ETH zürich

ETH Library

Monitoring DNA double-stranded break repair (DSBR) homology search in situ via RAD51 proximity labeling

Doctoral Thesis

Author(s):
Yeh, Charles D.

Publication date:
2024

Permanent link:
https://doi.org/https://doi.org/10.3929/ethz-b-000709733

Funding acknowledgement:
- Molecular mechanisms of human genome editing ()

This page was generated automatically upon download from the ETH Zurich Research Collection.
For more information, please consult the Terms of use.

DISS. ETH NO. 30555

Monitoring DNA double-stranded break repair (DSBR) homology search *in situ* via RAD51 proximity labeling

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES
(Dr. sc. ETH Zurich)

presented by

CHARLES D YEH
B. Sc., University of Maryland, Baltimore County

born on 28.04.1992

accepted on the recommendation of

Prof. Dr Jacob E. Corn
Prof. Dr. Gerald Schwank
Prof. Dr. Berend Snijder
Prof. Dr. Joao Matos

2024

ABSTRACT (ENGLISH)

Maintaining genome integrity is a vital function in living organisms, and accumulating uncorrected DNA damage may lead to malfunction and pathogenesis. However, DNA damage takes on a wide variety of forms and there are an equally diverse number of DNA repair pathways. Among the most toxic DNA damage events are double-stranded breaks (DSBs). In many forms of damage only one strand of DNA is affected, and this damage is readily corrected using the complementary strand as an intact donor template. In DSBs, this possibility does not exist, and it is then unknown whether mutations have occurred at the break site. A choice must then be made, find a template to restore any lost information or repair the DSB with the risk of mutations.

Whereas non-templated repair is fast, this process is often mutagenic, leading to insertion or deletion (indel) mutations. Templated repair—also called homology directed repair (HDR)—is significantly slower in human cells and requires a careful comparison of the DSB to many *candidate* DNA sequences to copy information from before ultimately *choosing* one. This process is called homology search and is a key rate limiting step during HDR. The ideal HDR donor template is often the sister chromatid, a duplicate of the genome made in preparation for cell division. However, only the small region exactly corresponding to the DSB site should be used, recombination with any other genetic sequence will cause mutagenesis. Therein lies the problem, finding a single short DNA sequence out of an enormous background, being 6.3 billion DNA bases in humans.

Despite the importance of homology search in HDR, limited methods exist to study this process. The available techniques are constrained by low-throughput, poor sensitivity, or lack of DNA sequence resolution. Thus, in this work we develop a novel method that enables highly sensitivity, *in situ*/*in vivo*, and sequence-resolved monitoring of HDR homology search. This method extends upon an earlier DNA-protein proximity labelling method called DamID. As the key protein involved in DSBR homology search is RAD51, we term our

novel method **RA**D51 **p**roximity **id**entification **seq**uencing (RaPID-seq). Using RaPID-seq, we provide new insight into homology search dynamics in human cells. In contrast to bacteria and budding yeast, we find that homology search is constrained to the region proximal to the DSB site. Furthermore, we extend the findings and method of RaPID-seq into studying the dynamics of genome editing and multiplexing with other important molecular biology techniques, demonstrating the potential of RaPID-seq as a core technique of DNA repair research.

ABSTRACT (DEUTSCH)†

Die Aufrechterhaltung der Integrität des Genoms ist eine kritische Funktion in lebenden Organismen, und die Anhäufung unkorrigierter DNA-Schäden kann zu Fehlfunktionen und pathogenen Resultaten führen. Da DNA-Schäden jedoch sehr unterschiedliche Formen annehmen können, gibt es eine ebenso große Anzahl von DNA-Reparaturwegen. Zu den gefährlichsten DNA-Schäden gehören die DNA-Doppelstrangbrüche (DSB). Bei vielen Formen von Schäden ist nur ein DNA-Strang betroffen welcher sich leicht mit dem komplementären Strang als intakter Vorlage korrigieren lässt. Bei DSBs gibt es diese Möglichkeit jedoch nicht und Information über allfällige Mutationen an der Bruchstelle sind aus diesem Grund nicht vorhanden. Die Zelle muss eine Entscheidung treffen: entweder eine Vorlage wird gefunden um die verlorene Information wiederherzustellen, oder der DSB wird mit dem Risiko von Mutationen repariert.

Während die Reparatur ohne Vorlage schnell ist, kann dieser Prozess oft fehlerhaft sein und zu Insertions- oder Deletionsmutationen (Indel) führen. Die Reparatur mit Vorlage hingegen - auch Homologie geleitete Reparatur (HDR) genannt - ist in menschlichen Zellen deutlich langsamer und erfordert einen sorgfältigen Vergleich des DSBs mit vielen potentiell kopierbaren DNA-Sequenzen, bevor eine ausgewählt werden kann. Dieser Vergleichsprozess wird als Homologiesuche bezeichnet und ist ein entscheidender, geschwindigkeitsbegrenzender Schritt bei der HDR. Die ideale HDR-Vorlage ist häufig das Schwester Chromatid, ein Duplikat des Genoms welches als Vorbereitung für die Zellteilung hergestellt wird. Jedoch sollte nur der kleine, genau der DSB-Stelle entsprechende Bereich verwendet werden, da die Rekombination mit jeder anderen genetischen Sequenz zu Mutationen führt. Genau hier liegt das Problem; eine einzelne kurze DNA-Sequenz in einem riesigen Hintergrund zu finden, der beim Menschen 6,3 Milliarden DNA-Basen umfasst.

† Translated by Lilly van de Venn.

Trotz der entscheidenden Bedeutung der Homologiesuche bei der HDR gibt es bis jetzt nur wenige Methoden zur Untersuchung dieses Prozesses und die momentan verfügbaren Techniken sind durch geringen Durchsatz, geringe Empfindlichkeit oder fehlende Auflösung der DNA-Sequenz eingeschränkt. In dieser Arbeit entwickeln wir daher eine neue, hochempfindliche, in situ/in vivo, sequenzaufgelöste Methode, die eine Überwachung der HDR-Homologiesuche ermöglicht. Diese Methode baut auf einer existierenden Methode zur Markierung von DNA-Protein-Kontakten namens DamID auf. Da das Schlüsselprotein bei der DSBR-Homologiesuche RAD51 ist, nennen wir unsere neue Methode RAD51 Proximity Identification Sequencing (RaPID-seq). Mithilfe von RaPID-seq erhalten wir neue Einblicke in die Dynamik der Homologiesuche in menschlichen Zellen. Im Gegensatz zu Bakterien und Hefen stellen wir fest, dass die Homologiesuche auf die Region in direkter Nähe zur DSB-Stelle beschränkt ist. Darüber hinaus erweitern wir die Erkenntnisse und die Methode von RaPID-seq auf die Untersuchung der Dynamik der Genom-Modifizierung und des Multiplexing mit anderen wichtigen molekularbiologischen Techniken, welches das Potenzial von RaPID-seq als Kerntechnik für die DNA-Reparaturforschung aufzeigt.