Wikipedia Page for XR-seq and Damage-seq

Su Sarlar susarlar@sabanciuniv.edu
BsBIO/FENS, Senior

Ali Kağan Akbaş akagan@sabanciuniv.edu
Undeclared/FENS, Freshman

Ogün Adebali
Molecular Biology, Genetics, Bioengineering

Abstract

The recent next-generation sequencing technologies, Damage-seq and XR-seq help us map DNA damage and nucleotide excision repair, respectively, onto genomes of interest. These state-of-the-art technologies developed by Aziz Sancar's team at the University of North Carolina, are only known by the limited scientific environment. We wrote semi-scientific and informal texts for each method to spread these techniques out. Wikipedia is one of the most appropriate platforms to publish such a write-up. For this project, we have merged the literature on XR-seq and Damage-seq and wrote two separate wiki pages for the techniques.

Keywords: XR-seq, Damage-seq, genomic damage, repair, DNA

1 Introduction

XR-seq is an excision repair sequencing protocol used in molecular biology. It is developed to study nucleotide excision repair of DNA adducts in humans, mice, Arabidopsis thaliana, yeast and Escherichia coli. Unlike the current genome-wide approaches that resolution limit caused by the loss of damage signal with time, XR-seq is favorable to detect repair signal above a background of zero. Damage-seq is a technology for mapping bulky base lesions such as cyclobutene pyrimidine dimers (CPD), pyrimidine-pyrimidone photoproducts (6-4 PP's) and cisplatin induced diadducts at single nucleotide resolution. Damage-seq can be used for complementing XR-seq libraries to investigate factors affecting damage repair. Damage-seq maps the location of bulky base lesions, using the lesions block replication. Repair is measured indirectly, as the difference of damage between a later time point and the initial time point, so the collective repair up to the later time point can be obtained (1).

2 Damage seq

2.1 Introduction:

DNA lesions such as: cyclobutane pyrimidine dimers, pyrimidine-pyrimidone photoproducts, caused by UV damage, cisplatin induced diadducts, intrastrand diadducts, interstrand cross-links, caused by anticancer drugs damaging DNA, are essentially distributed across genome in a uniform manner. These bulky DNA lesions are repaired by excision repair, and the ongoing repair at a single time can be mapped via XR-seq. Damage-seq and XR-seq maps complement each other.

Comparison of Damage-Seq and XR-seq shows that while the DNA damage is distributed rather homogenously, while the repair is heterogenous. Repair efficiency is dependent on chromatin states, transcription, transcription factor binding, and type of DNA damage. Repair efficiency is reduced in nucleosome center and increased in DNase hypersensitive sites (1).

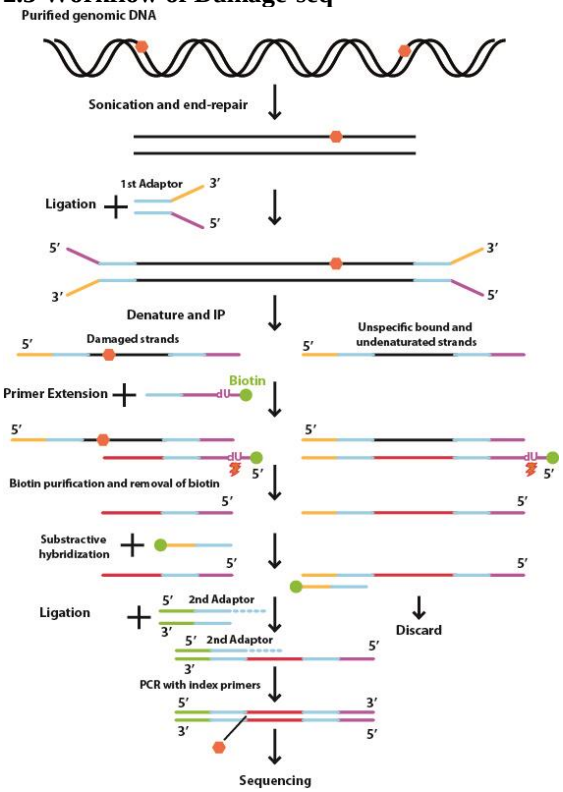
For damage-seq, cells or naked DNA is subjected to DNA damage causing factors, DNA is isolated, sheared, ligated to adapters, denatured and incubated with antibody coated beads. Primers are annealed to adapters and are extended until the presence of a DNA lesion, extended primers are captured and a second adapter is ligated, products are amplified using index primers. Samples are sequenced and unique aligned reads are produced. Reads are aligned to reference genome. Damage signal is viewed. Understanding how DNA repair efficiency changes across human genome would help cancer prevention and chemotherapy.(1)

2.2 Uses and current research:

Damage-seq maps provide information regarding positions of damage, damage formation and repair efficiency.
Transcription factor binding inhibits excision repair and may inhibit, stimulate or not affect photoproduct formation depending on type of transcription factor and strand. Transcription factor binding tends to inhibit cisplatin induced damage formation.
Comparison of XR-seq and Damage-seq maps shows that TCR for the CPD reaches a maximum in one hour. Between the damage profiles at the time points 0 and 1 hour there is no difference as the absolute repair levels in transcription sites are negligible compared to total damage. TCR peaks at 8th hour and gradually decreases until 48th hour. This decrease can be allocated to the decrease of damage in the transcription site and increase of repair in flanking regions. The difference in damage levels is bigger in transcription start sites than it is in transcription end sites. In human fibrioblasts, T-C 6-4 PP's are repaired faster than T-T 6-4 PP's.

For CPD damage, transcribed regions of DNA is repaired more efficiently compared to their complementary strands and non-transcribed regions. 6-4 PP is repaired efficiently by core repair machinery. Because 6-4 PP damage is less affected by TCR than CPD damage, both strands are repaired in a rather homogenous manner.

2.3 Workflow of Damage-seq



Damage Induction:

Genomic DNA is damaged using anticancer drugs, such as cisplatin or UV. For drug treatment, a solution is prepared by dissolving cisplatin or similar drug in DMSO to 20 µM. 5 µg DNA is incubated with 20 µM cisplatin solution in a final volume of 50 µL at room temperature for 15 minutes. Then DNA is purified and subjected to

Damage-seq:

For UV treatment different dosages lead to different types of DNA damage in different cell lines. In GM12878 cell line irradiation with 20 J/m2 causes CPD and 6-4 PP formation while in NHF1 line irradiation with 20 J/m2 causes 6-4 PP formation and 10 J/m2 causes CPD formation. Naked DNA is irradiated with 20 J/m2 UVC.

DNA preparation:
Genomic DNA is extracted by PureLink Genomic DNA kit (Thermo). Damaged DNA is fragmented down to an average of 400-600 base pair pieces by sonication. Fragments of 200-700 base pair in length are selected by HighPrep PCR beads. The fragments are end repaired, ligated to first adapters. Ligated DNA is denatured by addition of 20 μL 8 M urea and boiling for 1 minute to obtain single strands.
Immunoprecipitation
Obtained single stranded DNA's are incubated with antibody coated beads.
DNA amplification/Library preparation
A biotinylated primer is annealed to adapter. The primer is extended by a high-fidelity DNA polymerase. If there is a bulky lesion, extension stops. Extended primers are captured by streptavidin coated beads that capture biotin. A second adapter is ligated. PCR using index primers is done.
Library quality check
Using a small amount of ligation products, two PCR's with different primer sets are done. In the first, the primers anneal to the first adapter and in the second, primers anneal to the second adapter. In the first, only DNA strand without lesions are amplified. In the second, all DNA strands, damaged or not, are amplified. Then the first and second PCR results are compared. Products of second PCR will be considerably shorter than products of first PCR if a bigger proportion of products comes from DNA strands with bulky lesions. Products of second PCR will be slightly longer than products of first PCR if the bigger proportion of products comes from undamaged DNA strands.
Sequencing, Genome Alignments and Visualization
For reference genome sequencing, DNA from undamaged cell lines is sonicated. 100 ng of fragmented DNA is used for library preparation. Three percent of ligation products are amplified by PCR and products are sequenced from both ends. Samples are sequenced to at least 10 million mapped reads. Libraries are sequenced for producing paired end 50nt reads. This way, unique aligned reads that allow for discerning damage hotspots and amplification artifacts are established. Products of undamaged DNA are filtered out. Paired end reads are aligned to the reference genome. Duplicated reads are reduced to a single read. Damage position and nucleotide composition are determined relative to fragment. Common damage types are selected. To view damage signal, bigwig files are generated using read counts per 25-nt windows normalized to total reads of each chromosome.

3 XR-Seq

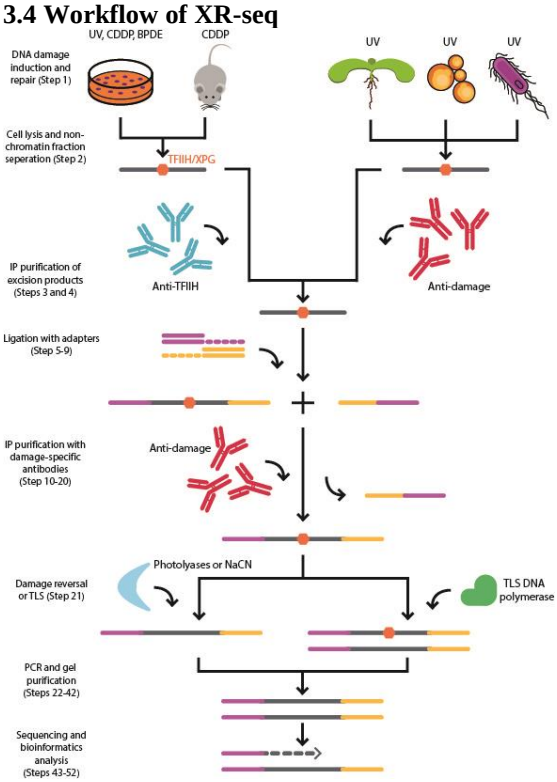
3.1 Introduction:
Nucleotide excision repair mechanism plays significant roles such as removing bulky DNA adducts from the genome, including the major UV photoproducts and lesions caused by other environmental mutagens. A flaw in this repair mechanism increases the lethality and mutagenicity of these agents enormously. As an example, deficiency of the nucleotide excision repair leads to xeroderma pigmentosum, which is a very rare skin disorder that creates high solar sensitivity in humans.
Nucleotide excision repair is influenced by various factors in vivo such as transcription, chromatin states, DNA replication, epigenetic modifications of DNA and histones, binding of regulatory proteins to DNA and other factors. An approach is necessary to study the effects of these factors on the repair with high resolution at a genome-wide scale. There are some high-resolution methods that have already exist for mapping DNA damage sites in the entire genome. However, they indirectly measure nucleotide excision repair by determining the disappearance of damage at the interval of two time points, which limits their sensitivity. Unlike the previous methods, XR-seq is an appropriate and a practicable genome-wide approach to study nucleotide excision repair of bulky DNA adducts in vivo at single-nucleotide resolution. First, XR-seq developed in cultured human cells but then it is successfully adapted to cellular organisms such as mice, A. thaliana, yeast and E. coli.
Therefore XR-seq can be applied to different species and biological systems damaged by multiple substances to investigate repair mechanism and other biological processes. These practices may help us to clarify basic biological functions. They can be helping to improve the efficacy of therapeutic interventions and create crops that have more efficient repair mechanisms to deal effectively with DNA-damaging environmental stresses.
3.2 Development and overview of XR-seq:
Previous mechanistic studies of nucleotide excision repair indicate that dual incisions are made in the damaged DNA strand on both sides of the damage. So, the product is a damage-containing single-stranded oligomer. These oligomers that emerged during repair reactions are the focal point of the XR-seq research.

3.3 Advantages and applications:

Genome-wide analysis of nucleotide excision repair of distinct damaging agents in a wide range of organism and formats is possible by using XR-seq protocol. Also, it is possible to probe both gene activity and regions of open chromatins. Another advantage of the XR-seq study is that it allows detailed illustrations of circadian rhythmicity in transcription and repair in plants and mice, at the organismal level.

It has been applied to investigate species-specific and damage-specific characteristics of repair, regulation of repair by chromatin structure and states besides. The 'chromatin state' of a genomic region can refer to the set of chromatin-associated proteins and histone modifications in that region. Data collected from XR-seq studies have been used to investigate "how repair heterogeneity determines the distribution of mutation frequency in some cancers".

3.4 Workflow of XR-seq



3.5 Limitations:

The most important limitation of XR-seq is the inability to measure true rates of repair. It is a good way to use XR-seq to determine relative rates, but actual repair rates cannot be calculated due to excised oligomers are concurrently formed and degraded, and rates of degradation in different species are unknown.

4 Conclusion and Future Work

Two drafts, on Damage seq and XR seq, were prepared. The writings should be periodically updated to reflect state of the art, as new implementations occur.

References:

Dynamic maps of UV damage formation and repair Hu, J, Adebali O, Adar S, Sancar S
Proceedings of the National Academy of Sciences Jun 2017, 114 (26) 67586763;
DOI:10.1073/pnas.1706522114

Cisplatin DNA damage and repair maps of the human genome at single-nucleotide resolution. Hu J, Lieb JD, Sancar A, Adar S (2016) Proc Natl Acad Sci USA 113:11507–11512.

.