



journal homepage: www.elsevier.com/locate/csbj



Review

Challenges in detecting and quantifying intron retention from next generation sequencing data



Lucile Broseus, William Ritchie*

IGH, Centre National de la Recherche Scientifique, University of Montpellier, Montpellier, France

ARTICLE INFO

Article history:

Received 3 September 2019
 Received in revised form 29 November 2019
 Accepted 15 February 2020
 Available online 26 February 2020

Keywords:

Intron retention
 mRNA splicing
 Bioinformatics
 RNA sequencing
 Gene expression

ABSTRACT

Intron retention (IR) occurs when an intron is transcribed into pre-mRNA and remains in the final mRNA. An increasing body of literature has demonstrated a major role for IR in numerous biological functions and in disease. Here we give an overview of the different computational approaches for detecting IR events from sequencing data. We show that these are based on different biological and computational assumptions that may lead to dramatically different results. We describe the various approaches for mitigating errors in detecting intron retention and for discovering IR signatures between different conditions.

© 2020 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Contents

1. Introduction	502
2. IR detection	502
2.1. Filtering sequencing data	502
2.2. Defining reference intronic sequences	502
2.2.1. The all-introns view	502
2.2.2. The measurable introns view	502
2.3. Overcoming sequencing and alignment artifacts	503
2.3.1. Overlapping features	503
2.3.2. Repeated regions	503
2.3.3. Low coverage of flanking exons	503
2.3.4. 3' Coverage bias	503
2.4. Implemented strategies to pinpoint reliable IR events	503
3. Quantifying IR levels	505
3.1. Percentage spliced-in	505
3.2. IR ratio	505
4. Cross-sample comparison	506
4.1. Intron-bin count-based methods	506
4.2. Average intron coverage method	506
5. Discussion	507
Declaration of Competing Interest	507
Acknowledgements	507
References	507

Abbreviations: IR, Intron retention; RNA-seq, RNA sequencing; AS, alternative splicing.

* Corresponding author.

E-mail addresses: lucile.broseus@igh.cnrs.fr (L. Broseus), william.ritchie@igh.cnrs.fr (W. Ritchie).

1. Introduction

Amongst the three major types of alternative splicing (AS) that include exon skipping/inclusion, alternative 5' and 3' splice-site selection and intron retention (IR), the latter has until recently been regarded as an oddity in mammals; IR was often added to the list for no other intent than to be exhaustive. However, recent discoveries about the role IR can play in fine-tuning gene expression [1–3] as well as the observation of characteristic IR patterns [4–9] highlight the value of investigating IR in transcriptomic studies. Numerous reports have demonstrated a regulatory role for IR in hematopoiesis [1,2], neuronal differentiation [10], germ cell differentiation [11] and CD4+ T cell activation [12] amongst others. In addition, a recent analysis of 1812 cancer patient samples showed that over 18% of splicing-associated single nucleotide variants caused IR and most of these events affected tumor suppressor genes [13]. Finally, the analysis of 2573 samples showed that IR occurs in all tissues analyzed and can affect over 80% of all coding genes [4]. Measuring IR can help decipher gene-level variations and inter-connections between transcriptional loads, structural variations and phenotypes [14,15].

IR has been demonstrated to downregulate gene expression in numerous systems by triggering the nonsense mediated mRNA decay (NMD) pathway. NMD recognizes transcripts with premature stop codons (PTC) that could potentially generate C-terminal truncated proteins and degrades them. This surveillance mechanism can thus rapidly degrade IR transcripts if they harbor a PTC. Given that introns are much longer than exons and under less selective pressure to conserve open reading frames, the probability that an IR event harbors a PTC is high and IR transcripts are thus good candidates for degradation via NMD. Initially, transcriptomic and bioinformatic analyses of NMD concluded that it was not coupled with mRNA splicing and that most PTC containing transcripts do not have major functional roles [16], thus relegating NMD to the role of scavenger [17]. The same team however recently revised their view on the functional importance of NMD [3] and the current consensus is that it couples with IR (and other forms of AS) to regulate gene expression in numerous systems [1,3,18–20]. The importance of NMD is further underscored by the fact that deletion of its core components result in embryonic lethality [21].

RNA-seq data is well suited for resolving local exon connectivity because sequencing reads are sufficiently long to cover exon-exon junctions. It is also well suited for measuring global gene expression because the high number of reads that map to genes generally enables the use of powerful statistical models. Detecting and measuring IR with RNA-seq is more complex. The technical biases that are known to distort gene expression levels (eg: GC content, amplification biases) and other types of splicing, also affect IR measurement. In addition, measurement of intronic expression is challenged by numerous factors. Within introns, highly expressed features such as small nucleolar RNAs, microRNAs or unannotated exons may erroneously inflate count-based measures of intronic expression. Conversely, low complexity regions, common in introns, prevent unique mapping of reads. Because retained introns are generally expressed at a fraction of their flanking exons, uncorrected biases can massively disrupt IR estimation.

Transcriptome-wide evaluation of IR by computational means is still a budding field of investigation. In many studies, IR was assessed via custom and briefly detailed procedures, which is probably due, in part, to the fact that few dedicated and comprehensive tools have been published so far. In the following, we give a survey of the technical biases that confound IR detection and available computational methods to tackle IR screening. We emphasize three crucial steps which are: the preparation and quality control of the sequencing data and the reference transcriptome; generat-

ing metrics that reflect the biological signal of IR transcripts and using a model to discover condition-specific IR events.

2. IR detection

2.1. Filtering sequencing data

Unlike sequencing reads that map to exon-exon junctions, reads that map to introns can originate from DNA contamination caused by ineffective DNase treatment or from pre-mature mRNA. One method to detect DNA contamination is to measure reads across splice sites and check that the majority of introns display high splicing efficiency (above 90%) [22]. Another approach, implemented in [4] is to verify that the ratio of the number of reads that map to intergenic regions to the number that maps to coding regions is less than 10%. Another source of bias is that intronic reads may originate from nascent and pre-mature RNAs [23]. So as to lessen signals due to unprocessed transcripts and overlapping antisense transcripts, it is recommended to use Poly-A enriched RNA-seq (or cytoplasmic fractionation) and strand-specific protocols [4].

Regarding library size, IR occurs at relatively low frequency in mammals, and introns tend to be substantially longer than exons. Determining an optimal library size obviously depends on many experiment-specific factors and on the IR effect size considered as biologically significant, 35 millions mapped reads for a one-versus-one experiment was suggested as an optimum for detection of differential intron usage (based on a resampling approach, [24]). In order to bypass intronic alignment biases, it has been suggested to consider only splice junction reads or similarly to focus only on a window centered on splice sites. Nonetheless, those junction-only analyses are likely to be more affected by splicing variations in flanking exons and lead to even more unstable estimates. Accordingly, previous studies pointed out that they require higher sequencing-depth (at least 70 million reads per sample, ideally more than 150 million reads) [23].

2.2. Defining reference intronic sequences

Sequencing reads that map to intronic intervals may originate from several different sources such as overlapping genes. These confound measurements of the magnitude of true IR. It is therefore crucial to correctly define the intronic intervals that will be used to measure IR. Here, two main approaches have been adopted (cf: Fig. 1), each calling for precautions for interpretation and specific processing to avoid false positive detection.

2.2.1. The all-introns view

A first possibility is to analyze all intronic intervals present in at least one annotated transcript model [4]. Although this allows to screen the largest set of candidates, this comes at the expense of having to deal with peaks of intronic alignments caused by expressed alternative exons and redundant IR calls due to overlapping introns.

2.2.2. The measurable introns view

Measurable (or independent) *introns* are (parts of) introns that do not overlap with any annotated exon [24,26,27,30]. They are obtained by subtracting merged exons from genes. This comes with the advantage to simplify the analysis of introns flanked by exons with known alternative donor sites, but ignores introns fully overlapped by annotated exons.

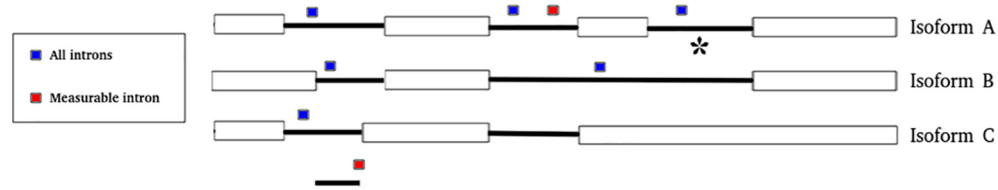


Fig. 1. Defining intronic intervals to be analyzed. Comprehensiveness of transcript annotation and the selection of reference intronic sequences have a major impact on IR detection. In the example, we consider a gene having three possible isoforms (A, B and C). Exons are represented as plain rectangles and introns as thick black lines. If only Isoforms B and C were annotated, the starred interval (*) would not be defined as an intron and most likely not detected as retained. Colored boxes indicate whether the annotated introns match the “all introns” or “independent/measurable intron” criteria used by current algorithms.

2.3. Overcoming sequencing and alignment artifacts

Even when appropriate sequencing protocols have been used, several sources of confusion remain that can only be overcome with computational means. These are detailed below and in Fig. 2.

2.3.1. Overlapping features

First, the sequencing protocol may capture molecules overlapping or mapping within introns, such as small nucleolar RNAs, micro RNAs, unannotated exons or alternative 5' and 3' splice-sites. They form characteristic peaks of reads within intronic regions that may induce false IR detections and result in inaccurate quantification if not properly identified and filtered.

2.3.2. Repeated regions

Relatively to exons, introns are long and poorly conserved, often contain low mappability regions (eg: duplicated regions like transposons [31–33] or repeated regions such as microsatellites [34] which impair correct IR level estimations.

2.3.3. Low coverage of flanking exons

Thirdly, sequencing and alignment artifacts may occur in flanking exons. For example, GC rich exons are under-covered and very

small exons are more difficult to map. This may perturb or inflate IR measures, especially those that only measure reads directly surrounding the intron. They may also lead to missed IR events as the junctions are weakly supported.

2.3.4. 3' Coverage bias

PolyA-enriched RNA-seq data usually display a marked 3' coverage bias, so that most 3' introns are likely to be more covered and thus more easily captured than 5' ones. This should be kept in mind for any inference regarding any positional bias of intron retention.

In practice, the prevailing strategy for classifying introns relies essentially on user-defined thresholds. Here, we summarize the different strategies used by IR detection tools, and indicate the parameter values suggested by their respective authors when they exist (Fig. 3 and Table 1).

2.4. Implemented strategies to pinpoint reliable IR events

Although numerous computational methods have been developed to estimate splicing efficiency and to model sequencing errors that may affect their estimation, we have decided to list here those approaches that specifically cater to the difficulties of detecting IR.

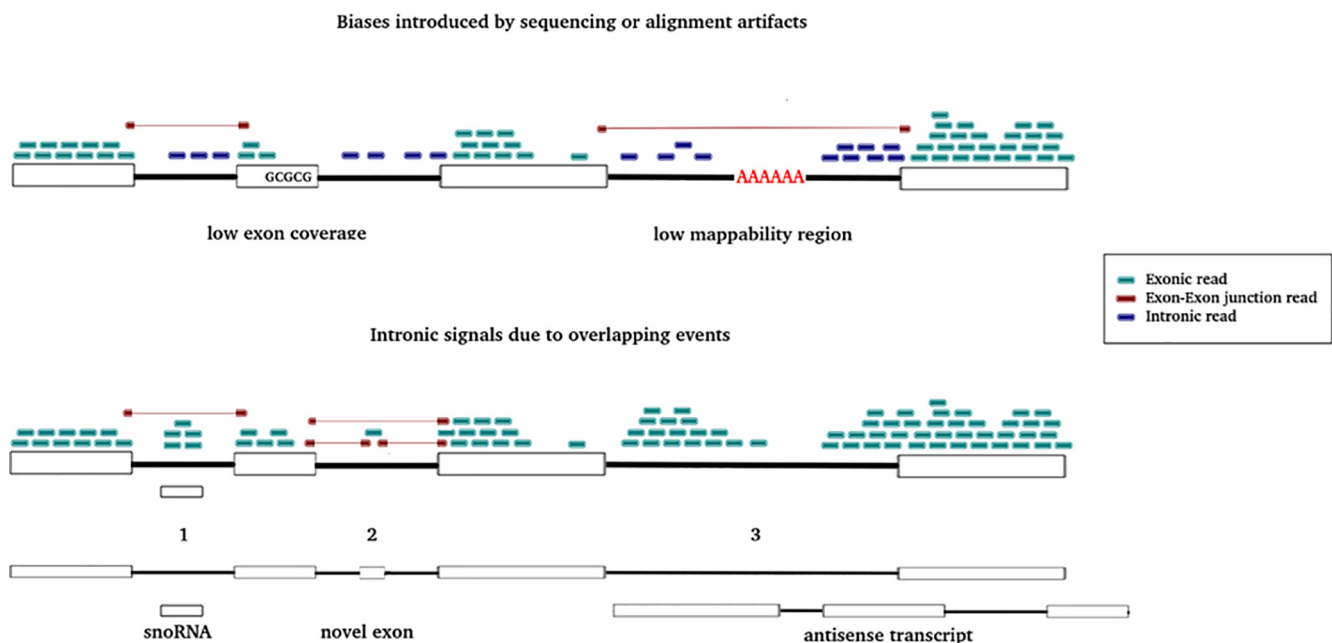


Fig. 2. Potential sources of bias and confusion: a very unfortunate gene. Only intron 3 is retained in this example. In intron 1: expression of an overlapping feature causes a peak in alignments, which can artificially inflate the estimation of IR. Intronic alignments in intron 2 originate from an unannotated exon. Intron 3 is retained but its detection is hampered by multiple biases. First, the presence of a low mappability region (repeated A sequence in red) would result either in a gap or in high uncertainty in read alignments in that region. Secondly, high GC content in the 5' exon explains the lack of exon-exon junctions and 3' exon-intron reads and may affect filtering and IR metrics based on them. Thirdly, due to its long length, it tends to be more sparsely covered. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

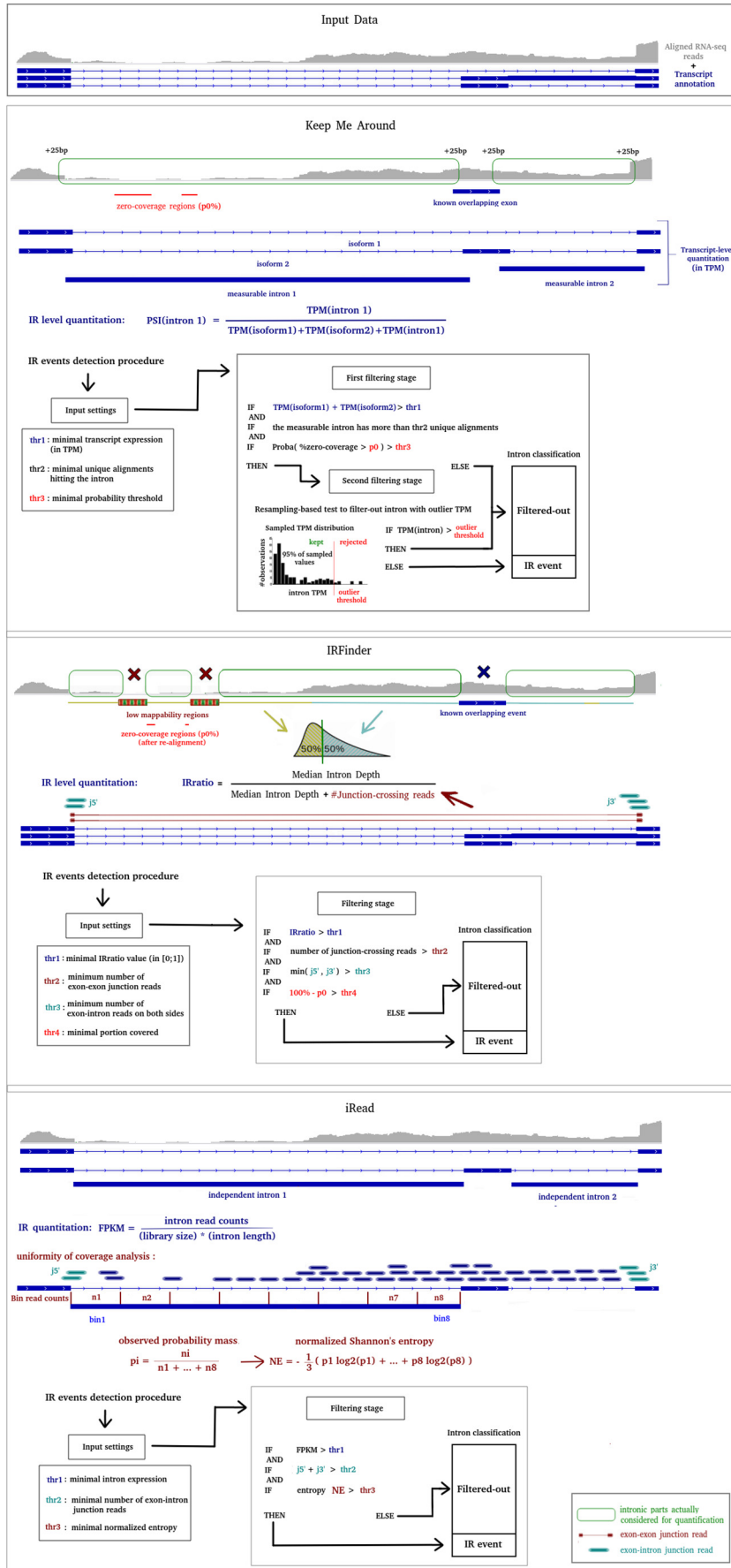


Fig. 3. Standard implementation of computational detection of IR events.

Table 1
Computational tools available to perform IR detection and their main features.

	Year	Publication	Language	Intron definition	IR measure	Low mappability correction	Unknown overlapping events detection
MISO [25]	2010	Nature Methods	Python	Independent introns	PSI	No	No
KMA [26]	2015	arXiv	Python and R	Measurable introns	PSI	No	Coverage analysis (Probabilistic test)
iRead [27]	2017	bioRxiv	Python	Independent introns	FPKM	No	Coverage Analysis (Shannon entropy)
IRFinder [4]	2017	Genome Biology	C++	All Introns	IRratio	Yes	Coverage Analysis (Detection of outlier regions)
IntEREest [24]	2018	BMC Bioinformatics	R	Independent introns	PSI or FPKM	Optional	No
ASDT [28]	2018	ATM	Perl	No (Reference-free)	No	No	Yes
JUM [29]	2018	PNAS	Perl	No (Reference-free)	No	No	Yes

Keep Me Around (KMA) [26] uses the measurable introns approach. Transcripts are quantified using eXpress [35] or Kallisto [36,37] and a PSI value is computed to evaluate IR levels (**cf: Quantifying IR levels**).

Spurious intronic signals are spotted by finding the longest alignment gap in an intron and calculating the probability of observing such a long gap given the intron's expression and if the distribution were uniform.

The authors selected introns with at least three uniquely mapped reads, cumulated TPM values for non-IR transcripts greater than 1 and introns with zero-coverage regions longer than 20% of the intron length [38].

IRFinder [4] makes the choice to screen all introns derived from the annotation. Introns that overlap with any known exon or RNA molecule are marked in the output.

A procedure is implemented to identify low mappability regions and exclude them and their reads from the subsequent calculation. Potential other artifacts are handled by discarding bases with outlier read depth value compared to the average intron depth.

The IRratio and several other complementary metrics are then computed to evaluate support for IR.

Suggested parameters values for IRFinder are: IRratio > 0.1, and at least 3 reads supporting intron exclusion on both sides.

In **iRead** [27], reference introns are provided by the user.

Suspect cases are likened to non-uniform coverage. Read coverage uniformity is quantified by the Shannon's entropy of read distribution along the intron and low entropy is associated to a non-uniform read coverage in the intron.

By default, iRead will select intron having FPKM > 3, at least one exon-intron junction read and the normalized entropy-score > 0.9.

IntEREest [24] computes FPKM and PSI values for independent introns. Optionally, low mappability regions can be excluded from the calculations. No specific guidelines are provided to select IR events, but data are formatted for some methods for performing differential analysis.

It is worth emphasizing that each approach makes use of pre-fixed threshold values for all intronic regions. Most of these values are defined according to the coverage profile expected for (well-behaved) short length medium-coverage introns, and are maybe the most straightforward way to guard from the bulk of artefactual detections. However, introns form a highly heterogeneous set of regions, hugely differing in length, inner and flanking coverage and sequence feature. For example, on genes having sparse coverage, chances to observe counts on a very small pre-specified area of the genome are quite low. Therefore, filters on the number of junction reads are likely to exclude most of their introns from further analysis. Very long introns are especially problematic and in practice these regions have little chance to be covered at their full extent, and well generally fail on the hard cutoffs set by these algorithms. It is thus clear that no universal threshold can be convenient in all cases, and that any rigid thresholding is likely to introduce a severe selection bias. We thus argue that a sensible

choice for the various parameters must be intron-specific and encourage the development and use of models that account explicitly for sequence features and coverage variations.

3. Quantifying IR levels

Though essential, devising a computable and robust metric that reflects "splicing efficiency" or oppositely the level of IR can be difficult. Three types of alignments should be taken into account [18]. Intronic reads and reads that span the flanking exons are informative of the level of IR. In addition, all the remaining alignments, can indicate how reliable the sequencing data is and what degree of confidence we may have in each IR event. The most commonly used ratios to quantify IR are the percentage spliced in and the IR-ratio, both described below.

3.1. Percentage spliced-in

Alternative splicing event frequencies are commonly quantified by the percentage spliced-in (PSI) ratio [29]. An intronic version has been suggested [19] as the number of reads supporting the retention of the intron against the number of reads supporting its exclusion. In practice, a transcript-level quantification is performed using an annotation of IR-free isoforms augmented with independent introns (taken as dummy transcripts). The PSI for a given intron can be formulated as:

$$PSI = \frac{IntronTPM}{IntronTPM + \sum transcriptsTPM}$$

where the sum is performed across all annotated transcripts of the same gene not retaining an intron.

3.2. IR ratio

This metric is to reflect splicing efficiency as the portion of informative reads which come from a transcript retaining the intron, that is:

$$IRratio = \frac{IntronicAbundance}{IntronicAbundance + NormalSplicingAbundance}$$

where *Intronic abundance* is measured by the median [4] or average [39] intron depth. The *abundance of normal splicing* is taken as the number of reads spliced across the intron.

These ratios tend to show high fluctuations and their behavior is difficult to model. This may explain why, so far, no approach has been developed to estimate dispersions and confidence intervals. Importantly, this hinders the identification of robust and reproducible patterns based on their observed values. Although these metrics can be employed, as a proxy for splicing efficiency, to call manifest IR events, additional statistics are required to infer intra- and cross-sample variation levels.

Table 2
Available computational methods to perform IR differential analysis.

Method	Year	Language	IR-specific	IR measure	Normalization For library size	Control for gene expression	Modeling of biological variability	Statistical Framework
edgeR-IR*	2010	R	No/Yes**	Intron bin count	TMM (ref)	No/Yes**	Yes	Generalized Linear Model
DESeq2-IR*	2014	R	No/Yes**	Intron bin count	Variance estimation and rescaling (ref)	No/Yes**	Yes	Generalized Linear Model
DEXSeq-IR*	2012	R	No/Yes**	Intron bin count	Variance estimation and rescaling (ref)	Yes	Yes	Generalized Linear Model
iDiffIR	2018	Python	Yes	Average per base read coverage	TMM (ref)	Yes	Yes	LogFC statistic and Z-test

* These refer to IR-tuned versions of existing software, and may require custom pre-processing.

** After IR-specific tuning.

One of the major difficulties for quantifying splicing efficiency is due to the fact that the exons flanking an intron may connect not only to each other but with other exons from the same gene to form different isoforms. This hampers the estimation of the portion of reads to attribute to the transcripts in which the intron is spliced. The two measures presented above (PSI and IRratio) address this problem differently. So as to overcome global variations in gene coverage caused by alternative exon usage, the strategy behind the IRratio is to only make use of the junction-crossing reads that hit any one of the two exons flanking the intron. The maximum value between the left and right quantities is then taken as a means to mitigate against the existence of multiple isoforms that connect to the flanking exons. However, the number of junction reads tend to be highly dispersed with high coverage, and to take zero values when the coverage is low. This may incidentally affect estimation accuracy. On the other hand, by using information across the whole transcript to evaluate the gene coverage, the PSI estimator might be more resistant to these local variations. Nonetheless, it would be of interest to assess the quality of the PSI estimates on genes which undergo manifold alternative splicing events [40,41].

4. Cross-sample comparison

Inferring differences in IR between conditions necessitates a statistical framework to combine biological replicates, assess dispersion of IR level estimates and control the false positive rate. Moreover, sample read abundances need to be normalized to account for variations in library size [42]. Moreover, the coverage depth of an intron is correlated to its gene coverage; sample comparison thus necessitates strategies to control for differences in gene expression [23].

To our knowledge, currently four implemented frameworks fulfill these requirements (cf: Table 2). In regards of their statistical methodology, we split them into three families of approaches.

4.1. Intron-bin count-based methods

The first two approaches re-use existing methods that were primarily devised either for gene expression [43,44] or exon usage [30,45] analyses, after some reworking of the data to adapt them to IR. They are count-based. To adjust for differences in library size, a gene-wise normalization factor is determined and applied to all (exon and intron) bins, as in usual gene expression differential analyses.

The authors of ASpli¹ suggest to adjust each intron bin counts B_i {i}, in each sample s , by biological condition through:

$$\frac{B_i}{G_{Condition}} * \bar{G}$$

where $G_{\{Condition\}}$ is the average gene count in condition and \bar{G} the average gene count across all samples. Classical testing procedures (either of edgeR or DESeq2) are then applied to these adjusted counts to infer a set of differential introns.

In the DEXSeq-IR method, for each intron, two count bins are considered: the intron bin and the union of all the remaining bins. The average bin count is modeled via a negative binomial generalized linear model with interaction term.

In more details, for an intron indexed by i in a sample j :

$$\log_2(\text{mean}_{ij}^{\text{Intron}}) = \beta_{ij}^{\text{Sample}} + l\beta_i^{\text{Intron}} + \beta_{ic}^{\text{Intron} \times \text{Condition}}$$

where $l = 1$ for the intron bin and 0 otherwise.

The sample parameter $\beta_{\{Sample\}}$ adjusts internally for the gene expression level and differences in intron usage between conditions is inferred by testing whether the interaction term is significantly different from zero².

As described previously, so as to limit spurious noise in intronic read counts, it may be worth refining further intron bin counts by removing reads that map to artifact-prone intervals. Several detection softwares already output these corrected read counts [4,24].

4.2. Average intron coverage method

The third and most recent approach, iDiffIR (implemented but not yet published, https://bitbucket.org/comp_bio/iddiffir, [39,46]) is primarily designed for IR.

IR levels are quantified, from genomic unique alignments, by the average per base read coverage (over the intron interval). To account for library size, TMM normalization [47] is applied, separately, for intron and exon per base read counts. Per base counts are further normalized to force overall gene coverage to be equal across conditions.

The test statistic is a corrected log fold change between the average read coverage in each of the conditions compared:

$$\log_{FC}(I) = \log_2\left(\frac{a + \text{mean}(I1)}{a + \text{mean}(I2)}\right)$$

The correction parameter a is a pseudo-count whose value is chosen to minimize the log fold-change and control large values caused by lowly covered introns.

The biological intuition of *splicing efficiency* translates quantitatively as the proportion of emitted transcripts which retain an intron. However, owing to short read size and rarity of IR events, this frequency cannot be reliably estimated from RNA-seq data.

¹ <https://www.bioconductor.org/packages/release/bioc/html/ASpli.html>.

² <https://www.bioconductor.org/packages/release/bioc/vignettes/DEXSeq/inst/doc/DEXSeq.html>.

Metrics that can actually be computed (eg: PSI, IRratio) would only be proxies, and their properties and meaning are still poorly understood. Importantly, it is not sure whether intronic expression can be compared based on these metrics. All the problems discussed previously in the measurement of IR levels will have direct repercussions on the detection of alternate IR levels between samples. Ironically, despite numerous efforts to quantify IR events, the cases best suited to most of these approaches remains introns with high coverage and short length, exon-like introns.

5. Discussion

The recent interest in expressed introns has led to a flourishing number of examples of regulation through intron retention. The accurate detection of retained introns and precise measurement of intronic expression are crucial to these studies. Numerous factors impede the detection of IR from next generation sequencing data. Introns are much longer than exons and thus have a much higher probability of containing overlapping features that may confound the estimation of intronic expression. In addition, introns are enriched in low complexity and repeat sequences that may prevent sequencing data from being uniquely mapped. These factors must be accounted for when detecting IR events. Most computational approaches however will introduce a selection bias as only introns with sufficient coverage can be detected and the statistical power required to detect differences between conditions increases with coverage depth and the read count [48,49]. As a result of this bias, gene enrichment tests of genes derived from IR signatures [50] are heavily skewed towards the more expressed genes and towards introns that do not contain these confounding features.

Despite the recent results that demonstrate a crucial role for IR, very few IR events have been validated in the wetlab and amongst these an even smaller portion have been investigated for their functional impact. As a consequence, no reliable benchmark of IR detection or differential intronic expression has been published. This lack of reliable controls is however temporary because long read technologies capable of sequencing entire IR transcripts help resolve most of the detection problems. However, due to their low coverage, these technologies are far from allowing a comprehensive detection of IR events and even further from allowing a reliable quantification of IR levels between different tissues.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Agence Nationale de la Recherche [ANRJCJC - WIRED], the Labex EpiGenMed [ANR-10-LABX-12-01] and the MUSE initiative [GECKO].

References

- [1] Wong JJ-L, Ritchie W, Ebner OA, Selbach M, Wong JWH, Huang Y, et al. Orchestrated intron retention regulates normal granulocyte differentiation. *Cell* 2013;154:583–95. <https://doi.org/10.1016/j.cell.2013.06.052>.
- [2] Edwards CR, Ritchie W, Wong JJ-L, Schmitz U, Middleton R, An X, et al. A dynamic intron retention program in the mammalian megakaryocyte and erythrocyte lineages. *Blood* 2016;127:e24–34. <https://doi.org/10.1182/blood-2016-01-692764>.
- [3] Braunschweig U, Barbosa-Morais NL, Pan Q, Nachman EN, Alipanahi B, Gonatopoulos-Pournatzis T, et al. Widespread intron retention in mammals functionally tunes transcripts. *Genome Res* 2014;24:1774–86. <https://doi.org/10.1101/gr.177790.114>.

- [4] Middleton R, Gao D, Thomas A, Singh B, Au A, Wong JJ-L, et al. IRFinder: assessing the impact of intron retention on mammalian gene expression. *Genome Biol* 2017;18:51. <https://doi.org/10.1186/s13059-017-1184-4>.
- [5] Giannopoulou AF, Konstantakou EG, Velentzas AD, Avgeris SN, Avgeris M, Papandreou NC, et al. Gene-specific intron retention serves as molecular signature that distinguishes melanoma from non-melanoma cancer cells in greek patients. *Int J Mol Sci* 2019;20:937. <https://doi.org/10.3390/ijms20040937>.
- [6] Braun CJ, Stanciu M, Boutz PL, Patterson JC, Calligaris D, Higuchi F, et al. Coordinated splicing of regulatory retained introns within oncogenic transcripts creates an exploitable vulnerability in malignant glioma. *Cancer Cell* 2017;32(411–426):. <https://doi.org/10.1016/j.ccell.2017.08.018e11>.
- [7] Jangi M, Fleet C, Cullen P, Gupta SV, Mekhoubad S, Chiao E, et al. SMN deficiency in severe models of spinal muscular atrophy causes widespread intron retention and DNA damage. *PNAS* 2017;114:E2347–56. <https://doi.org/10.1073/pnas.1613181114>.
- [8] Harahap NIF, Niba ETE, Rochmah MA, Wijaya YOS, Saito T, Saito K, et al. Intron-retained transcripts of the spinal muscular atrophy genes, SMN1 and SMN2. *Brain Develop* 2018;40:670–7. <https://doi.org/10.1016/j.braindev.2018.03.001>.
- [9] Adusumalli S, Ngian Z, Lin W, Benoukraf T, Ong C. Increased intron retention is a post-transcriptional signature associated with progressive aging and Alzheimer's disease. *Aging Cell* 2019;18. <https://doi.org/10.1111/acel.12928>.
- [10] Mauger O, Lemoine F, Scheiffele P. Targeted intron retention and excision for rapid gene regulation in response to neuronal activity. *Neuron* 2016;92:1266–78. <https://doi.org/10.1016/j.neuron.2016.11.032>.
- [11] Naro C, Pellegrini L, Jolly A, Farini D, Cesari E, Bielli P, et al. Functional interaction between U1snRNP and Sam68 insures proper 3' end Pre-mRNA processing during germ cell differentiation. *Cell Reports* 2019;26(2929–2941):. <https://doi.org/10.1016/j.celrep.2019.02.058e5>.
- [12] Ni T, Yang W, Han M, Zhang Y, Shen T, Nie H, et al. Global intron retention mediated gene regulation during CD4+ T cell activation. *Nucleic Acids Res* 2016;44:6817–29. <https://doi.org/10.1093/nar/gkw591>.
- [13] Boutz PL, Bhutkar A, Sharp PA. Retained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev* 2015;29:63–80. <https://doi.org/10.1101/gad.247361.114>.
- [14] Odhams CA, Graham DSC, Vyse TJ. Profiling RNA-Seq at multiple resolutions markedly increases the number of causal eQTLs in autoimmune disease. *PLoS Genet* 2017;13:. <https://doi.org/10.1371/journal.pgen.1007071>.
- [15] A data-driven approach to characterising intron signal in RNA-seq data | bioRxiv n.d. <https://www.biorxiv.org/content/10.1101/352823v1.article-info> (accessed November 29, 2019).
- [16] Pan Q, Saltzman AL, Kim YK, Misquitta C, Shai O, Maquat LE, et al. Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. *Genes Dev* 2006;20:153–8. <https://doi.org/10.1101/gad.1382806>.
- [17] Ge Y, Porse BT. The functional consequences of intron retention: alternative splicing coupled to NMD as a regulator of gene expression. *BioEssays* 2014;36:236–43. <https://doi.org/10.1002/bies.201300156>.
- [18] Jacob AG, Smith CWJ. Intron retention as a component of regulated gene expression programs. *Hum Genet* 2017;136:1043–57. <https://doi.org/10.1007/s00439-017-1791-x>.
- [19] Hussein SMI, Puri MC, Tonge PD, Benevento M, Corso AJ, Clancy JL, et al. Genome-wide characterization of the routes to pluripotency. *Nature* 2014;516:198–206. <https://doi.org/10.1038/nature14046>.
- [20] Wong JJ-L, Au AYM, Ritchie W, Rasko JEJ. Intron retention in mRNA: no longer nonsense. *BioEssays* 2016;38:41–9. <https://doi.org/10.1002/bies.201500117>.
- [21] Medghalchi SM, Frischmeyer PA, Mendell JT, Kelly AG, Lawler AM, Dietz HC. Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum Mol Genet* 2001;10:99–105. <https://doi.org/10.1093/hmg/10.2.99>.
- [22] Khodor YL, Rodriguez J, Abruzzi KC, Tang C-HA, Marr MT, Rosbash M. Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in *Drosophila*. *Genes Dev* 2011;25:2502–12. <https://doi.org/10.1101/gad.178962.111>.
- [23] Vanichkina DP, Schmitz U, Wong JJ-L, Rasko JEJ. Challenges in defining the role of intron retention in normal biology and disease. *Semin Cell Dev Biol* 2018;75:40–9. <https://doi.org/10.1016/j.semcdb.2017.07.030>.
- [24] Oghabian A, Greco D, Frilander MJ. IntERESt: intron-exon retention estimator. *BMC Bioinf* 2018;19:130. <https://doi.org/10.1186/s12859-018-2122-5>.
- [25] Katz Y, Wang ET, Airoldi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods* 2010;7:1009–15. <https://doi.org/10.1038/nmeth.1528>.
- [26] [1510.00696] Keep Me Around: Intron Retention Detection and Analysis n.d. <https://arxiv.org/abs/1510.00696> (accessed November 29, 2019).
- [27] iREAD: A Tool for Intron Retention Detection from RNA-seq Data | bioRxiv n.d. <https://www.biorxiv.org/content/10.1101/135624v2> (accessed November 29, 2019).
- [28] Adamopoulos PG, Theodoropoulou MC, Scorialas A. Alternative splicing detection tool—a novel PERL algorithm for sensitive detection of splicing events, based on next-generation sequencing data analysis. *Ann Transl Med* 2018;6.
- [29] Wang Q, Rio DC. JUM is a computational method for comprehensive annotation-free analysis of alternative pre-mRNA splicing patterns. *PNAS* 2018;115:E8181–90. <https://doi.org/10.1073/pnas.1806018115>.
- [30] Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. *Genome Res* 2012;22:2008–17. <https://doi.org/10.1101/gr.133744.111>.

- [31] Li Y, Bor Y, Misawa Y, Xue Y, Rekosh D, Hammarskjöld M-L. An intron with a constitutive transport element is retained in a Tap messenger RNA. *Nature* 2006;443:234–7. <https://doi.org/10.1038/nature05107>.
- [32] Rekosh D, Hammarskjöld M-L. Intron retention in viruses and cellular genes: detention, border controls and passports. *Wiley Interdiscip Rev RNA* 2018;9: <https://doi.org/10.1002/wrna.1470>.
- [33] Li Y, Bor Y, Fitzgerald MP, Lee KS, Rekosh D, Hammarskjöld M-L. An NXF1 mRNA with a retained intron is expressed in hippocampal and neocortical neurons and is translated into a protein that functions as an Nxf1 cofactor. *MBoC* 2016;27:3903–12. <https://doi.org/10.1091/mbc.F16-07-0515>.
- [34] Sznajder Ł, Thomas JD, Carrell EM, Reid T, McFarland KN, Cleary JD, et al. Intron retention induced by microsatellite expansions as a disease biomarker. *PNAS* 2018;115:4234–9. <https://doi.org/10.1073/pnas.1716617115>.
- [35] Roberts A, Pachter L. Streaming fragment assignment for real-time analysis of sequencing experiments. *Nat Methods* 2013;10:71–3. <https://doi.org/10.1038/nmeth.2251>.
- [36] Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016;34:525–7. <https://doi.org/10.1038/nbt.3519>.
- [37] Smart AC, Margolis CA, Pimentel H, He MX, Miao D, Adeegbe D, et al. Intron retention is a source of neoepitopes in cancer. *Nat Biotechnol* 2018;36:1056–8. <https://doi.org/10.1038/nbt.4239>.
- [38] Sakharkar MK, Chow VTK, Kanguane P. Distributions of exons and introns in the human genome. *Silico Biology* 2004;4:387–93.
- [39] Filichkin SA, Hamilton M, Dharmawardhana PD, Singh SK, Sullivan C, Ben-Hur A, et al. Abiotic stresses modulate landscape of poplar transcriptome via alternative splicing, differential intron retention, and isoform ratio switching. *Front Plant Sci* 2018;9. <https://doi.org/10.3389/fpls.2018.00005>.
- [40] Teng M, Love MI, Davis CA, Djebali S, Dobin A, Graveley BR, et al. A benchmark for RNA-seq quantification pipelines. *Genome Biol* 2016;17:74. <https://doi.org/10.1186/s13059-016-0940-1>.
- [41] Angelini C, Canditiis DD, Feis ID. Computational approaches for isoform detection and estimation: good and bad news. *BMC Bioinf* 2014;15:135. <https://doi.org/10.1186/1471-2105-15-135>.
- [42] Dillies M-A, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, et al. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform* 2013;14:671–83. <https://doi.org/10.1093/bib/bbs046>.
- [43] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15. <https://doi.org/10.1186/s13059-014-0550-8>.
- [44] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
- [45] Li Y, Rao X, Mattox WW, Amos CI, Liu B. RNA-Seq analysis of differential splice junction usage and intron retentions by DEXSeq. *PLoS ONE* 2015;10: <https://doi.org/10.1371/journal.pone.0136653>.
- [46] Collins JE, White RJ, Staudt N, Sealy IM, Packham I, Wali N, et al. Common and distinct transcriptional signatures of mammalian embryonic lethality. *Nat Commun* 2019;10:1–16. <https://doi.org/10.1038/s41467-019-10642-x>.
- [47] Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 2010;11:R25. <https://doi.org/10.1186/gb-2010-11-3-r25>.
- [48] Oshlack A, Wakefield MJ. Transcript length bias in RNA-seq data confounds systems biology. *Biology Direct* 2009;4:14. <https://doi.org/10.1186/1745-6150-4-14>.
- [49] Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* 2010;11:R14. <https://doi.org/10.1186/gb-2010-11-2-r14>.
- [50] Timmons JA, Szkop KJ, Gallagher IJ. Multiple sources of bias confound functional enrichment analysis of global-omics data. *Genome Biol* 2015;16:186. <https://doi.org/10.1186/s13059-015-0761-7>.