickelt. Die Nukleosomen sind ihrerseits durch dazwischenliegende Segmente aus proteinfreier Verbindungs-DNA miteinander verbunden. Wie in Abbildung 1 dargestellt, ist die DNA innerhalb der Nukleosomen für Transkriptionsfaktoren schlechter zugänglich als die Verbindungs-DNA.

Abbildung 1: Nukleosomenkette mit einem Protein, das an die freie DNA zwischen zwei Nukleosomen bindet



Nukleosomenkette mit einem zwischen den Nukleosomen (blau) an die Verbindungs-DNA (rot) gebundenen Transkriptionsfaktor (gelb). Die offene Konformation, in der die Nukleosomen wie in einer Perlenkette aneinandergereiht sind, kann sich zu einer Chromatinfaser falten. Daher beeinflussen sowohl die Positionierung eines Nukleosoms auf der DNA sowie die übergeordnete Faltstruktur der Nukleosomenkette die Zugänglichkeit zur DNA (Quelle: Gernot Längst).

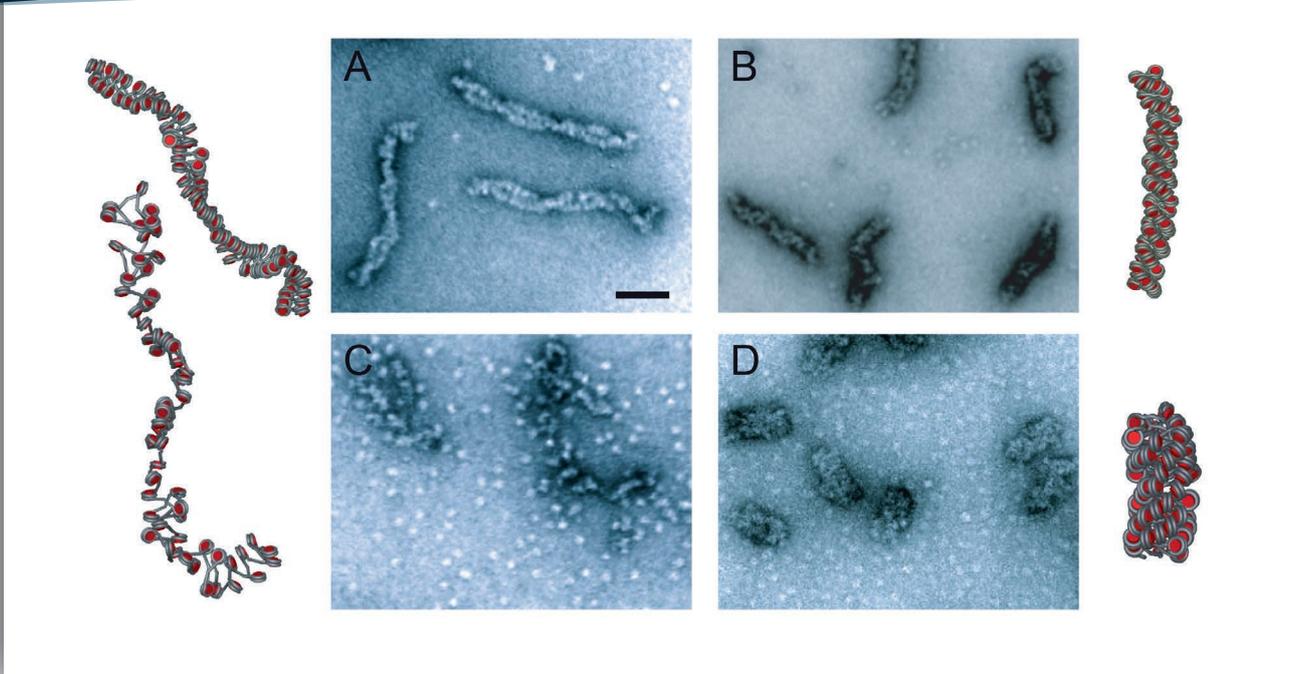


Abbildung 2: Elektronenmikroskopische Aufnahmen von Chromatinfasern und Modellstrukturen aus Monte-Carlo-Simulationen

Die Proben unterscheiden sich in Bezug auf den Abstand der Nukleosomen, d.h. die nukleosomale Wiederholungslänge (nucleosome repeat length, NRL) und die Anwesenheit bzw. das Fehlen des Verbindungs-Histons H1. Maßstab: 50 nm.

- A)** 167-bp-NRL-Faser ohne Verbindungs-Histon H1.
- B)** Zugabe von H1 löst die Kompaktierung der 167-bp-NRL-Faser aus.
- C)** Ungefaltete 197-bp-NRL-Nukleosomenkette mit einer unregelmäßigeren Struktur ohne Verbindungs-Histon.
- D)** Vollständig gefaltete 197-bp-NRL-Faserstruktur mit Verbindungs-Histon H1.

[Quelle: Daniela Rhodes (Elektronenmikroskopie), Nick Kepper (Chromatinfasersimulation)].

dungs-DNA. Folglich bestimmt die Positionierung der Nukleosomen, ob sich bestimmte DNA-Sequenzen im Bereich der leichter zugänglichen Verbindungs-DNA zwischen den Nukleosomen befinden, oder diese durch die Interaktionen von Histon und DNA verborgen werden. Insofern entscheiden die Nukleosomenpositionen zusammen mit anderen Regulationsmechanismen, ob Gene aktiv oder inaktiv sind.

Modelle, Experimente und neue Modelle...

Unsere theoretischen Studien an Nukleosomen enthüllten einen komplexen Zusammenhang zwischen der dynamischen Struktur des Nukleosoms, Interaktionen der unstrukturierten Enden der Histonproteine mit DNA und der Bindung von Transkriptionsfaktoren (Ettig *et al.*, 2011; Teif *et al.*, 2010). In diesem Zusammenhang entwickelten wir Modelle zur Vorhersage der Zugänglichkeit von Verbindungs-DNA sowie der partiellen Abwicklung der nukleosomalen DNA, um längere Ketten aus bis zu 1.000 Nukleosomen zu untersuchen (Rippe *et al.*, 2012). Diese Modelle wurden anhand von experimentellen Daten aus *in-vitro*-Analysen kurzer Nukleosomenketten mit regelmäßigem Nukleosomenabstand konstruiert und parametrisiert. Wie aus Abbildung 2 ersichtlich, gibt es eine sehr gute Übereinstimmung zwischen experimentell unter dem Elektronenmikroskop untersuchten

Faserstrukturen und den durch die Monte-Carlo-Simulation vorhergesagten Konformationen.

Bis vor kurzem war es jedoch nicht möglich, den Zusammenhang zwischen der Positionierung von Nukleosomen und der Bindung der Transkriptionsfaktoren in Säugetiergenomen genomweit zu untersuchen. Stattdessen konzentrierten sich die meisten Studien auf die Analyse einfacher Modellorganismen wie der Hefe oder der Fruchtfliege *Drosophila* oder beschränkten sich auf die Positionierung von Nukleosomen bei ausgewählten Genen. Aufgrund der jüngsten Fortschritte bei den Hochdurchsatz-DNA-Sequenzierungsmethoden wurde es möglich, alle Nukleosomenpositionen des Säugetiergenoms mit einer Auflösung von einem Basenpaar zu bestimmen und eine Verknüpfung zwischen Modellen und Experimenten herzustellen (Abbildung 3). Im Rahmen des BMBF-geförderten Verbundprojekts *EpiGenSys* wurde diese spannende neue Möglichkeit genutzt. Ziel war es, erstmalig alle Nukleosomenpositionen in embryonalen Stammzellen (also noch undifferenzierten Zellen) der Maus im Vergleich zu deren differenzierten Gegenstücken (also beispielsweise Hautzellen) zu bestimmen (Teif *et al.*, 2012). Die Ergebnisse enthüllten zahlreiche Eigenschaften der Nukleosomenpositionierung an DNA-Stellen, die von großer Bedeutung für die Differenzierung von Zellen sind. Anfang und Ende aktiver Gene

zeichneten sich durch nukleosomenarme Regionen aus, während in diesen Bereichen bei abgeschalteten Genen zahlreiche Nukleosomen zu finden sind.

Es stellte sich außerdem heraus, dass sich die Positionsprofile der Nukleosomen je nach spezifischer Modifikation des Histonproteins verändern. Zu diesen Modifikationen zählen u. a. posttranslationale (also nach der Proteinsynthese aus der genetischen Information stattfindende) Modifikationen mit Methyl- und Acetylgruppen, die an die Histone geheftet bzw. von ihnen entfernt werden können. Diese sogenannten „epigenetischen“ Signale dienen der Übertragung eines bestimmten Chromatinzustands während der Zellteilung.

Wie erwartet konnte festgestellt werden, dass viele der Proteine, die bei der Zellentwicklung eine zentrale Rolle übernehmen, in embryonalen Stammzellen tatsächlich an die freie Verbindungs-DNA zwischen Nukleosomen binden. Dieses Ergebnis bestätigte die erwartete Rolle des Nukleosoms bei der Regulierung des DNA-Zugangs. Allerdings zeigten einige Proteine, wie drei der übergeordneten Regulatoren des Stammzellenzustands, nämlich die Proteine Nanog, Oct4 und Sox2, ein anderes Bindungsmuster. Diese wurden auch in Bereichen mit hoher Nukleosomenkonzentration gefunden, was nahelegt, dass sie als Initiationsfaktoren für weitere Aktivierungsschritte an die nukleosomale DNA binden können. Mit dieser Erkenntnis muss die vereinfachte Sichtweise, dass Nukleosomen stets die Bindung von Transkriptionsfaktoren erschweren, überdacht und in verbesserten theoretischen Modellen berücksichtigt werden.

Eine zusätzliche Herausforderung bei der Modellierung, die sich der experimentell bestimmten Positionen der Nukleosomen zur Erstellung von Monte-Carlo-Simulationen der 3-D-Faltung des Genoms bedient, entsteht durch die Heterogenität der untersuchten Zellpopulationen. In vielen Fällen waren die experimentell ermittelten Nukleosomenpositionen nicht mit einer einzelnen Nukleosomenkettenkonformation kompatibel. Um spezifische Nukleosomenanordnungen und optimierte Lösungen für die komplexen Muster aus experimentellen Daten zu erhalten, entwickelten wir ein Konzept, das die Analyse binärer Variablen und das Monte-Carlo-Verfahren mit einem als „*simulated annealing*“ bezeichneten Optimierungsverfahren

kombiniert (Schöpflin *et al.*, 2013). Mit diesem Ansatz führen wir nun basierend auf den experimentellen Daten des Populationsdurchschnitts die Computersimulationen für spezifische genomische Regionen von bis zu 1.000 Nukleosomen durch, um deren räumliche Anordnung und die Zugänglichkeit der DNA zu untersuchen (siehe Abbildung 3).

Wie geht es weiter?

Die Nukleosomenpositionierung ist ein dynamischer Prozess, der in Zellen durch energieverbrauchende molekulare Maschinen reguliert wird, die als „Chromatin-Remodellierungskomplexe“ (CRs) bezeichnet werden. Sie können Nukleosomen entlang des DNA-Strangs bewegen oder sogar von diesem entfernen und so zwischen „aktivem“ und „inaktivem“ Zustand der DNA umschalten. Interessanterweise zeigen genomweite Sequenzierungsstudien, wie sie beispielsweise im Internationalen Krebsgenomkonsortium (<http://dcc.icgc.org>) durchgeführt werden, bei Krebserkrankungen eine überraschend hohe Anzahl von Mutationen (also Genveränderungen) in verschiedenen CRs. Bei Brust-, Eierstock-, Nieren-, Lungen-, Gebärmutter- und Leberkrebs wurden potenzielle „Treibermutationen“ (also maßgebliche Mutationen, die weitere Veränderungen im Genom auslösen) bei einigen CR-Familien des Typs CHD und SMARCA identifiziert. Darüber hinaus ist bekannt, dass abweichende epigenetische Muster die Aktivität der CRs fehlleiten und somit pathologische Zustände des Chromatins in Tumoren induzieren.

Daher ergeben sich aus den im Rahmen des *EpiGenSys*-Projekts erhaltenen Einblicken neue wichtige Fragen: Wie wird die aktive Translokation (also Bewegung entlang des DNA-Strangs) von Nukleosomen in den Zellen gesteuert? Unser aktuelles Modell legt nahe, dass Signale, die an die CRs gerichtet sind und deren Aktivität verändern, in der nukleosomalen DNA-Sequenz, sowie in den posttranslationalen Modifikationen der Histone H3 und H4 und anderen Chromatineigenschaften wie der DNA-Cytosin-Methylierung verschlüsselt sind. Es scheint, als gäbe es einen bisher unbekanntes „Chromatin-Remodellierungs-Code“, der die CRs auf Genombereiche lenkt, in welchen sie die Positionen der Nukleosomen verändern. Auf diese Weise können CRs spezielle lokale Chromatinstrukturen etablieren und sollten daher nicht länger als Faktoren betrachtet werden, die Nukleosomen unspezifisch verschieben und das Chromatin „flüssiger“ machen. Zudem deutet unsere bisher unveröffentlichte Modellierungsarbeit auf eine

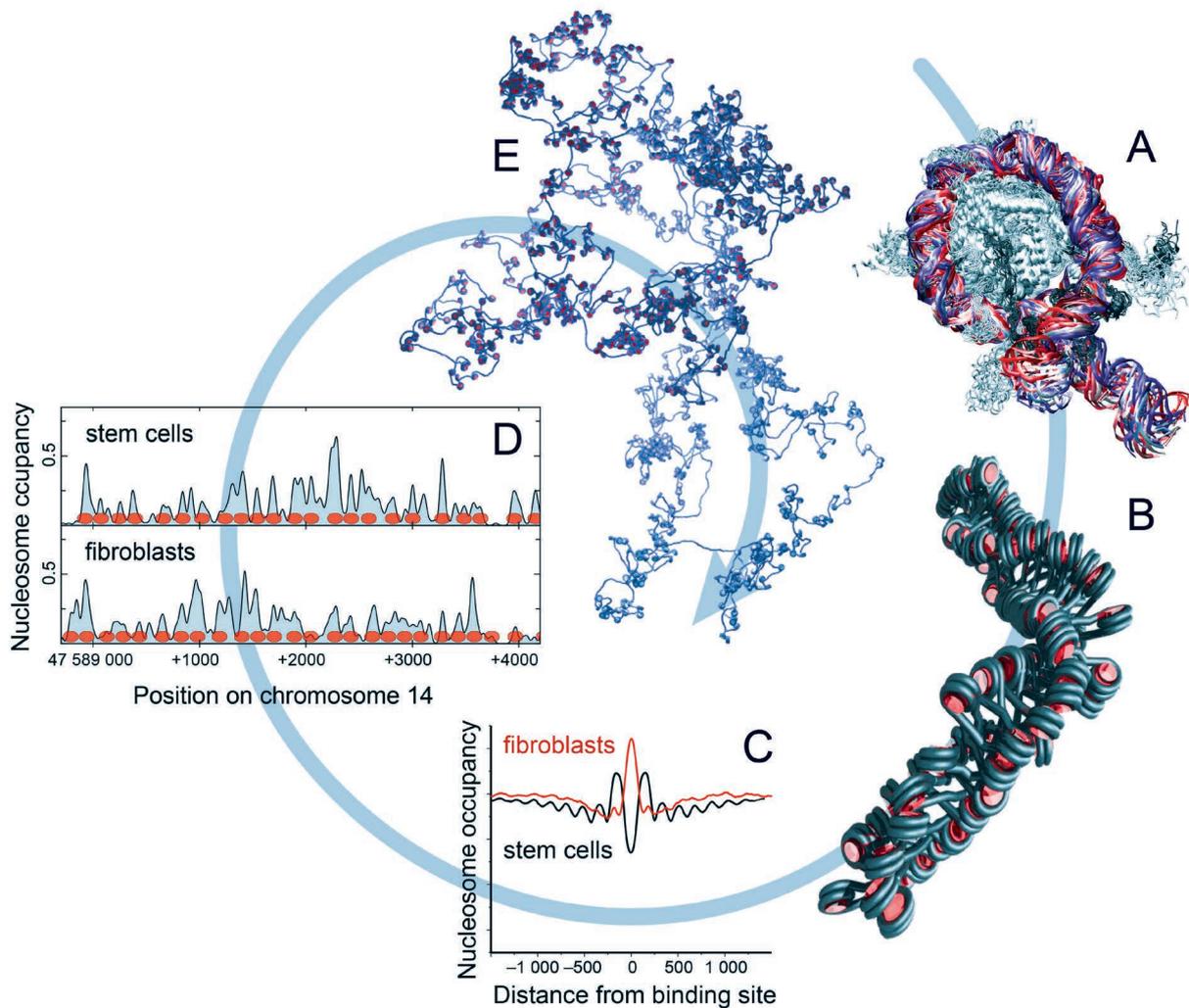


Abbildung 3: Informationsschleifen zwischen Modellierung und experimentellen Studien geben Aufschluss über das Verhältnis zwischen Nucleosomen und der DNA-Zugänglichkeit für die Transkriptionsfaktorbindung

- A)** Mithilfe von Molekulardynamik-Simulationen wurde der Zugang zur nucleosomalen DNA in Bezug auf die kompetitive DNA-Bindung des Histonproteinkerns und der Transkriptionsfaktoren untersucht (Ettig *et al.*, 2011; Teif *et al.*, 2010). Die Abbildung zeigt Nucleosomenkonformationen, die in 0,2-Nanosekunden-Intervallen übereinandergelegt wurden. Die DNA ist entsprechend der Simulationszeit zunehmend von Rot über Weiß zu Blau dargestellt. Die Kernhistonproteine sind weiß dargestellt. Bereits während des sehr kurzen Simulationszeitraums von zwei Nanosekunden zeigt sich die Nucleosomenkonformation sehr dynamisch.
- B)** Die Monte-Carlo-Simulationen der Faltung der Nucleosomenkette liefern Informationen zum Zugang zur Verbindungs-DNA in den Chromatinfasern (Rippe *et al.*, 2012).
- C)** Experimentell erstellte Besetzungsprofile der Nucleosomen an den Bindungsstellen des CTCF-Transkriptionsfaktors (Teif *et al.*, 2012). Bei embryonalen Stammzellen sorgt das gebundene CTCF-Protein für regelmäßige Abstände zwischen den Nucleosomen im Bereich um dessen Bindungsstelle. Bei differenzierten Fibroblastenzellen besetzen Nucleosomen einige der Stellen, an die in Stammzellen CTCF gebunden ist.
- D)** Aus experimentellen Datensätzen konnten eindeutige Positionierungen der Nucleosomen für die anschließende Modellierung von Chromatin bei spezifischen Genorten bestimmt werden (Schöpflin *et al.*, 2013).
- E)** Monte-Carlo-Simulation der Chromatinanordnung am Locus des 200 kb großen SAMD4-Gens basierend auf der experimentellen Bestimmung der Nucleosomenpositionierung (Quelle: Robert Schöpflin).

weitere Verbindung zwischen den Positionen der Nucleosomen und der Chromatinstruktur hin: Bereits durch die Veränderung einer einzelnen Nucleosomenposition kann die übergeordnete Faltung der Nucleosomenkette erheblich modifiziert werden. Folglich regulieren die lokalen Positionen der Nucleosomen den Zugang zur Verbindungs-DNA über die Veränderung der Chromatinfaserkonformation. Gleichzeitig können sie Interaktionen zwischen Proteinen, die an weit voneinander entfernten Stellen gebunden sind, fördern oder erschweren.

Angesichts der Bedeutung der Positionierung von Nucleosomen für die Zugänglichkeit und das Auslesen des Genoms ergibt sich die Frage, ob Fehlsteuerungen der Nucleosomenpositionierung bei Krebszellen mit Vorhersagen zum Therapieansprechen in Verbindung gebracht werden können. Mit diesem Punkt beschäftigt sich das Verbundprojekt *CancerEpiSys* (www.CancerEpiSys.org) im Rahmen des BMBF-geförderten CancerSys-Programms. Die Nucleosomenpositionen und epigenetischen Modifikationen in primären Tumorzellen von Patienten mit chronischer lymph-

tischer Leukämie werden kartiert, um Muster der deregulierten Chromatinanordnung zu identifizieren, die dem Stadium der Krebserkrankung zugeordnet werden können. Die bisherigen Ergebnisse lassen darauf schließen, dass die Einbeziehung der zahlreichen epigenetischen Signale des Genoms in Bezug auf die Veränderungen des DNA-Zugangs wertvolle Informationen für neue individualisierte Therapieansätze ergibt. Daher wird sich die Integration der Nukleosomenpositionen zu einem wesentlichen Bestandteil quantitativer Beschreibungen der Genregulation entwickeln. Diese erlauben Vorhersagen darüber, wie Zellen ihr aktives genetisches Programm auswählen, und wie dieser Prozess in Krebszellen dereguliert wird.

Steckbrief Forschungsprojekt:

Die Arbeiten zur Modellierung von Nukleosomen und Chromatin sowie die experimentelle Bestimmung der Nukleosomenpositionierungen waren Teil des dreijährigen, BMBF-geförderten Verbundprojekts *EpiGenSys* – Systembiologische Bestimmung der epigenomischen Struktur-Funktionsbeziehung (www.EpiGenSys.org) im Rahmen der europaweiten Initiative ERASysBioPlus im EU FP7 ERA-NET Plus Programm (Förderkennzeichen 0315712A, Karsten Rippe, und 0315712C, Gero Wedemann).

Im Rahmen des Projekts *EpiGenSys* wurden (epi-)genomische Struktur-Funktionsbeziehungen zwischen Chromatin und Transkription in einem mehrstufigen Ansatz analysiert und modelliert. Die Arbeit befasste sich mit der Struktur einzelner Nukleosomen, der Organisation der Nukleosomenkette in übergeordnete Faserstrukturen sowie der globalen 3-D-Architektur des Genoms. Innerhalb des Verbunds führten die Gruppen von Karsten Rippe und Gero Wedemann Modellierungsarbeiten auf Ebene einzelner Nukleosomen in Bezug auf den Zugang zur nukleosomalen DNA aus (Ettig *et al.*, 2011; Teif *et al.*, 2010). Außerdem entwickelten sie grobkörnige Modelle für die Nukleosomenkette mithilfe von Monte-Carlo-Simulationen (s. Übersichtsartikel Rippe *et al.*, 2012), der experimentell durch Hochdurchsatzsequenzierung bestimmten Positionierung der Nukleosomen (Teif *et al.*, 2012) sowie bioinformatischer Ansätze (Schöpflin *et al.*, 2013).

Referenzen:

- Ettig, R., Kepper, N., Stehr, R., Wedemann, G., und Rippe, K. (2011). Dissecting DNA-histone interactions in the nucleosome by molecular dynamics simulations of DNA unwrapping. *Biophys J* 101, 1999 – 2008.
- Rippe, K., Stehr, R., und Wedemann, G. (2012). Monte Carlo simulations of nucleosome chains to identify factors that control DNA compaction and access. In *Innovations in Biomolecular Modeling and Simulations*, T. Schlick, ed. (Cambridge: Royal Society of Chemistry), pp. 198 – 235.
- Schöpflin, R., Teif, V. B., Müller, O., Weinberg, C., Rippe, K., und Wedemann, G. (2013). Modeling nucleosome position distributions from experimental nucleosome positioning maps. *Bioinformatics* 29, 2380 – 2386.
- Teif, V. B., Ettig, R., und Rippe, K. (2010). A lattice model for transcription factor access to nucleosomal DNA. *Biophys J* 99, 2597 – 2607.
- Teif, V. B., Vainshtein, Y., Caudron-Herger, M., Mallm, J.-P., Marth, C., Höfer, T., und Rippe, K. (2012). Genome-wide nucleosome positioning during embryonic stem cell development. *Nat Struct Mol Biol* 19, 1185 – 1192.

Kontakt:



Karsten Rippe

Deutsches Krebsforschungszentrum (DKFZ)
& BioQuant, Heidelberg

Leiter der AG „Genomorganisation
und -funktion“

Karsten.Rippe@dkfz.de

<http://malone.bioquant.uni-heidelberg.de>



Gero Wedemann

Fachhochschule Stralsund

Leiter des Institute for Applied Computer
Science

Leiter der AG „Competence Center
Bioinformatics“

gero.wedemann@fh-stralsund.de

www.bioinformatics.fh-stralsund.de