INDUCE-seq: a novel tool for next-generation risk assessment

National Centre for the Replacement Refinement & Reduction of Animals in Research

Cardiff University / Prifysgol Caerdydd — School of Medicine / Ysgol Meddygaeth

Introduction

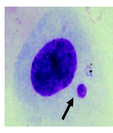
Rationale: INDUCE-seq was originally developed for the detection of double strand breaks (DSBs) induced by CRISPR-Cas9 nucleases. However, as an unbiased break detection method it can also be utilised for the detection of chemically-induced DSBs.

Aim: Develop INDUCE-seq for the quantitative detection of chemically-induced DSBs, with a focus on developing bioinformatic tools for characterising induced breaks and distinguishing these breaks from the background of endogenous breaks.

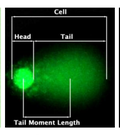
Approach: Treatment of cells with genotoxic chemicals, followed by break detection with INDUCE-seq. Analysis of break data as to determine break patterns associated with various chemicals. Extraction of latent information in order to elucidate the mechanisms of break induction.

Cell-based approaches

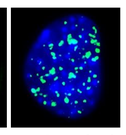
- Imaging/fluorescence-based
- Indirectly measures breaks
- Limited data output



Micronucleus test



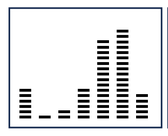
Comet assay



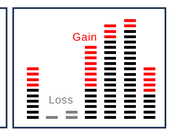
γH2AX foci imaging

Genomic techniques

- Sequencing-based
- Directly measures breaks
- Extensive data output



INDUCE-seq



Other NGS approaches

Next-generation risk assessment

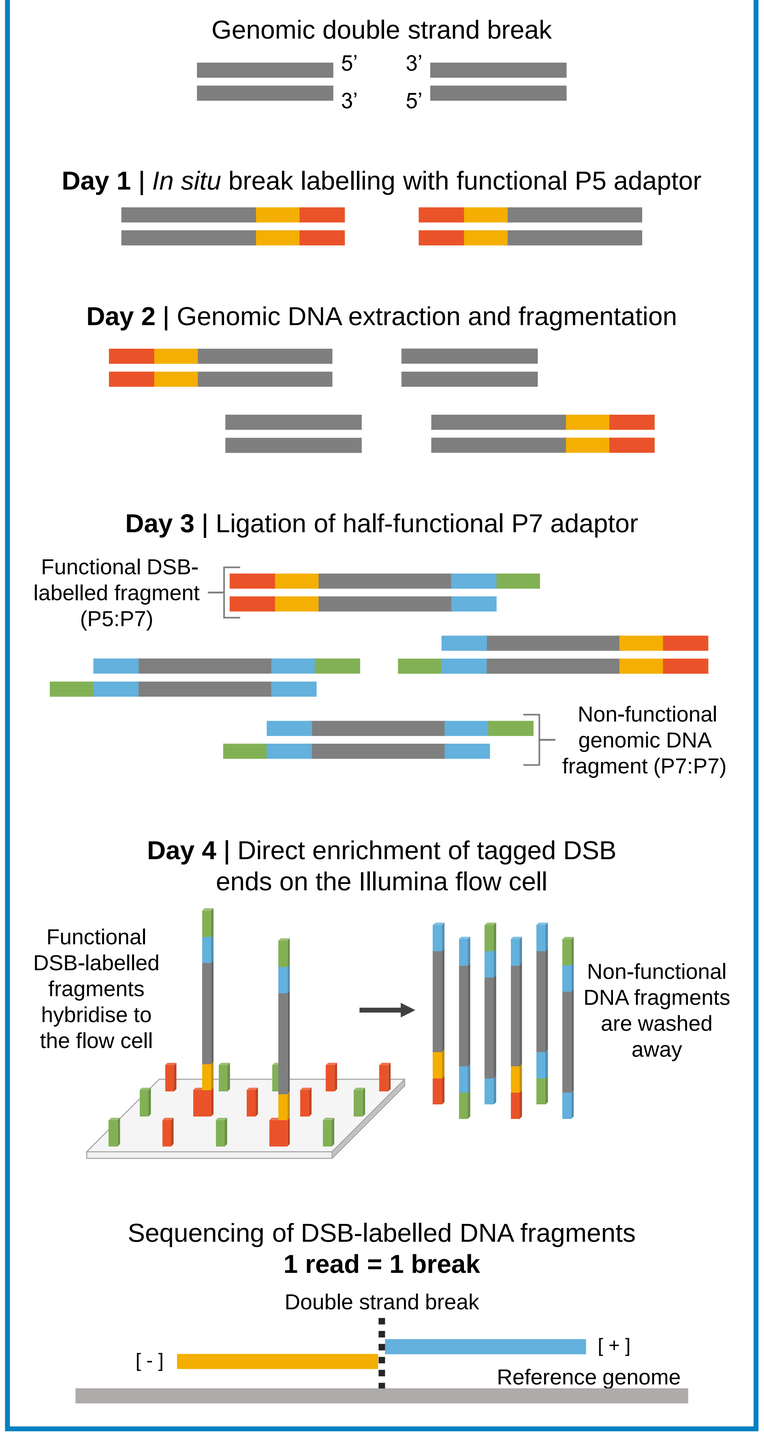
Integration of *in silico*, *in chemico* and *in vitro* approaches

Human-relevant

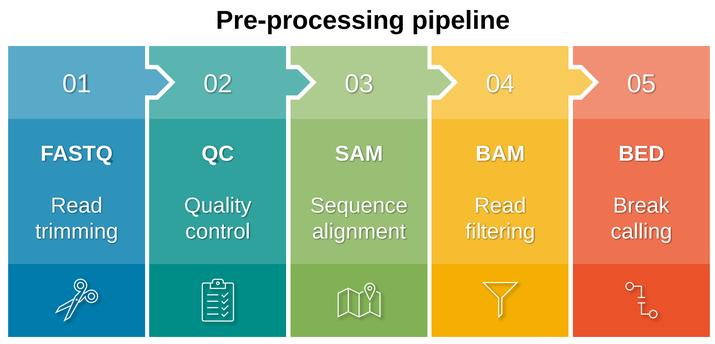
Exposure-led

Hypothesis-driven

INDUCE-seq methodology

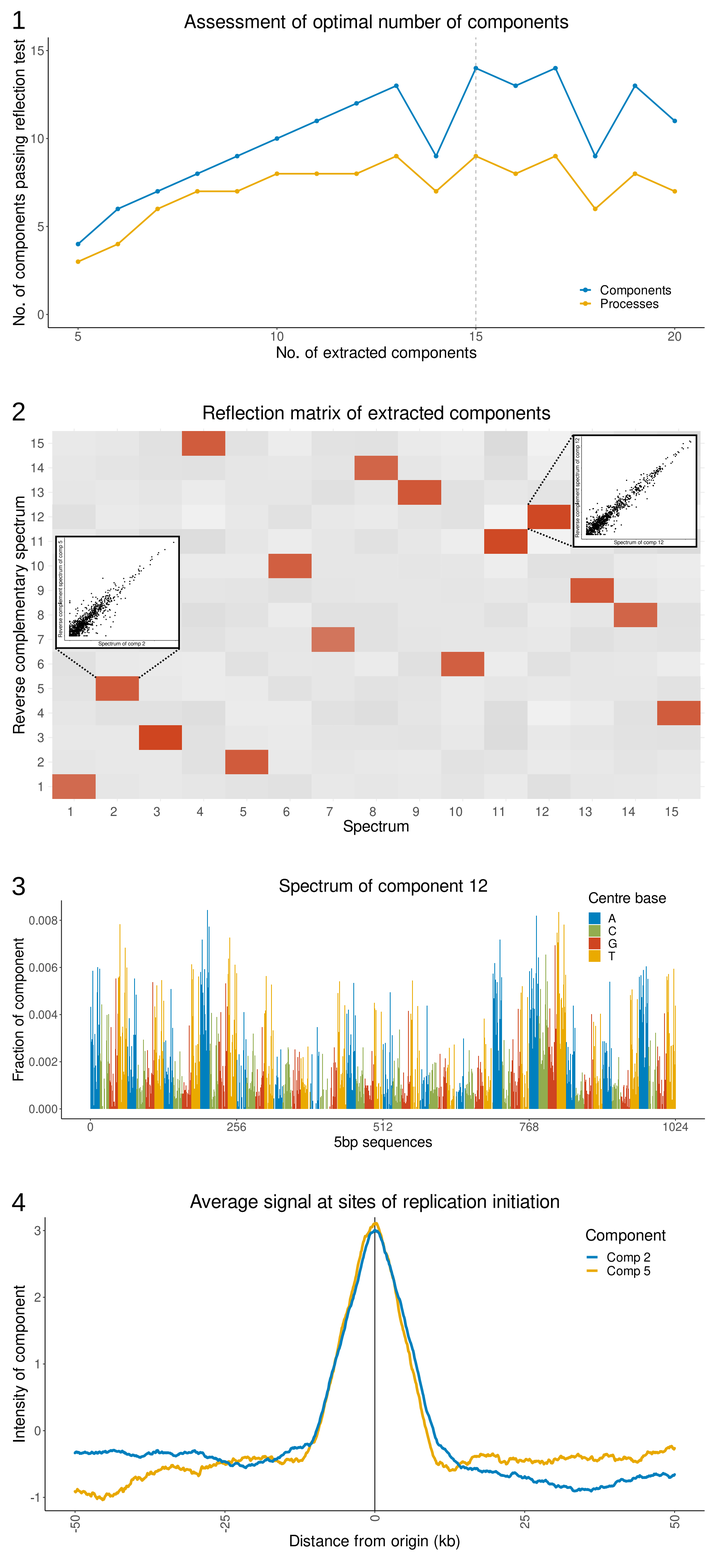


Data analysis



Latent features represent meaningful components extracted from break data and as such can be utilised to **classify breaks** based on the **mechanism** by which they were induced

Results



3Rs - Replacement

Development of **INDUCE-seq** for the quantitative detection of **chemically-induced breaks** in human cells

Provides additional genomic DNA damage to add to the various **omics methods** in development

Replace the use of rodents in industry sectors where animal safety studies are still used

These approaches represent the necessary tools for establishing **NGRA**

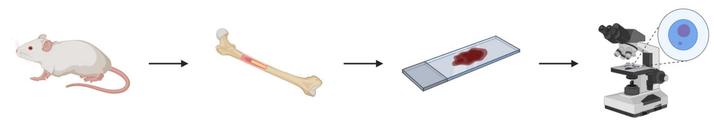
Most-widely used genotoxicity test is currently the **micronucleus assay**

The aim of NGRA is to increase acceptance of **non-animal data** in safety decision making

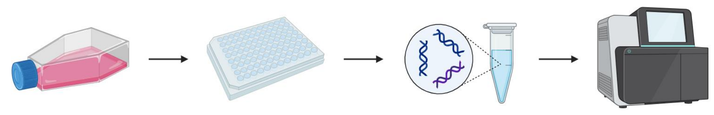
MN test involves **dosing rodents** with chemicals and sacrificing them to examine the bone marrow

INDUCE-seq is an alternative approach for **genotoxicity risk assessment studies** that utilises next-generation sequencing to directly measure DSBs

***In vivo* micronucleus assay** (historical approach)



INDUCE-seq (next-generation approach)



Discussion

INDUCE-seq is a novel genomic approach capable of unbiased genome-wide detection of chemically-induced breaks at a single-nucleotide resolution in human cells.

NMF has previously been used to extract signatures from catalogues of cancer mutations. These signatures have been associated with various mechanisms of DNA damage.

Volume-regularised NMF extracts a small number of signatures from an inventory of breaks and classifies the underlying processes as strand-independent or strand-dependent.

We hypothesise that chemicals with a direct mechanism of DNA damage will result in a new signature arising, associated with that chemical’s mechanism of action.

Conversely, chemicals with an indirect mechanism of DNA damage, disrupting endogenous processes, will change the distribution of pre-existing signatures.

Figure legends

Figure 1. The optimal number of components to extract by volume-regularised non-negative matrix factorisation (VRNMF). A range of components from 5 to 20 was evaluated, with 15 determined to be the optimal number to extract. Of these 15 components, 14 pass the reflection test, corresponding to 9 underlying processes. **Figure 2.** Reflection matrix of extracted components shows the correlation between components’ spectra and reverse complement spectra. Components with a spectrum that correlates with their own reverse complement are symmetrical. Components whose spectrum correlates with the reverse complement of another component are described as a pair of asymmetrical components. **Figure 3**. The complete spectrum of component 12. The spectrum shows the contribution of each of the 1024 5bp sequences to the overall profile of the signature. **Figure 4.** The average signal of components 2 and 5 relative to sites of replication initiation. The intensity of components 2 and 5 show an increase relative to sites of replication initiation.

References

Dent, M. et al. 2018. Principles underpinning the use of new methodologies in the risk assessment of cosmetic ingredients. *Computational Toxicology* 7, pp. 20–26

Dobbs, F.M., van Eijk, P., Fellows, M.D. *et al.* Precision digital mapping of endogenous and induced genomic DNA breaks by INDUCE-seq. *Nat Commun* 13, 3989

Liu, Z., Huang, R., Roberts, R. and Tong, W. 2019. Toxicogenomics: A 2020 vision. *Trends Pharmacol* Sci 40(2), pp. 92–103

Vladimir B. Seplyarskiy et al. ,Population sequencing data reveal a compendium of mutational processes in the human germ line.*Science* 373,1030-1035

Authors

Kierney O’Dare
Dr Patrick Van Eijk
Professor Simon Reed

Institutes

Division of Cancer and Genetics, School of Medicine, Cardiff University
Safety and Environmental Assurance Centre, Unilever

Acknowledgments

Many thanks to Dr Andrew Scott and Sophie Malcomber, from Unilever, for their continued support throughout this project. Funding for this project is provided by Unilever and the NC3Rs.

Unilever