

How statistics helps to unravel plant development

Bayesian inference in long-distance mRNA movement in plants

14 December 2022, Norwich, UK

Long-distance movement of mRNAs is important for plants to develop correctly. But these movements are rare events and the sequencing used to detect moving mRNAs is not perfect. So how do you know, if a promising signal in sequencing data reflects a moving mRNA, or is only a sequencing error? Scientists from the John Innes Centre in Norwich just reported key advances in differentiating between sequencing errors and truly mobile mRNAs with a statistical method called Bayesian inference.

Grafting the shoot of one plant on top of roots of another is frequently used to study long-distance signalling in plants. Some of the molecules that travel from one part of a graft to another are messenger RNAs (mRNAs), the long macromolecules made up of short pieces of genetic code that are copied from DNA and carry the information for making proteins. Sequencing is used to identify mRNAs. Most of the mRNAs you will find are those that have been locally copied from the DNA. But a few ones might have moved there from the other, grafted part of the plant.

These are the mRNAs that scientists in the EU-funded PLAMORF project want to study, but they are challenging to detect. Melissa Tomkins, Franziska Hoerbst and fellow scientists are part of the PLAMORF team that just published a solution to distinguish between “the noise” of sequencing errors of local mRNAs, and those that have moved in from other parts of the plant. They developed a method based on Bayesian inference to analyse sequencing data and accurately point out mRNA originating from a grafted part, as they just reported in the Royal Society Interface journal.

Finding the (mRNA) needle in the (sequencing) haystack

The genes of plants that belong to same species are highly similar, but rarely 100% identical – they have different “genetic backgrounds”. The positions in the DNA code that are distinct between plants from different genetic backgrounds are called single-nucleotide polymorphisms (SNPs). Scientists make use of these SNPs to differentiate between mRNAs that have been copied from the DNA of one plant, and mRNAs originating from a grafted plant with a different genetic background that might have moved there. But in sequencing data, these SNPs look like sequencing errors, which are also changes in single positions of genes.

Existing statistical methods differentiate poorly between SNPs and sequencing errors. Until now, this made it difficult for scientists to identify mobile mRNAs with confidence. A novel analysis method for sequencing data of grafted plants is about to change this. PLAMORF scientist Melissa Tomkins from the John Innes Centre comments: “Introducing Bayesian inference in sequence analysis is a game-changer. This mathematical method is our key that opens the door to finding mobile mRNAs in large sequencing datasets.”

100% accurate, 100% efficient

“Our Bayesian framework does not even need huge computing power because we were able to derive analytical solutions” explains Franziska Hoerbst from the John Innes Centre. The new framework is not only resource-efficient, but it also outperforms all existing methods to identify mobile mRNA. Using it on simulated data and published data from grafts, the framework exceeded all expectations – it performed nearly 100% accurately. Prof Richard Morris, Lead scientist in the John Innes Centre, sums up: “With this exciting advance, we are starting to uncover some really interesting patterns in those mRNAs for which we found statistical support for their movement. Beyond this specific application, we think the same methodology will be useful in other areas where detecting signals in noisy data is a challenge.”

For more information, please visit our website www.plamorf.eu or contact:

Prof. Richard J Morris
Lead scientist
John Innes Centre
richard.morris(at)jic.ac.uk

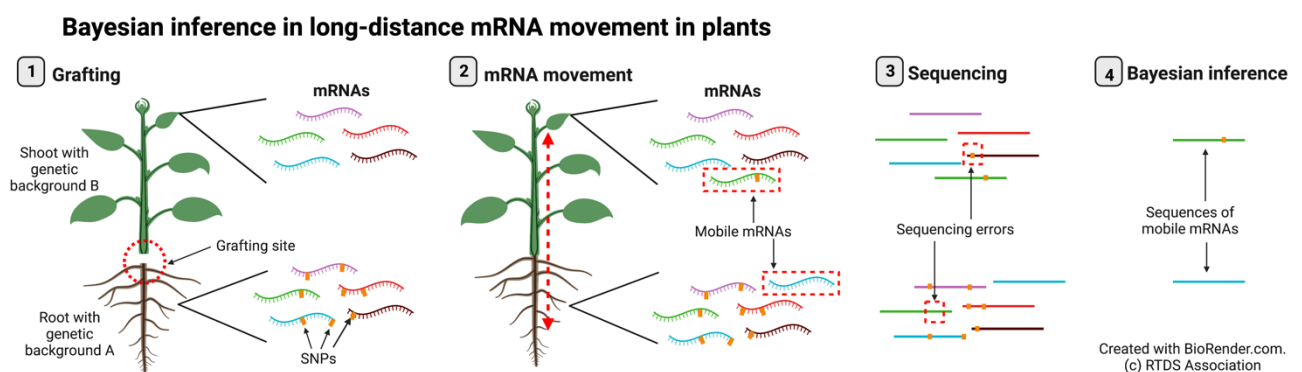
Adrian Galvin
Communications & Media Manager
John Innes Centre
adrian.galvin(at)jic.ac.uk

Original Publication:

Tomkins M*, Hoerbst F*, Gupta S, Apelt F, Kehr J, Kragler F, Morris RJ. 2022, Exact Bayesian inference for the detection of graft-mobile transcripts from sequencing data. J. R. Soc. Interface 20220644. <https://doi.org/10.1098/rsif.2022.0644>

*These authors contributed equally to this study.

Graphical abstract



(1) A shoot of plant with genetic background B is grafted on roots from genetic background A. The mRNAs differ in some positions in the genetic code, so-called SNPs. (2) Some mRNAs move long-distance over the graft junction. (3) Sequencing reveals the SNPs of mobile mRNAs, but they are hard to distinguish from sequencing errors in local mRNAs. (4) Bayesian inference helps to find the mRNA signals in the noise of sequencing errors.