



# Testing the Genomic Shock Hypothesis Using Transposable Element Expression in Yeast Hybrids

Marika Drouin<sup>1,2,3,4\*</sup>, Mathieu Hénault<sup>1,2,3,4</sup>, Johan Hallin<sup>1,2,3,4,5</sup> and Christian R. Landry<sup>1,2,3,4,5\*</sup>

<sup>1</sup> Institut de Biologie Intégrative et des Systèmes, Université Laval, Québec, QC, Canada, <sup>2</sup> Département de Biochimie, de Microbiologie et de Bio-Informatique, Université Laval, Québec, QC, Canada, <sup>3</sup> PROTEO - Regroupement Québécois de Recherche sur la Fonction, l'Ingénierie et les Applications des Protéines, Québec, QC, Canada, <sup>4</sup> Centre de Recherche en Données Massives de l'Université Laval, Université Laval, Québec, QC, Canada, <sup>5</sup> Département de Biologie, Université Laval, Québec, QC, Canada

## OPEN ACCESS

### Edited by:

Toni Gabaldón,  
Barcelona Supercomputing  
Center, Spain

### Reviewed by:

Fanny Hartmann,  
Universite Paris-Saclay, France  
Teeratas Kijpornyongpan,  
National Renewable Energy  
Laboratory (DOE), United States

### \*Correspondence:

Marika Drouin  
marika.drouin.1@ulaval.ca  
Christian R. Landry  
christian.landry@bio.ulaval.ca

### Specialty section:

This article was submitted to  
Fungal Genomics and Evolution,  
a section of the journal  
Frontiers in Fungal Biology

**Received:** 22 June 2021

**Accepted:** 26 July 2021

**Published:** 23 August 2021

### Citation:

Drouin M, Hénault M, Hallin J and  
Landry CR (2021) Testing the  
Genomic Shock Hypothesis Using  
Transposable Element Expression in  
Yeast Hybrids.  
Front. Fungal Biol. 2:729264.  
doi: 10.3389/ffunb.2021.729264

Transposable element (TE) insertions are a source of structural variation and can cause genetic instability and gene expression changes. A host can limit the spread of TEs with various repression mechanisms. Many examples of plant and animal interspecific hybrids show disrupted TE repression leading to TE propagation. Recent studies in yeast did not find any increase in transposition rate in hybrids. However, this does not rule out the possibility that the transcriptional or translational activity of TEs increases following hybridization because of a disruption of the host TE control mechanisms. Thus, whether total expression of a TE family is higher in hybrids than in their parental species remains to be examined. We leveraged publically available RNA-seq and ribosomal profiling data on yeast artificial hybrids of the *Saccharomyces* genus and performed differential expression analysis of their LTR retrotransposons (Ty elements). Our analyses of total mRNA levels show that Ty elements are generally not differentially expressed in hybrids, even when the hybrids are exposed to a low temperature stress condition. Overall, only 2/26 Ty families show significantly higher expression in the *S. cerevisiae* × *S. uvarum* hybrids while there are 3/26 showing significantly lower expression in the *S. cerevisiae* × *S. paradoxus* hybrids. Our analysis of ribosome profiling data of *S. cerevisiae* × *S. paradoxus* hybrids shows similar translation efficiency of Ty in both parents and hybrids, except for Ty1<sub>cer</sub> showing higher translation efficiency. Overall, our results do not support the hypothesis that hybridization could act as a systematic trigger of TE expression in yeast and suggest that the impact of hybridization on TE activity is strain and TE specific.

**Keywords:** transposable element, hybridization, genomic shock, differential expression analysis, RNA sequencing, retrotransposon, yeast

## INTRODUCTION

Genome diversification resulting from interspecific hybridization allows the development of novel phenotypes, creating an adaptive potential that can empower hybrids to colonize new niches (Runemark et al., 2019; Steensels et al., 2021). However, combining divergent genomes can also lead to hybrid sterility or inviability caused by genetic incompatibilities (Muller, 1939; Maheshwari and Barbash, 2011). Other consequences of hybridization include genetic instability that can affect

hybrid fitness and hybrid genome evolution. One example is an increased rate of proliferation of transposable elements (TEs) in hybrid genomes (McClintock, 1984).

TEs are dispersed repeated sequences that can propagate within a genome (Bourque et al., 2018). They are ubiquitous in eukaryotes but represent various genome proportions in different species. For instance, they represent from 3% of the yeast *Saccharomyces cerevisiae* genome and up to 80% of some plant genomes (Meyers et al., 2001; Carr et al., 2012; Platt et al., 2018). TE abundance explains much of the variation in genome size among eukaryotes (Elliott and Gregory, 2015). TEs generate genomic insertions that are mostly selectively neutral or slightly deleterious (Doolittle and Sapienza, 1980). Because of their detrimental effects, TEs are often compared to parasites spreading selfishly within their host genome (Orgel and Crick, 1980). However, beneficial consequences of TE insertions have also been reported (González and Petrov, 2009; Arkhipova, 2018). One source of deleteriousness of TE multiplication is the resulting increase in genome size and complexity, which can lead to increased instability and thus to phenotypic changes (Kidwell and Lisch, 1997). Instability can come from ectopic genome recombination that can cause chromosomal instability and structural variants, such as translocations and inversions (Luning Prak and Kazazian, 2000). TEs can also affect gene expression by disrupting existing regulatory sequences at insertion sites or by providing new ones (Niu et al., 2019). For instance, some TEs contain promoters and enhancers that regulate their own expression and that can affect expression of neighboring genes (Naito et al., 2009; Lisch, 2013; Chuong et al., 2017). Hence, TE mobilization is likely to reduce the fitness of the host.

Hosts have TE repression mechanisms that limit the spread of TEs and that vary depending on the organism and the type of element (Saha et al., 2015; Sigman and Slotkin, 2016; Luo and Lu, 2017). However, the ability to regulate TEs could be disrupted in certain environmental or genomic conditions, including in hybrids (Horváth et al., 2017). According to the genomic shock hypothesis proposed by McClintock in 1984, stress and regulatory interference triggered by interspecific hybridization could lead to TE mobilization (McClintock, 1984). One possible cause of such interference could be the antagonistic coevolution between TEs and host repression. The merger of two divergent genomes brings together two TE populations that may differ in terms of composition and relative abundance, and which may form co-adapted units with their respective host genome. Hybridization may thus perturb these coevolved units, leading to a disruption of transposition regulation mechanisms and TE mobilization. Many studies comparing TE transcription level in parental species and their hybrids found an overexpression of some TEs in hybrids (Kelleher et al., 2012; Dion-Côté et al., 2014; Renaut et al., 2014; Lopez-Maestre et al., 2017), whereas others found no overexpression in hybrids for the majority of the TEs (Josefsson et al., 2006; Goebel et al., 2018). Thus, McClintock's hypothesis has been both confirmed and refuted depending on the system, so the phenomenon of genomic shock resulting from hybridization is not universal.

One of the model systems harboring active TEs but in which TE activity has not been found to increase in hybrids

is the budding yeast. Two recent studies examined whether hybridization causes an increase in the transposition rate of TEs in yeast hybrids and did not find evidence for it (Hénault et al., 2020; Smukowski Heil et al., 2021). The relative simplicity of the yeast genome, along with extensive knowledge of the life cycle of its TEs, makes it an ideal system to yield deeper insights into the regulation of transposition in hybrids. Yeast TEs are long terminal repeat (LTR) retrotransposons also called Ty elements, which replicate via a copy-and-paste mechanism including an RNA intermediate. Five main Ty families are found in the model species *S. cerevisiae*: Ty1, Ty2, Ty3, Ty4, and Ty5 (Kim et al., 1998; Carr et al., 2012). A larger diversity of TE families are found in other *Saccharomyces* species (Bleykasten-Grosshans and Neuvéglise, 2011). LTR retrotransposons comprise an internal coding sequence flanked by two LTRs in direct orientation. While active families have full-length copies, the vast majority of genomic insertions are non-coding solo LTR elements remaining from LTR-LTR recombination of full-length elements (Kim et al., 1998). The abundance of solo LTR and full-length Ty sequences varies depending on Ty families and strains (Bleykasten-Grosshans et al., 2013). Much knowledge on the biology of Ty elements (and LTR retrotransposons in general) stems from studies of the Ty1 and Ty3 families from *S. cerevisiae* (Curcio et al., 2015; Sandmeyer et al., 2015). The life cycle of the Ty elements is analogous to that of retroviruses (Rowley, 2017; Czaja et al., 2020). Ty elements are transcribed into mRNAs encoding two overlapping open reading frames (ORFs), GAG, and POL, which are, respectively, translated in the Gag and Gag-Pol proteins. Gag is a structural protein, while Gag-Pol is a catalytic protein containing protease, integrase, and reverse transcriptase domains. These Ty-encoded components are necessary for the reverse-transcription of the mRNA into cDNA and for the genomic integration of a new Ty copy, completing the transposition cycle.

The life cycle of Ty elements is complex, and measures of transposition are accordingly diverse. The recent investigations performed on yeast tested mobilization by measuring Ty copy number changes after experimental evolution (Hénault et al., 2020), or by experimentally measuring transposition rates with reporter assays (Smukowski Heil et al., 2021). Although these studies concluded in an absence of mobilization, they have not ruled out the possibility that TEs are transcriptionally derepressed, as seen for some animal and plant species (Kelleher et al., 2012; Dion-Côté et al., 2014; Renaut et al., 2014; Lopez-Maestre et al., 2017; Laporte et al., 2019). For instance, transcriptional derepression could be compensated by Ty regulation at the post-transcriptional or post-translational levels. Studying an intermediate stage of the Ty life cycle allows one to better understand the host mechanisms enabling or repressing transposons activity. Transcript abundance and intensity of translation are good proxies to evaluate the activity of these TEs in hybrids.

Our objective was to test if TEs are more expressed in F1 hybrids of yeast. We performed differential expression analysis on two publicly available RNA sequencing (RNA-seq) datasets of *Saccharomyces* species (*S. cerevisiae* × *S. uvarum* and *S. cerevisiae* × *S. paradoxus* diploid hybrids and their diploid

parents). We also analyzed a published ribosome profiling dataset of *S. cerevisiae* × *S. paradoxus* hybrids to evaluate Ty translation efficiency. We tested whether the total expression of each active Ty family is higher in hybrids than in their parental species at the transcriptional and the translational level. Our analyses show that Ty elements are generally not differentially expressed in hybrids, even when the hybrids are exposed to a low temperature stress condition.

## MATERIALS AND METHODS

### Dataset and Strain Description

We performed differential expression analysis using three RNA-seq datasets of diploid parental and lab constructed interspecific diploid hybrids of the *Saccharomyces* genus. The first dataset, which we will call DS1 throughout the remainder of the article, is from Schraiber and collaborators (Schraiber et al., 2013). DS1 includes two biological replicates of *S. uvarum* (CBS 7001) and *S. paradoxus* (CBS432) parents hybridized with *S. cerevisiae* (YHL068) and grown at 25°C. The second dataset (DS2) is from Hovhannisyan et al. (2020). DS2 includes three biological replicates of *S. cerevisiae* (YPS128), *S. uvarum* (UWOPS99-807.1.1) and their hybrids grown at two temperatures: 30°C and 12°C. The third dataset (DS3) is a ribosome profiling dataset that includes two replicates of hybrids of *S. cerevisiae* (S288C) and *S. paradoxus* (CBS432) and the matched total mRNA RNA-seq libraries (McManus et al., 2014). Sample description and accession numbers of DS1, DS2, and DS3 can be found in **Supplementary Tables 1, 2**. Raw RNA-seq data were retrieved from NCBI Sequence Read Archive (Leinonen et al., 2011) with SRA Toolkit version 2.10.8 with `–split-files` option <https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?cmd=show&f=software&m=software&s=software>.

Downstream analysis was slightly different for the datasets because different sequencing methods were used to generate them. For DS1, the authors used 3'-end RNA-seq on samples with polyadenylated transcripts enriched twice. cDNAs were sequenced using paired-end (PE) 36 bp mode on an Illumina IIX Genome Analyzer (Yoon and Brem, 2010). For DS2, libraries were sequenced in PE 50 bp (30°C) and PE 75 bp (12°C) mode on an Illumina HiSeq 2500 platform after a single poly-A mRNA selection. These data are also reverse stranded. For DS3, stranded mRNA libraries were prepared by poly-A enrichment and random fragmentation before sequencing in single-end 50 bp mode on an Illumina HiSeq 2000 instrument.

### Quality Control and Data Filtering

For DS1 and DS2, we used FastQC v0.11.8 <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> and MultiQC v1.9 (Ewels et al., 2016) for raw sequencing data quality control. For DS1, we used a custom script based on the ShortRead R package (Morgan et al., 2009) to only keep read pairs where one of the reads has two or more T nucleotides at the 5' end. This way, we discarded reads that had no sign of mRNA poly-A tail. For DS2, no filtering was done because we confirmed that there were no poly-A tails in raw reads. For both DS1 and DS2, we used Trimmomatic v0.36 (Bolger et al., 2014) with Truseq3

adaptors and the following parameters: 2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:30 for adaptor trimming. For DS3, reads were trimmed for base quality and adapter sequences using fastp v0.21.0 (Chen et al., 2018) with options `-l 20-5-3-r` and a custom Illumina adapter database.

### TE Reference Sequences and Annotations

Our custom library of representative internal and LTR sequences of each Ty family inserted in the reference genomes includes Ty1, Ty2, Ty3, Ty4, and Ty5 from *S. cerevisiae*, Ty1, Ty3, and Ty5 from *S. paradoxus* and Tsu4 from *S. uvarum*. The library of internal and LTR reference sequences of Ty1-Ty5 in *S. cerevisiae* was retrieved from file S1 of Carr and collaborators (Carr et al., 2012). For the Tsu4\_uva sequence, the only Ty family found in *S. uvarum*, we used the GenBank reference sequence under accession number AJ439550.1 (Neuvéglise et al., 2002). For the Ty1, Ty3, and Ty5 families in *S. paradoxus*, we generated representative reference sequences from the genome of the strain CBS432 (Yue et al., 2017) using custom Python v3.9.1 scripts (Van Rossum and Drake, 2009). Internal and LTR sequences were extracted from the whole-genome assembly using reference annotations (Yue et al., 2017). Multiple sequence alignments were produced using MUSCLE v3.8.31 (Edgar, 2004) and consensus sequences were derived using the BioPython package v1.76 (Cock et al., 2009) with a minimal frequency threshold of 0.5 and replacing ambiguous positions with Ns. BLASTN v2.7.0 (Camacho et al., 2009) was used to find the best hit (highest bit score) for each consensus among the annotated sequences, which was chosen as a reference. In the case of *S. cerevisiae* and *S. paradoxus*, the reference genomes from which the Ty reference sequences were retrieved in our study are often the exact same strains employed in the transcriptomics datasets.

Nucleotide divergence between *S. cerevisiae* and *S. paradoxus*, the two most closely related parental species of our datasets, is generally high enough to allow unambiguous short read mapping. To ensure this is the case for Ty families, we analyzed the sequence similarity profiles between the Ty1 and Ty3 families, which are orthologous between both species (namely Ty1\_cer-Ty1\_par and Ty3\_cer-Ty3\_par). We extracted all full-length annotated Ty1 and Ty3 sequences from the S288c and CBS432 genomes (Yue et al., 2017). We computed multiple sequence alignments for Ty1\_cer-Ty1\_par sequences and Ty3\_cer-Ty3\_par sequences using MUSCLE v3.8.31 (Edgar, 2004). Pairwise nucleotide identity between *S. cerevisiae* and *S. paradoxus* sequences was computed using custom Python v3.9.1 scripts (Van Rossum and Drake, 2009). Alignments were split into 100 bp non-overlapping windows. For each window, nucleotide identity between each pair of sequences was approximated by counting the number of mismatches, ignoring insertions and deletions. The average nucleotide identity was computed across all pairwise comparisons for a given window.

### Genome Sequences and Annotations

For DS2, whole genome assembly and annotations of *S. cerevisiae* YPS128 strain were downloaded in GenBank under accession code Bioproject PRJEB7245 (Yue et al., 2017). The reference genomes and annotations for the other strains in DS1 and

DS2 were downloaded from [www.saccharomycessensustricto.org](http://www.saccharomycessensustricto.org) (Scannell et al., 2011). For each species, we combined ultrascaffolds and unplaced regions in the same fasta file to get one complete reference genome, as described in Hovhannisyan et al. (2020). The annotation file format conversion was done with `gffread v0.12.4` (Pertea and Pertea, 2020) and the code in `GFF_GTF_conversion.txt` file found at [https://github.com/Gabalardonlab/Hybrid\\_project](https://github.com/Gabalardonlab/Hybrid_project). For DS1 and DS2, we used `RepeatMasker v4.1.0` (Smit et al., 2015) with `-gff` and `-lib 2` “custom TEs library” options to detect and mask interspersed repeats and low complexity DNA sequences by replacing them by Ns in the reference parental genome sequences. After masking Ty copies, we added our custom library of representative internal and LTR sequences of each Ty family to the fasta files and annotation files as if Ty sequences were additional chromosomes. Hybrid genomes were constructed by concatenating the corresponding masked parental genomes and adding reference sequences of Tys present in parental species. This was done so that each hybrid reference genome contains only a single copy of Ty from each family on which the RNA-seq reads could be aligned. We added the concatenated internal and LTR sequences to reference genomes and genome annotation for DS1, whereas we added our library of internal and LTR Ty sequences as separate fasta entries for DS2. Ty internal sequences were concatenated with their 3′ LTR in DS1 in order to select the read pairs including only one read mapping on the internal sequence and the other read mapping on the 3′ LTR along with read pairs mapping on the internal sequence. Since this dataset was generated with a 3′-end sequencing method, it is expected that the mapping should extend from the 3′ extremity of the Ty internal sequence to the end of the transcript, in the 3′ LTR.

For DS3, we generated a concatenated hybrid reference genome by merging chromosomes from *S. paradoxus* CBS432 and *S. cerevisiae* S288C reference genomes (Yue et al., 2017) and adding reference sequence for each Ty family that comprised the 5′ LTR and the complete internal sequence. Chromosomal sequences were hard-masked (replaced by Ns) for all Ty sequences using reference annotations and a custom Python script.

## RNA-seq Mapping

Alignments of RNA-seq reads of hybrids and their parents were performed on concatenated masked parental genomes supplemented with Ty reference sequences. For DS1, we used `Bowtie2 v2.3.4.1` (Langmead and Salzberg, 2012) with `-local` option to do a soft clipped alignment with reads having poly-T at 5′ end, as described in Schraiber et al. (2013). To exclude spurious transcription from solo LTRs, we kept only pairs for which the first read mapped inside an internal sequence using `samtools view` with `-f 67 -h` options and the coordinates of internal sequences on the pseudo-chromosomes. Lists of unique read IDs were extracted from the resulting bam files. These lists allowed us to retrieve the corresponding pairs using the `FilterSamReads` tool from `picard-tools v2.23.2` <https://broadinstitute.github.io/picard/> with option `FILTER=includeReadList`. Bam files of Ty elements and of all host chromosomes were combined with

`GatherBamFiles` of `picard-tools`. For DS2, we performed RNA-Seq read mapping with the splice-junction aware mapper `STAR` (Dobin et al., 2013) with default parameters, as described in Hovhannisyan et al. (2020). For DS3, reads were mapped on the concatenated references using `HISAT2 v2.2.1` (Kim et al., 2019). Secondary alignments were filtered out using `samtools view v1.9` (Li et al., 2009) with option `-F256`.

## Differential Expression Analysis

For genes and TE total expression quantification of DS1 and DS2, we used the `Rsubread featureCounts` R function (Liao et al., 2014) to assign mapped reads to genomic features and generate raw read count matrices of Ty. We used `featureCounts` with `minFragLength=5`, `GTF.attrType=“transcript_id”` options for the non-stranded DS1. For the reverse-stranded DS2, we used `countMultiMappingReads=TRUE`, `fraction=TRUE`, `strandSpecific=2`, `minFragLength=5`, `GTF.attrType=“transcript_id.”`

We normalized data of raw read count matrices with the median of ratios normalisation method implemented in `DESeq2` (Love et al., 2014). This normalization accounts for sequencing depth and RNA composition using the negative binomial distribution. To evaluate if TE expression differs between hybrid and parental genomes, we performed differential expression analysis with `DESeq2` (Love et al., 2014), which uses Wald test for significance testing. For both datasets, we compared Ty expression from their internal sequences in hybrids with the ones in the parent having full-length elements. Parental species were set as the first level in the treatment factor and only the subgenomes of these parental species (including their corresponding Ty sequences) were kept for `DESeq2` analysis of each correspondent hybrid and parental samples. For DS2, we performed a multi-factor analysis with `design = counts ~ temperature + species + species:temperature`. Since the temperature effect was only significant for the Ty1 family, we normalized and analysed Ty1 expression data at 30 and 12°C separately. To identify Ty families differentially expressed in hybrids, we used an adjusted *p*-value threshold of 0.05.

For DS3, analyses were conducted with scripts from the `Plastid Python` library (Dunn and Weissman, 2016). Since reference transcript annotations included no untranslated regions (UTRs), gene annotations were modified to add a 50 bp padding upstream of each CDS to simulate 5′ UTRs, allowing to capture 5′ offsets of ribosome protected fragments (RPF) reads. For RPF libraries, 5′ offsets were estimated using the `psite` script from `Plastid` (Dunn and Weissman, 2016) to yield the position of ribosomal P-sites for each read length. The `phase_by_size` script was used to determine which read lengths yielded optimal phasing. We restricted read lengths to the 27–32 bp range inclusively. Read counts per position and per gene were computed using the `get_count_vectors` and `counts_in_region` scripts, respectively, from `Plastid` (Dunn and Weissman, 2016), with variable 5′ offsets. For total mRNA libraries, read counts per gene were computed using the `counts_in_region` script from `Plastid` (Dunn and Weissman, 2016) with the central position of each gene. Counts were analyzed in `DESeq2` (Love et al., 2014) by fitting the model `counts ~ species +`

experiment + species:experiment, where the species categorical variable includes the hybrid and one of its parental species, and the experiment categorical variable comprises total mRNA and RPF. The latter interaction term corresponds to translation efficiency. Two independent models of this class were run for sets of genes/Ty and samples corresponding to each parental subgenome. Similarly, we ran a separate class of DESeq analyses, but restricting to mRNA libraries and using the model counts ~ species. For all models, log<sub>2</sub> fold change values were corrected with the default apeglm shrinking method (Zhu et al., 2019).

## Coverage Depth Uniformity

For DS1 and DS2, we generated heatmaps of RNA-seq coverage depth along Ty sequences by generating bed files with samtools depth -a to get per-base coverage depth values along Ty sequences. For DS3, the output of the get\_count\_vectors script from Plastid was used to get position-wise coverage depth values (Dunn and Weissman, 2016). Custom Python scripts were used to compute z-scores of mean coverage depth in 75 bp-wide non-overlapping windows. For the three datasets, we also calculated Transcripts Per Million (TPM) for each sample.

## Code

All bash, R and Python code used in this study can be found at [https://github.com/Landrylab/Drouin\\_et\\_al\\_2021](https://github.com/Landrylab/Drouin_et_al_2021).

## RESULTS

In order to determine if TEs are derepressed in newly formed hybrids, either at the transcription or the translation level, we performed differential expression analysis on three publicly available expression datasets: DS1, DS2 and DS3 (Schraiber et al., 2013; McManus et al., 2014; Hovhannisyan et al., 2020). We tested whether the expression of each active Ty family is different in hybrids compared to their parental species. These datasets include *S. cerevisiae* x *S. uvarum* (Sc x Su) and *S. cerevisiae* x *S. paradoxus* (Sc x Sp) hybrids, as well as different growth temperatures (30 and 12°C) (Figure 1A).

## Coverage Along Ty Sequences

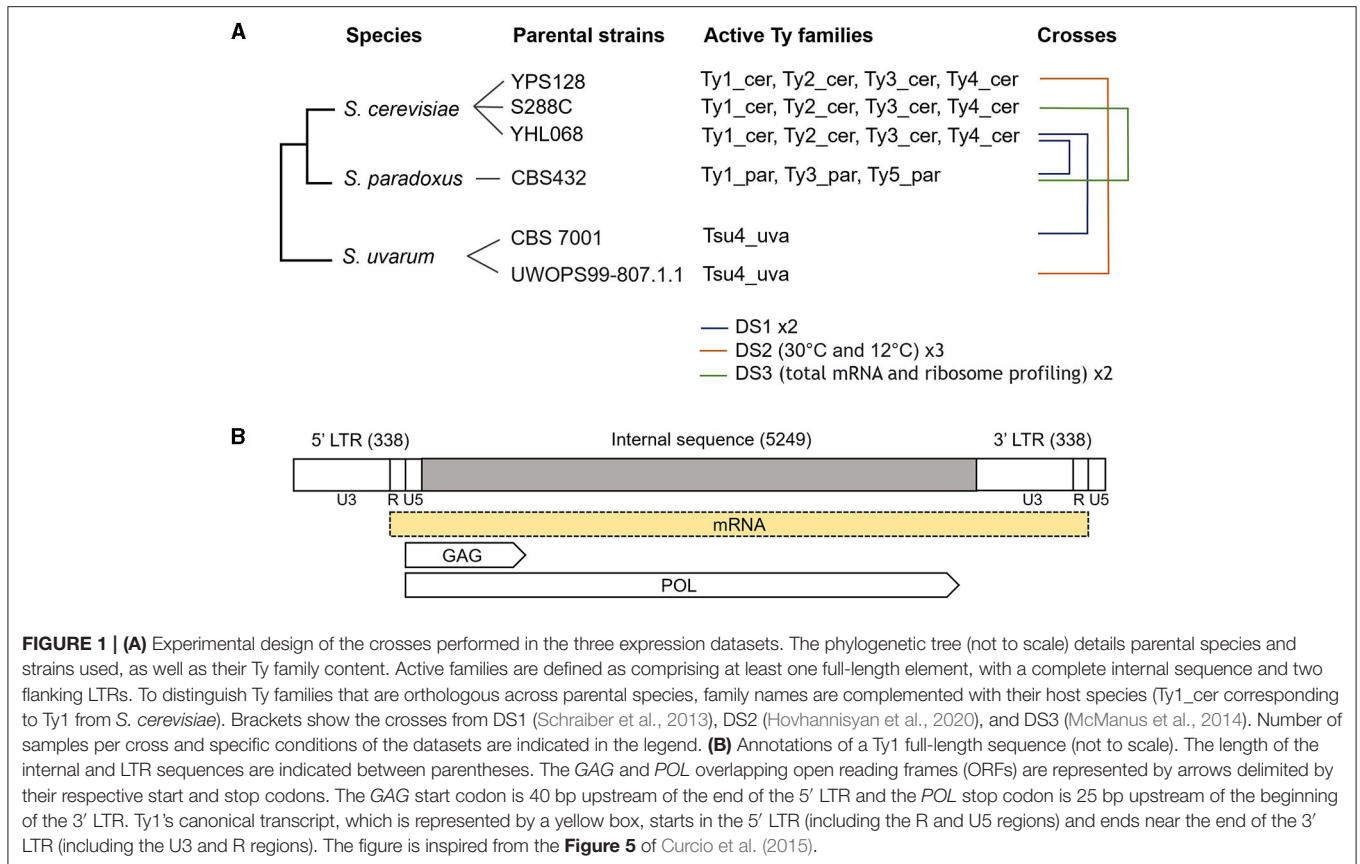
To confirm that alignments of RNA-seq reads on Ty sequences were comparable between parental and hybrid samples, we first examined the coverage along Ty sequences (Figure 1B). For DS1, coverage is concentrated in narrow peaks along the sequences (Figure 2A). The distributions of peaks differ between Ty families and contrast the expectation from the 3' end RNA-seq, in which coverage should be limited to the 3' portion of transcripts (Tandonnet and Torres, 2017; Ma et al., 2019). Since we select for A nucleotides at the 3' end of RNA-seq reads, homopolymers of A in the reference sequences could be associated with higher mapping, for instance because of library contamination by fragments that would not correspond to the 3' portion of transcripts. However, we generally observe no poly-A stretches that could explain the high coverage in these regions. This suggests that major alternative polyadenylation sites may exist in Ty internal sequences. Nevertheless, these distributions are consistent within Ty families, thus allowing

us to compare hybrid and parental samples. For DS2 and DS3, the coverage is fairly uniform along the sequence, as expected from mRNA sequencing with a single poly-A selection (Figure 2B, Supplementary Figure 1). Both replicates of parents and hybrids also have similar coverage profiles between the two temperatures in DS2. Furthermore, for all datasets, some parental species known to be devoid of full-length elements for a given family exhibit low and irregular coverage, consistent with non-specific mapping. These results suggest that RNA-seq data can be harnessed for differential expression analysis of Ty elements between hybrids and parental samples, and between environmental conditions.

One exception is the Ty1 family of *S. cerevisiae* and *S. paradoxus*, which shows evidence of substantial non-specific mapping (i.e., *S. paradoxus* reads mapping on Ty1\_cer and vice-versa; Figure 2, Supplementary Figure 1). Some of this non-specific mapping is due to the close sequence similarity of the Ty1 GAG region between the two species (Supplementary Figure 2), consistent with interspecific introgression for this locus (Czaja et al., 2020). We thus excluded the GAG region from the analysis of Ty1 in Sc x Sp hybrids. Nevertheless, a substantial fraction of non-specific mapping occurred outside of GAG. Using libraries of the parental species, we estimated that between 5 and 15% of Ty1 reads of a given species map on the reference sequence of the cognate species. While this non-specific mapping biases both Ty1\_cer and Ty1\_par expression levels, the latter has at least five times lower expression levels across library types (Figure 2; Supplementary Figure 1) and, as such, is disproportionately more affected by non-specific mapping coming from highly abundant *S. cerevisiae* reads. As a consequence, we excluded Ty1\_par from our analyses.

## Most Ty Show no Differential Expression Between Parents and Hybrids

We performed differential expression analysis of the whole transcriptomes of DS1, DS2, and DS3, comparing hybrids and their parents. Analysis of DS1 reveals differences in Ty expression patterns between Sc x Su and Sc x Sp hybrids and their respective parental species (Figures 3A,B, 4A,B). Transcript levels are significantly higher in the Sc x Su hybrid for Tsu4\_uva (log<sub>2</sub> fold change: 0.383, FDR corrected *p*-value: 0.0474) and Ty1\_cer (log<sub>2</sub> fold change: 0.433, FDR corrected *p*-value: 1.95 × 10<sup>-4</sup>), whereas other Ty families show similar expression (Figures 3A, 4A). No Ty family is overexpressed in Sc x Sp hybrids (Figures 3B, 4B). Instead, Ty1\_cer, Ty2\_cer, and Ty4\_cer are significantly underexpressed in hybrids, while the other Tys show no significant difference. Since Ty1\_cer is subject to asymmetrical non-specific mapping (Figure 2B), we aimed to assess the robustness of our differential expression analysis for this family. Only ~2% of the *S. paradoxus* Ty1 reads are estimated to map on Ty1\_cer, so we considered this non-specific mapping negligible. We manually raised the Ty1\_cer read counts by an amount corresponding to the estimated rate of non-specific mapping of *S. cerevisiae* Ty1 reads on Ty1\_par (~12%). This analysis showed that, even after our manual count correction, Ty1\_cer is still significantly underexpressed in Sc



× *Sp* hybrids. The multi-factor analysis of DS2 reveals that, when grouping samples grown at both 30 and 12°C, there is no Ty family overexpressed in *Sc* × *Su* hybrids compared to their parents (**Figure 3C**). Except for Ty1, temperature does not significantly affect the expression of Ty families (**Figure 3D**) and there is no significant interaction between species and temperature factors (**Figure 3E**). We also analysed samples of both temperatures independently (**Figures 4C,D**). Our results reveal that all the Ty families, either grown at 12 or 30°C, show similar expression in *Sc* × *Su* hybrids and in *S. cerevisiae* or in *S. uvarum* (**Figures 3C, 4D**). Even if Ty1 expression is slightly higher in hybrids, the difference is not significant. Our results show that Ty elements are not more often differentially expressed in hybrids compared to in their parents when they are grown at 12°C. This suggests that Ty regulation remains unaffected by hybridization, even in the low temperature condition. Finally, *Sc* × *Sp* hybrids in DS3 show no significant difference in transcript levels for any Ty family (**Figure 5B, Supplementary Figure 3**).

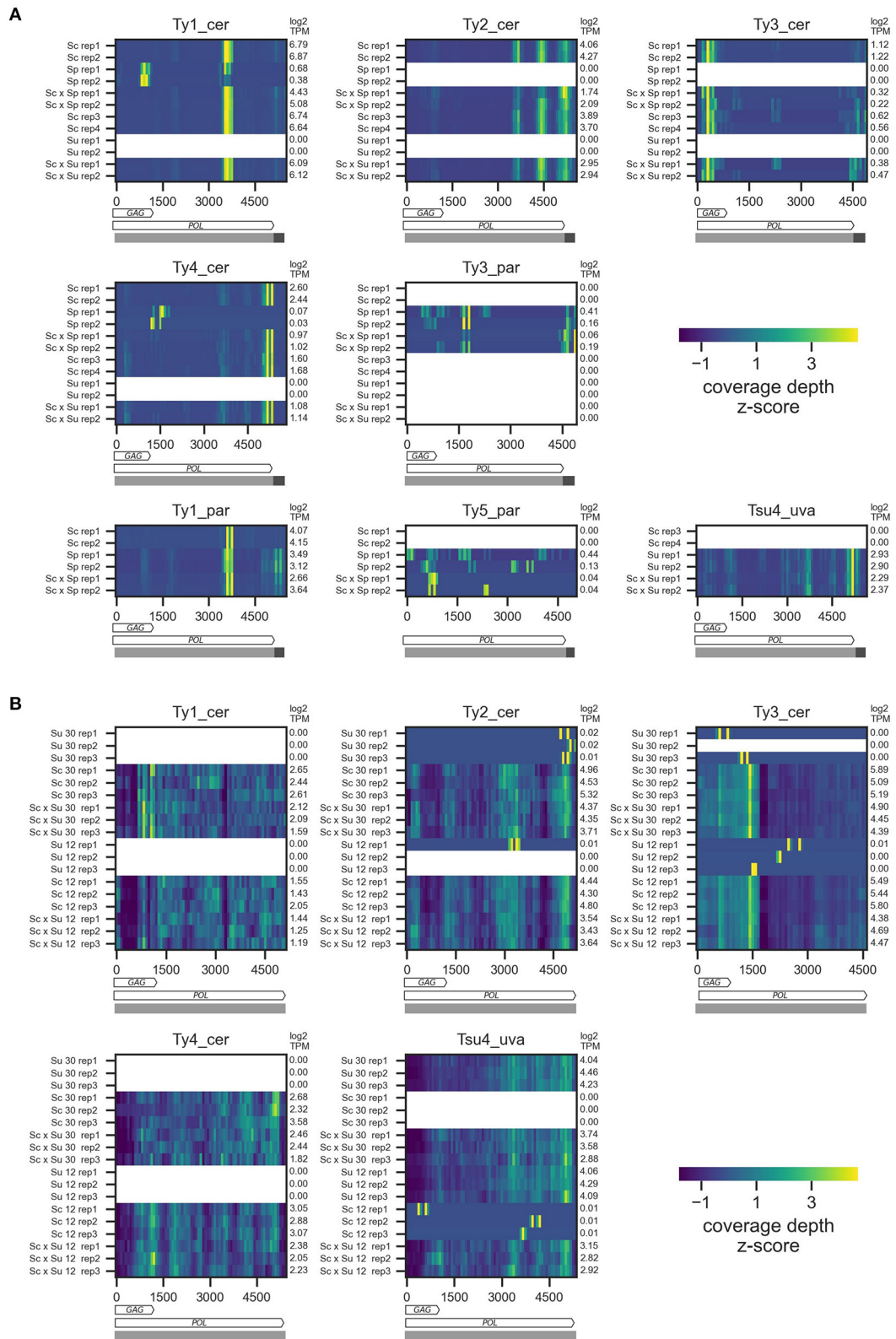
In the three datasets, the normalized read counts for each Ty family is consistent with the copy number of Ty complete sequences in parental genomes (**Figures 2, 5, Supplementary Table 3**). Since the total normalized read counts correspond to the sum of the transcript levels of each individual Ty copy within a given family, this result suggests that average transcript levels per copy are fairly consistent across families. Ty3\_cer of DS2 is an exception

because it is more expressed than Ty2\_cer even though the *S. cerevisiae* genome contains fewer copies of Ty3 than Ty2.

In summary, differential expression analysis reveals no sign of systematic genomic shock effect on Ty expression in hybrids. Among the 26 hybrid-parent comparisons, only 2 Ty families show a significant expression increase in hybrids (Ty1\_cer and Tsu4\_uva in *Sc* × *Su* hybrids of DS1).

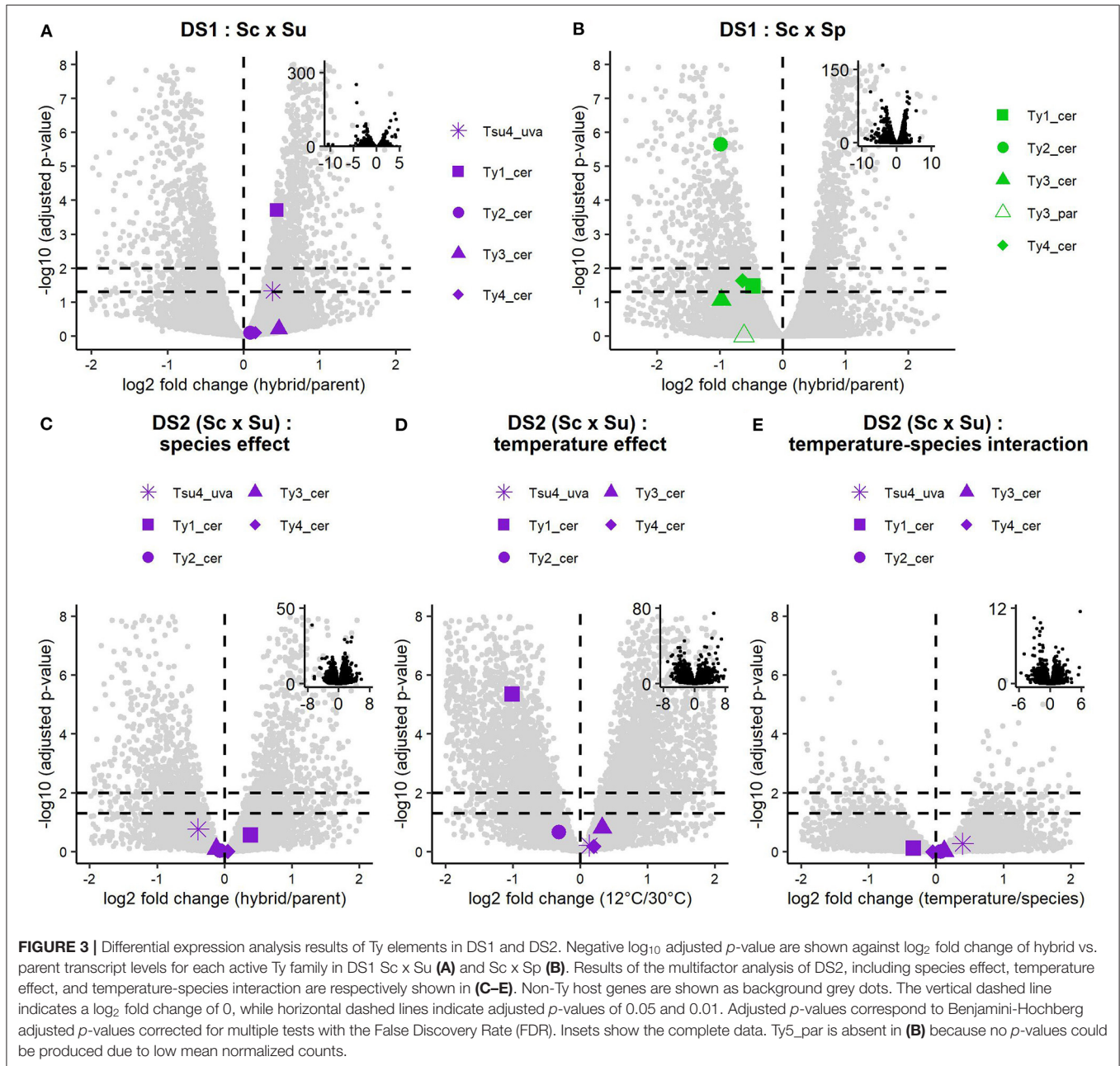
## Translation Efficiency

While Ty families generally show no significant difference in transcript levels, hybridization could modulate Ty expression at other levels. The level immediately downstream of transcription is translation efficiency, which corresponds to the intensity of ribosome occupancy per mRNA molecule. To examine this, we re-analyzed the ribosome profiling data of *Sc* × *Sp* hybrids in DS3 (McManus et al., 2014) to test for differences in translation efficiency between hybrids and their parents (**Figure 5; Supplementary Figures 1, 3, 4**). When measured for the initial 60 nucleotides of the coding sequence (CDS) of each Ty family, none exhibit a significant difference in translation efficiency between the parents and the hybrid (**Figure 6**). The same analysis on full-length CDS revealed significantly higher translation efficiency in the hybrid for the Ty1 family of *S. cerevisiae*. It also revealed substantial expression of Ty1\_cer in the *S.*



**FIGURE 2 |** RNA sequencing coverage depth along Ty reference sequences for DS1 (A) and DS2 (B). Z-scores of mean coverage depth are shown for each sample by 75 bp-wide non-overlapping bins along Ty reference sequences. In (A), rows show the two replicates for each parental species and hybrids. In (B), rows show the (Continued)

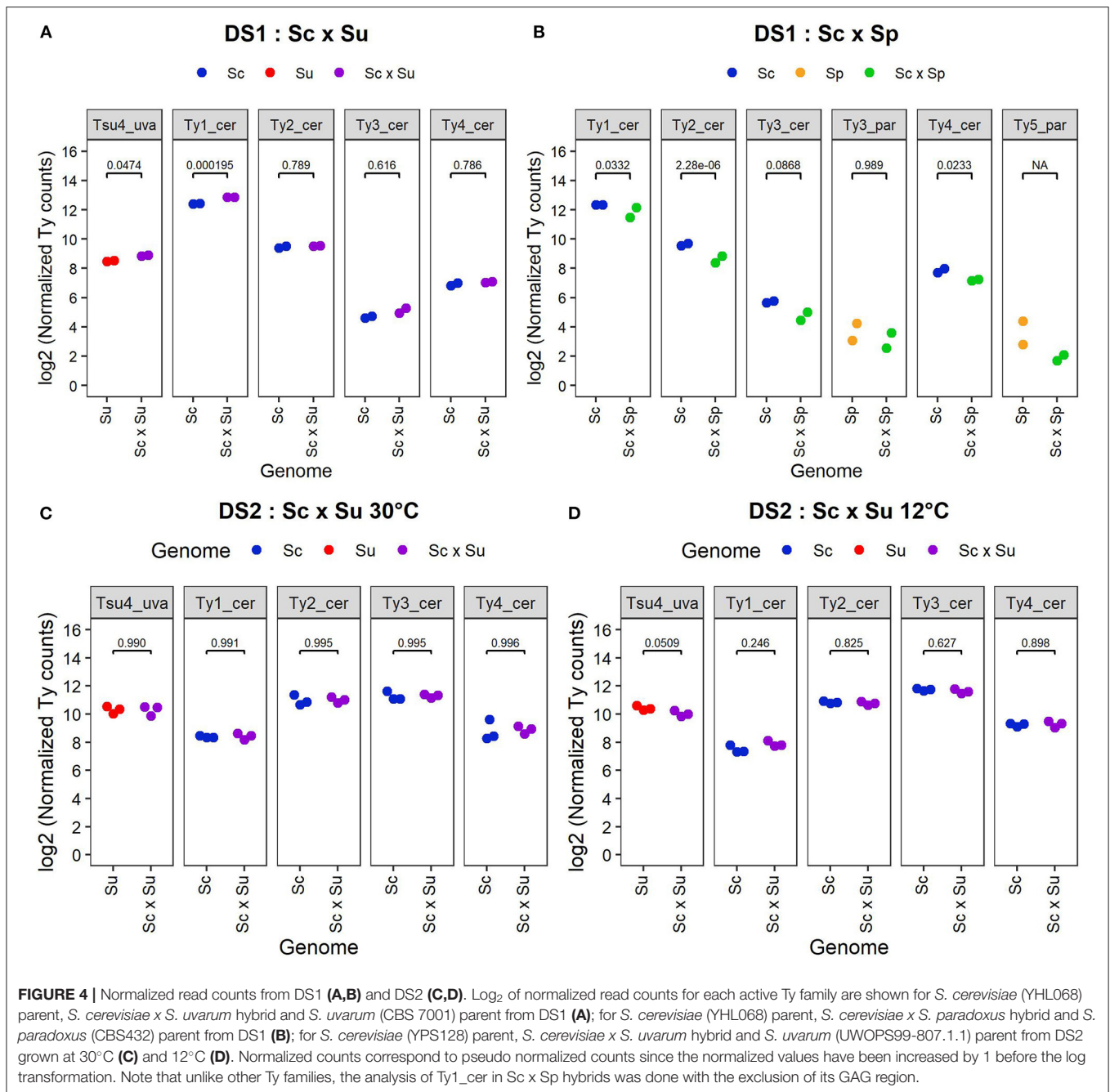
**FIGURE 2** | three replicates for each parental species and hybrids at the two temperatures. For each Ty family, crosses for which one of the parents is known to harbor this family are shown. As Ty3\_par was included in the Ty analysis of *S. cerevisiae* in Carr et al. (2012), we also investigated its content in Sc x Su samples. Transcripts Per Million (TPM) of each sample are indicated at the right of the heatmaps. Horizontal bars below the heatmaps represent the annotation of the reference sequences used for read mapping, with internal sequences and LTRs shown, respectively, in light and dark grey. The POL and GAG (except for Ty5\_par) ORFs are shown as arrows.



*paradoxus* parent as observed in DS1. This signal is consistent with the close relatedness of the GAG Ty1 ORF between species (**Supplementary Figure 2**, Czaja et al., 2020), as its exclusion revealed markedly decreased expression for the POL-exclusive region only, particularly for Ty1\_cer (**Figure 5A**). In Ty1, the

expression of the full-length POL ORF relies on a programmed frameshift that escapes the premature stop codon of the GAG ORF. Thus, our results suggest that the regulation of this frameshift in hybrids may lead to a significant increase in translational efficiency.



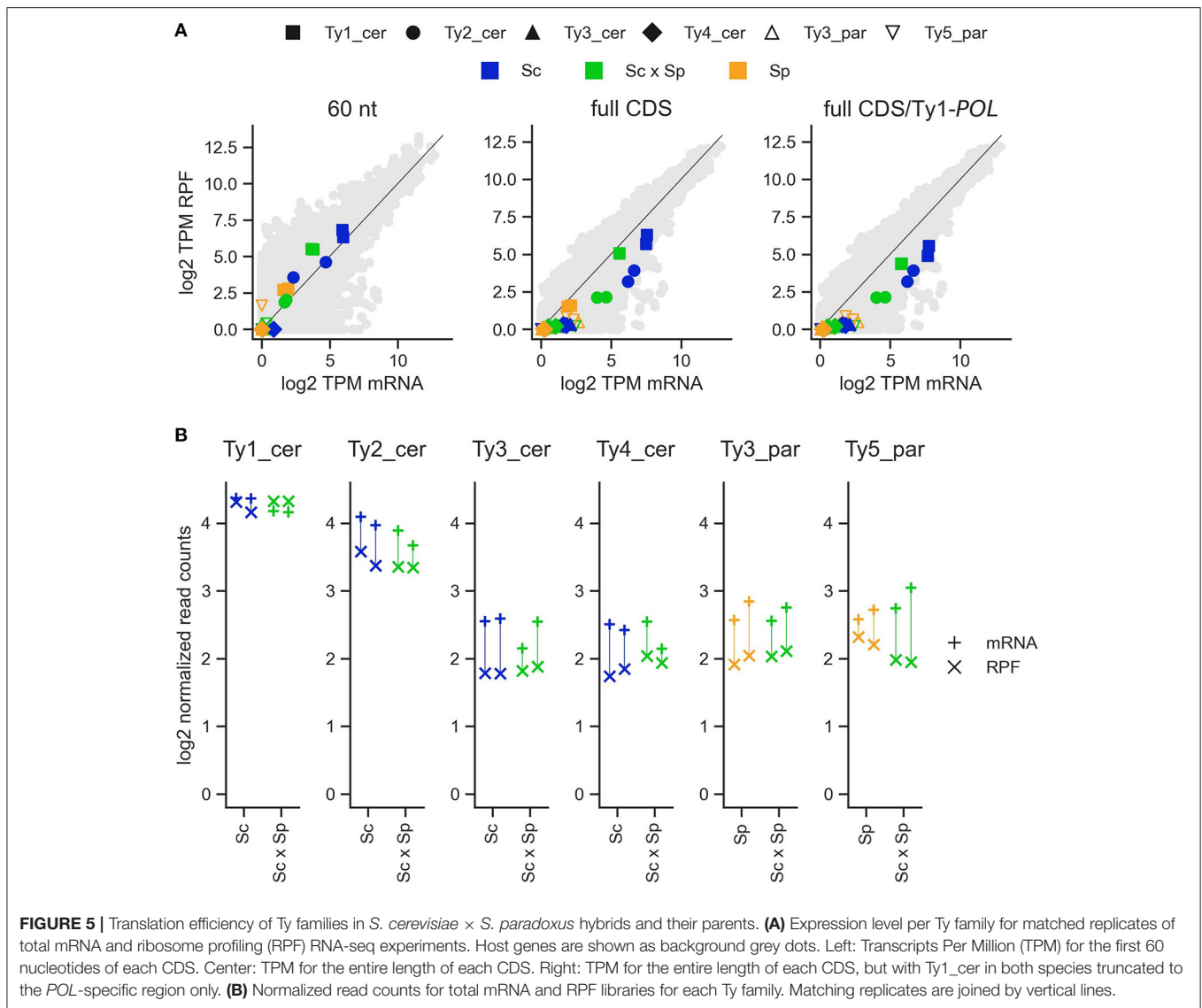


## DISCUSSION

In this study, we investigated the transcriptional and translational activity of yeast TE families in different genomic contexts and environmental conditions to examine whether they were more expressed in hybrids. We focused on TE transcription and translation levels since yeast retrotransposon activity has mostly been studied by measuring transposition rate in genomes. We performed differential expression analysis on two RNA-seq datasets (DS1, Schraiber et al., 2013; and DS2, Hovhannisyanyan et al., 2020) and one ribosome profiling dataset (DS3, McManus

et al., 2014) of *Saccharomyces* yeast hybrids. To investigate if total expression of Ty families are higher in hybrids than in their parents, we quantified RNA-seq reads mapping on a library of reference sequences comprising every known Ty family. We also evaluated translation efficiency in Sc x Sp hybrids (DS3) by quantifying read mappings from ribosome-protected mRNA fragments.

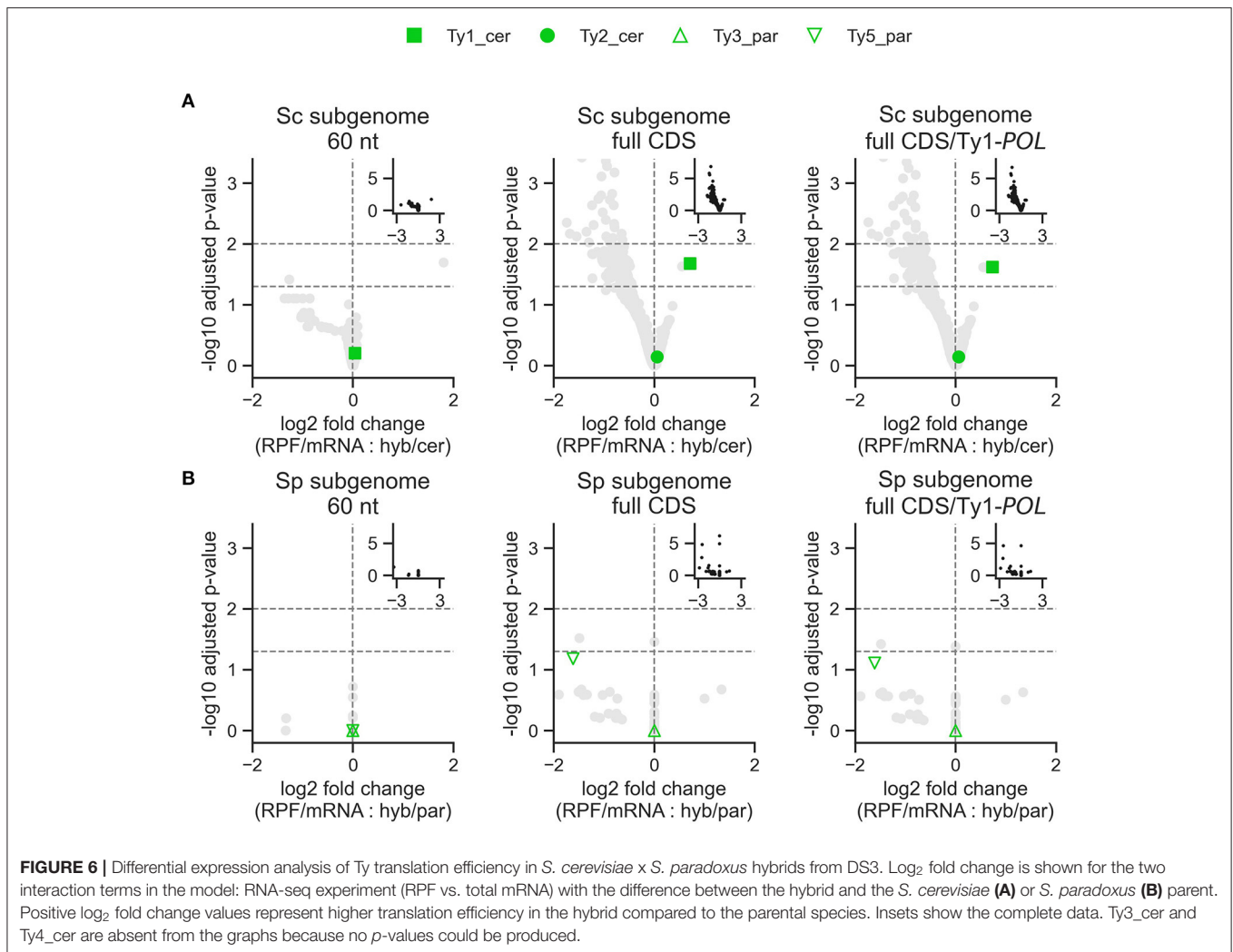
Overall, our results demonstrate that hybridization generally does not alter TE expression level, with few exceptions of small effect sizes. Among the 26 hybrid-parent total mRNA comparisons, only two Ty families show a significant but small



expression increase in hybrids (Ty1\_cer and Ty4\_uva in Sc × Su hybrids of DS1). Other Ty families are either not differentially expressed (21/26) or significantly less expressed (3/26) in hybrids. The only overexpressed families were found in the hybrids having the highest parental divergence level, as *S. uvarum* is more distant phylogenetically to *S. cerevisiae* than *S. paradoxus* is (Alsammar and Delneri, 2020). *S. uvarum* is 20 million years divergent from *S. cerevisiae* (Smukowski Heil et al., 2021) whereas *S. paradoxus* and *S. cerevisiae* diverged 5–10 million years ago (Tirosch et al., 2009). This result appears consistent with regulatory divergence increasing with time and causing an accumulation of incompatibilities. Although we would expect to see the impact of a genomic shock at higher frequency in the more divergent species, unlike DS1, our analyses of the same species cross (Sc × Su hybrids) in DS2 revealed no overexpression. The three datasets include different parental strains, growth conditions and RNA-seq library preparation methods. While our analyses included two Sc × Sp hybrids, they correspond

to the same parental strains (YHL068 being a close derivative of S288C) and were grown at similar temperatures (25 and 30°C). In contrast, the three Sc × Su hybrids spanned more diversity in terms of parental strains and culture conditions. The higher diversity sampled in Sc × Su hybrids likely increases the odds of observing significant differences in Ty expression. Thus, we found no systematic effect of genomic shock following hybridization on Ty transcriptional activity, and no strong support for a role of regulatory divergence in hybrids.

Our results are consistent with two recent studies showing no transposition rate increase in yeast hybrids (Hénault et al., 2020; Smukowski Heil et al., 2021), as well as with many studies in plants showing no expression changes in hybrids for most TE families (Josefsson et al., 2006; Goebel et al., 2018). Our findings contrast with studies performed in numerous systems such as *Drosophila* (Kelleher et al., 2012; Lopez-Maestre et al., 2017), fish (Dion-Côté et al., 2014; Laporte et al., 2019), and plant (Renaut et al., 2014) showing an overexpression of some



TE families in hybrids. Our work also contrasts with studies that found a significantly higher retrotransposon transposition rate in *Drosophila* and mammalian hybrids (O'Neill et al., 1998; Labrador et al., 1999). Patterns of TE expression seems to vary depending on species and types of elements.

TE transcriptional reactivation in diverse fungi has been shown in many stress conditions, including nutrient starvation and host infection stress (Esnault et al., 2019; Fouch et al., 2020). In plants, de-repression of TEs under stress usually impacts TE transcription levels and can increase transpositional activity (Dubin et al., 2018). *S. cerevisiae*'s Ty1 can be transcriptionally activated following adenylic nucleotide depletion (Servant et al., 2012) and have transposition rate 100 times higher at 15°C or 20°C compared to 30°C (Paquin and Williamson, 1984; Garfinkel et al., 2005). In addition to samples grown near the optimal growth temperature of *S. cerevisiae* in the laboratory (25 or 30°C), the datasets we reanalyzed comprised samples grown at a lower temperature (12°C), allowing us to investigate the thermosensitivity of Ty1 (Boeke et al., 1986). Again, we found Ty expression to be the same in hybrids and parents. Sc x Su hybrids include the cryotolerant species *S. uvarum*

(Salvadó et al., 2011), which could mitigate Ty mobilization at low temperatures. Our results are similar to the ones reported by Goebel and collaborators, who found no increase in TE expression in *Arabidopsis* hybrids, even when they were exposed to dehydration stress (Goebel et al., 2018). Our results suggest that changing the environment in which yeast hybridization occurs has a limited impact on retrotransposon activity in hybrids, although we cannot exclude that other conditions or stresses may have an impact.

To extend our analysis beyond transcription, we analyzed a published dataset comprising standard mRNA-seq and RPF sequencing of Sc x Sp artificial hybrids (McManus et al., 2014). Our analyses revealed contrasting results depending on the method employed. Quantification of RPF reads mapping to the first 60 nucleotides downstream of start codons should provide a readout that approximates the level of translation initiation (Ingolia, 2014). Using this metric, no Ty family showed differential translation efficiency in hybrids. However, when considering RPF levels over the whole length of coding sequences, Ty1\_cer was translated more efficiently in hybrids. The POL-specific section of Ty1's coding sequence appears

sufficient to yield this pattern. The *POL* ORF is translated following a programmed frameshift which escapes the stop codon of *GAG* (Curcio et al., 2015). Our results may be explained by a higher frameshifting rate in hybrids, yielding more *POL* translation per initiation event. In *S. cerevisiae*, one factor known to regulate the frequency of frameshifting is the abundance of tRNA-Arg(CCU), which binds the codon right before the frameshift (Kawakami et al., 1993). Alternatively, translation of the *POL*-specific may be slower in hybrids, yielding higher ribosome occupancy levels. Disentangling these hypotheses will require direct quantification of *Gag* and *Gag-Pol* protein abundance in hybrids. Overall, these results suggest that the regulation of Ty expression can be affected by other means than transcription in hybrids.

We see four main limitations to our study. The first one is the challenge to estimate Ty transcription level in some datasets in which sequence coverage of the elements is highly heterogeneous. The fact that an important proportion of the mRNA signal comes from small regions of the Ty sequences could be explained by the presence of alternative polyadenylation sites in Ty sequences. For example, the expression peak at around 800–1,000 bp in Ty1\_cer of DS2 corresponds to the 5' end of the Ty1 internal (Ty1i) transcript, a shorter sense strand RNA encoding Gag protein p22 and p18 (Saha et al., 2015). The peak is in accordance with previously published northern blots reporting Ty alternative transcripts (Fulton et al., 1988; Hug and Feldmann, 1996). Further analysis would be needed to confirm that the different patterns of expression observed in other Ty families also correspond to alternative transcripts. The second limitation is that we quantified transcript abundance at the level of whole families instead of individual elements. In *Saccharomyces*, most full-length Ty elements not yet lost by LTR-LTR recombination are young and highly similar (Carr et al., 2012; Hénault et al., 2020), hindering their discrimination by short read mapping. Individual copies may exhibit wide variation in expression within a Ty family, as illustrated by the 50-fold expression range measured for Ty1 copies in *S. cerevisiae* (Morillon et al., 2002). Thus, we cannot exclude that some Ty copies are overexpressed in hybrids. Whether hybridization alters the variance in transcript levels across individual copies, rather than the mean, remains an open question. The third limitation is that TE mobilization in interspecific hybrids could also be explained by mechanisms that we did not investigate. For instance, a recent study suggests that mitochondrial DNA inheritance from one parent species or the other can affect transposition rate (Smukowski Heil et al., 2021). None of the artificial hybrids reported in the studies we re-analyzed were characterized for mitochondrial DNA inheritance. The final limitation is that the genomic stability of yeast hybrids may depend on the genotypes of the parental strains used (Marsit et al., 2021) and impact the accumulation of TEs (Hénault et al., 2020). Taken individually, our analyses used a single strain per species, offering an incomplete assessment of the intraspecific diversity even if taken together they cover some parental strains of *Sc* and *Su*.

The nature of TE repression mechanisms in yeast may provide limited opportunity for the deregulation of TE activity levels in hybrids (Hénault, 2021). TE overexpression has

been mostly observed in species regulating TE propagation at the transcriptional level. Mammals and plants mostly repress transposition via epigenetic modifications affecting chromatin accessibility such as DNA methylation and histone modification (Slotkin and Martienssen, 2007). In *Drosophila*, there are two principal small RNA pathways repressing TE mobilization, the piRNA (piwi-interacting RNA) pathway and the siRNA (small interfering RNA) pathway (Merel et al., 2020). Unlike many elements and species, yeast Ty1 elements are self-repressed by a post-translational copy number control (CNC) mechanism that prevents the transposition of an already abundant Ty family (Saha et al., 2015; Czaja et al., 2020). This means that the Ty1 elements from the parental genomes come with their own regulatory mechanisms repressing their own transposition. The fact that Ty1 repression is genome dependent could explain why Ty1 is only overexpressed in some hybrids on DS1. For example, if the strength of self-repression depends on Ty copy number (Garfinkel et al., 2003), variation in Ty copy number among parental strains could determine the strength of the regulation in the resulting hybrids.

In conclusion, our results suggest that Ty transcriptional regulation is quite robust to interspecific hybridization in yeast since we found only a few cases of Ty differential expression. Further studies of TE mobilization incorporating diverse species, TE family repertoires and environmental conditions are needed to assess whether our findings can be generalized to fungal hybridization.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found on NCBI Sequence Read Archive (SRA). The accession numbers of DS1 (Schraiber et al., 2013) correspond to SRR515220-SRR515231 and SRR835213-SRR835218 (GEO accession number: GSE38875). The accession numbers of DS2 (Hovhannisyann et al., 2020) are SRR10246851-SRR10246868 (Bioproject number: PRJNA576452). The accession number of DS3 (McManus et al., 2014) are SRR948549-SRR948564 (Bioproject number: PRJNA213844).

## AUTHOR CONTRIBUTIONS

MD, MH, JH, and CRL contributed to the conception and design of the study. MD performed the analysis of DS1 and DS2 and wrote the first draft of the manuscript. MH performed the analysis of DS3. MD, MH, and CRL wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

MD was supported by an Alexander Graham Bell master's scholarship from NSERC (BESC M), a graduate scholarship from PROTEO and a Leadership and Sustainable Development Scholarship from Université Laval. MH was supported by an Alexander Graham Bell doctoral scholarship from NSERC. JH was funded by the Human Frontiers Science Program fellowship

(LT000182/2019-L). The work was funded by a NSERC Discovery grant (RGPIN-2020-04844) to CRL. CRL holds the Canada Research Chair in Cellular Systems and Synthetic Biology.

## ACKNOWLEDGMENTS

We thank Simon Aubé, Philippe Després, Romain Durand, and Anna Fijarczyk for providing useful

comments. We also thank all the Landry Lab members for their support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/ffunb.2021.729264/full#supplementary-material>

## REFERENCES

- Alsammar, H., and Delneri, D. (2020). An update on the diversity, ecology and biogeography of the *Saccharomyces* genus. *FEMS Yeast Res.* 20:foaa013. doi: 10.1093/femsyr/foaa013
- Arkhipova, I. R. (2018). Neutral theory, transposable elements, and eukaryotic genome evolution. *Mol. Biol. Evol.* 35, 1332–1337. doi: 10.1093/molbev/msy083
- Bleykasten-Grosshans, C., Friedrich, A., and Schacherer, J. (2013). Genome-wide analysis of intraspecific transposon diversity in yeast. *BMC Genomics* 14:399. doi: 10.1186/1471-2164-14-399
- Bleykasten-Grosshans, C., and Neuvéglise, C. (2011). Transposable elements in yeasts. *C. R. Biol.* 334, 679–686. doi: 10.1016/j.crvi.2011.05.017
- Boeke, J. D., Styles, C. A., and Fink, G. R. (1986). *Saccharomyces cerevisiae* SPT3 gene is required for transposition and transpositional recombination of chromosomal Ty elements. *Mol. Cell. Biol.* 6, 3575–3581. doi: 10.1128/MCB.6.11.3575
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Bourque, G., Burns, K. H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., et al. (2018). Ten things you should know about transposable elements. *Genome Biol.* 19:199. doi: 10.1186/s13059-018-1577-z
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform* 10:421. doi: 10.1186/1471-2105-10-421
- Carr, M., Bensasson, D., and Bergman, C. M. (2012). Evolutionary genomics of transposable elements in *Saccharomyces cerevisiae*. *PLoS ONE* 7:e50978. doi: 10.1371/journal.pone.0050978
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinform. Oxf. Engl.* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Chuong, E. B., Elde, N. C., and Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* 18, 71–86. doi: 10.1038/nrg.2016.139
- Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinform. Oxf. Engl.* 25, 1422–1423. doi: 10.1093/bioinformatics/btp163
- Curcio, M. J., Lutz, S., and Lesage, P. (2015). The Ty1 LTR-retrotransposon of budding yeast, *Saccharomyces cerevisiae*. *Microbiol. Spectr.* 3, 1–35. doi: 10.1128/microbiolspec.MDNA3-0053-2014
- Czaja, W., Bensasson, D., Ahn, H. W., Garfinkel, D. J., and Bergman, C. M. (2020). Evolution of Ty1 copy number control in yeast by horizontal transfer and recombination. *PLOS Genet.* 16:e1008632. doi: 10.1371/journal.pgen.1008632
- Dion-Côté, A.-M., Renaut, S., Normandeau, E., and Bernatchez, L. (2014). RNA-seq reveals transcriptomic shock involving transposable elements reactivation in hybrids of young lake whitefish species. *Mol. Biol. Evol.* 31, 1188–1199. doi: 10.1093/molbev/msu069
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635
- Doolittle, W. F., and Sapienza, C. (1980). Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284, 601–603. doi: 10.1038/284601a0
- Dubin, M. J., Mittelsten Scheid, O., and Becker, C. (2018). Transposons: a blessing curse. *Curr. Opin. Plant Biol.* 42, 23–29. doi: 10.1016/j.pbi.2018.01.003
- Dunn, J. G., and Weissman, J. S. (2016). Plastid: nucleotide-resolution analysis of next-generation sequencing and genomics data. *BMC Genomics* 17:958. doi: 10.1186/s12864-016-3278-x
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Elliott, T. A., and Gregory, T. R. (2015). What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. *Philos. Trans. R. Soc. B Biol. Sci.* 370:20140331. doi: 10.1098/rstb.2014.0331
- Esnault, C., Lee, M., Ham, C., and Levin, H. L. (2019). Transposable element insertions in fission yeast drive adaptation to environmental stress. *Genome Res.* 29, 85–95. doi: 10.1101/gr.239699.118
- Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354
- Fouché, S., Badet, T., Oggenfuss, U., Plissonneau, C., Francisco, C. S., and Croll, D. (2020). Stress-driven transposable element de-repression dynamics and virulence evolution in a fungal pathogen. *Mol. Biol. Evol.* 37, 221–239. doi: 10.1093/molbev/msz216
- Fulton, A. M., Rathjen, P. D., Kingsman, S. M., and Kingsman, A. J. (1988). Upstream and downstream transcriptional control signals in the yeast retrotransposon, TY. *Nucleic Acids Res.* 16, 5439–5458. doi: 10.1093/nar/16.12.5439
- Garfinkel, D. J., Nyswaner, K., Wang, J., and Cho, J.-Y. (2003). Post-transcriptional cosuppression of Ty1 retrotransposition. *Genetics* 165, 83–99. doi: 10.1093/genetics/165.1.83
- Garfinkel, D. J., Nyswaner, K. M., Stefanisko, K. M., Chang, C., and Moore, S. P. (2005). Ty1 copy number dynamics in *saccharomyces*. *Genetics* 169, 1845–1857. doi: 10.1534/genetics.104.037317
- Goebel, U., Arce, A., He, F., Rico, A., Schmitz, G., and Meaux, J. de. (2018). Robustness of transposable element regulation but no genomic shock observed in interspecific arabidopsis hybrids. *Genome Biol. Evol.* 2018, 1403–1415. doi: 10.1093/gbe/evy095
- González, J., and Petrov, D. A. (2009). The adaptive role of transposable elements in the *Drosophila* genome. *Gene* 448, 124–133. doi: 10.1016/j.gene.2009.06.008
- Hénault, M. (2021). The challenges of predicting transposable element activity in hybrids. *Curr. Genet.* 67, 567–572. doi: 10.1007/s00294-021-01169-0
- Hénault, M., Marsit, S., Charron, G., and Landry, C. R. (2020). The effect of hybridization on transposable element accumulation in an undomesticated fungal species. *eLife* 9:e60474. doi: 10.7554/eLife.60474.sa2
- Horváth, V., Merenciano, M., and González, J. (2017). Revisiting the relationship between transposable elements and the eukaryotic stress response. *Trends Genet. TIG* 33:832–841. doi: 10.1016/j.tig.2017.08.007
- Hovhannisyán, H., Saus, E., Ksiezopolska, E., Hinks Roberts, A. J., Louis, E. J., and Gabaldón, T. (2020). Integrative omics analysis reveals a limited transcriptional shock after yeast interspecies hybridization. *Front. Genet.* 11:404. doi: 10.3389/fgene.2020.00404
- Hug, A. M., and Feldmann, H. (1996). Yeast retrotransposon Ty4: the majority of the rare transcripts lack a U3-R sequence. *Nucleic Acids Res.* 24, 2338–2346. doi: 10.1093/nar/24.12.2338
- Ingolia, N. T. (2014). Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.* 15, 205–213. doi: 10.1038/nrg3645

- Josefsson, C., Dilkes, B., and Comai, L. (2006). Parent-dependent loss of gene silencing during interspecies hybridization. *Curr. Biol.* 16, 1322–1328. doi: 10.1016/j.cub.2006.05.045
- Kawakami, K., Pande, S., Faiola, B., Moore, D. P., Boeke, J. D., Farabaugh, P. J., et al. (1993). A rare tRNA-Arg(CCU) that regulates Ty1 element ribosomal frameshifting is essential for Ty1 retrotransposition in *Saccharomyces cerevisiae*. *Genetics* 135, 309–320. doi: 10.1093/genetics/135.2.309
- Kelleher, E. S., Edelman, N. B., and Barbash, D. A. (2012). *Drosophila* interspecific hybrids phenocopy piRNA-pathway mutants. *PLoS Biol.* 10:e1001428. doi: 10.1371/journal.pbio.1001428
- Kidwell, M. G., and Lisch, D. (1997). Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci. U.S.A.* 94, 7704–7711. doi: 10.1073/pnas.94.15.7704
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4
- Kim, J. M., Vanguri, S., Boeke, J. D., Gabriel, A., and Voytas, D. F. (1998). Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res.* 8, 464–478. doi: 10.1101/gr.8.5.464
- Labrador, M., Farre, M., Utzet, F., and Fontdevila, A. (1999). Interspecific hybridization increases transposition rates of Osvaldo. *Mol. Biol. Evol.* 16, 931–937. doi: 10.1093/oxfordjournals.molbev.a026182
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Laporte, M., Le Luyer, J., Rougeux, C., Dion-Côté, A.-M., Krick, M., and Bernatchez, L. (2019). DNA methylation reprogramming, TE derepression, and postzygotic isolation of nascent animal species. *Sci. Adv.* 5:eaaw1644. doi: 10.1126/sciadv.aaw1644
- Leinonen, R., Sugawara, H., Shumway, M., and International Nucleotide Sequence Database Collaboration (2011). The sequence read archive. *Nucleic Acids Res.* 39, D19–21. doi: 10.1093/nar/gkq1019
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma. Oxf. Engl.* 30, 923–930. doi: 10.1093/bioinformatics/btt656
- Lisch, D. (2013). How important are transposons for plant evolution? *Nat. Rev. Genet.* 14, 49–61. doi: 10.1038/nrg3374
- Lopez-Maestre, H., Carnelossi, E. A. G., Lacroix, V., Burlet, N., Mugat, B., Chambeyron, S., et al. (2017). Identification of misexpressed genetic elements in hybrids between *Drosophila*-related species. *Sci. Rep.* 7:40618. doi: 10.1038/srep40618
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Luning Prak, E. T., and Kazazian, H. H. (2000). Mobile elements and the human genome. *Nat. Rev. Genet.* 1, 134–144. doi: 10.1038/35038572
- Luo, S., and Lu, J. (2017). Silencing of transposable elements by piRNAs in *Drosophila*: an evolutionary perspective. *Genom. Proteom. Bioinform.* 15, 164–176. doi: 10.1016/j.gpb.2017.01.006
- Ma, F., Fuqua, B. K., Hasin, Y., Yukhtman, C., Vulpe, C. D., Lusic, A. J., et al. (2019). A comparison between whole transcript and 3' RNA sequencing methods using Kapa and Lexogen library preparation methods. *BMC Genomics* 20:9. doi: 10.1186/s12864-018-5393-3
- Maheshwari, S., and Barbash, D. A. (2011). The genetics of hybrid incompatibilities. *Annu. Rev. Genet.* 45, 331–355. doi: 10.1146/annurev-genet-110410-132514
- Marsit, S., Hénault, M., Charron, G., Fijarczyk, A., and Landry, C. R. (2021). The neutral rate of whole-genome duplication varies among yeast species and their hybrids. *Nat. Commun.* 12:3126. doi: 10.1038/s41467-021-23231-8
- McClintock, B. (1984). The significance of responses of the genome to challenge. *Science* 226, 792–801. doi: 10.1126/science.15739260
- McManus, C. J., May, G. E., Spealman, P., and Shteyman, A. (2014). Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. *Genome Res.* 24, 422–430. doi: 10.1101/gr.164996.113
- Merel, V., Boulesteix, M., Fablet, M., and Vieira, C. (2020). Transposable elements in *Drosophila*. *Mob. Dna* 11:23. doi: 10.1186/s13100-020-00213-z
- Meyers, B. C., Tingey, S. V., and Morgante, M. (2001). Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. *Genome Res.* 11, 1660–1676. doi: 10.1101/gr.188201
- Morgan, M., Anders, S., Lawrence, M., Abouyou, P., Pagès, H., and Gentleman, R. (2009). ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* 25, 2607–2608. doi: 10.1093/bioinformatics/btp450
- Morillon, A., Bénard, L., Springer, M., and Lesage, P. (2002). Differential effects of chromatin and Gcn4 on the 50-fold range of expression among individual yeast Ty1 retrotransposons. *Mol. Cell. Biol.* 22, 2078–2088. doi: 10.1128/MCB.22.7.2078-2088.2002
- Muller, H. J. (1939). Reversibility in evolution considered from the standpoint of genetics. *Biol. Rev.* 14, 261–280. doi: 10.1111/j.1469-185X.1939.tb00934.x
- Naito, K., Zhang, F., Tsukiyama, T., Saito, H., Hancock, C. N., Richardson, A. O., et al. (2009). Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature* 461, 1130–U232. doi: 10.1038/nature08479
- Neuvéglise, C., Feldmann, H., Bon, E., Gaillardin, C., and Casaregola, S. (2002). Genomic evolution of the long terminal repeat retrotransposons in hemiascomycetous yeasts. *Genome Res.* 12, 930–943. doi: 10.1101/gr.219202
- Niu, X.-M., Xu, Y.-C., Li, Z.-W., Bian, Y.-T., Hou, X.-H., Chen, J.-F., et al. (2019). Transposable elements drive rapid phenotypic variation in *Capsella rubella*. *Proc. Natl. Acad. Sci. U.S.A.* 116, 6908–6913. doi: 10.1073/pnas.1811498116
- O'Neill, R. J. W., O'Neill, M. J., and Graves, J. A. M. (1998). Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* 393, 68–72. doi: 10.1038/29985
- Orgel, L. E., and Crick, F. H. C. (1980). Selfish DNA: the ultimate parasite. *Nature* 284, 604–607. doi: 10.1038/284604a0
- Paquin, C. E., and Williamson, V. M. (1984). Temperature effects on the rate of ty transposition. *Science* 226, 53–55. doi: 10.1126/science.226.4670.53
- Perteau, G., and Perteau, M. (2020). GFF Utilities: GffRead and GffCompare. *F1000Research* 9:304. doi: 10.12688/f1000research.23297.1
- Platt, R. N., Vandeweghe, M. W., and Ray, D. A. (2018). Mammalian transposable elements and their impacts on genome evolution. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 26, 25–43. doi: 10.1007/s10577-017-9570-z
- Renaut, S., Rowe, H. C., Ungerer, M. C., and Rieseberg, L. H. (2014). Genomics of homoploid hybrid speciation: diversity and transcriptional activity of long terminal repeat retrotransposons in hybrid sunflowers. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20130345. doi: 10.1098/rstb.2013.0345
- Rowley, P. A. (2017). The frenemies within: viruses, retrotransposons and plasmids that naturally infect *Saccharomyces* yeasts: viruses and virus-like elements of *Saccharomyces* yeasts. *Yeast* 34, 279–292. doi: 10.1002/yea.3234
- Runemark, A., Vallejo-Marin, M., and Meier, J. I. (2019). Eukaryote hybrid genomes. *PLoS Genet.* 15:e1008404. doi: 10.1371/journal.pgen.1008404
- Saha, A., Mitchell, J. A., Nishida, Y., Hildreth, J. E., Ariberre, J. A., Gilbert, W. V., et al. (2015). A trans-dominant form of gag restricts Ty1 retrotransposition and mediates copy number control. *J. Virol.* 89, 3922–3938. doi: 10.1128/JVI.03060-14
- Salvadó, Z., Arroyo-López, F. N., Guillamón, J. M., Salazar, G., Querol, A., and Barrio, E. (2011). Temperature adaptation markedly determines evolution within the genus *Saccharomyces*. *Appl. Environ. Microbiol.* 77, 2292–2302. doi: 10.1128/AEM.01861-10
- Sandmeyer, S., Patterson, K., and Bilanchone, V. (2015). Ty3, a position-specific retrotransposon in budding yeast. *Microbiol. Spectr.* 3:MDNA3-0057–2014. doi: 10.1128/9781555819217.ch42
- Scannell, D. R., Zill, O. A., Rokas, A., Payen, C., Dunham, M. J., Eisen, M. B., et al. (2011). The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *saccharomyces sensu stricto* genus. *G3 Genes Genomes Genet.* 1, 11–25. doi: 10.1534/g3.111.000273
- Schraiber, J. G., Mostovoy, Y., Hsu, T. Y., and Brem, R. B. (2013). Inferring evolutionary histories of pathway regulation from transcriptional profiling data. *PLoS Comput. Biol.* 9:e1003255. doi: 10.1371/journal.pcbi.1003255
- Servant, G., Pinson, B., Tchalikian-Cosson, A., Couplier, F., Lemoine, S., Pennetier, C., et al. (2012). Tye7 regulates yeast Ty1 retrotransposon sense and antisense

- transcription in response to adenylic nucleotides stress. *Nucleic Acids Res.* 40, 5271–5282. doi: 10.1093/nar/gks166
- Sigman, M. J., and Slotkin, R. K. (2016). The first rule of plant transposable element silencing: location, location, location. *Plant Cell* 28, 304–313. doi: 10.1105/tpc.15.00869
- Slotkin, R. K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* 8, 272–285. doi: 10.1038/nrg2072
- Smit, A. F. A., Hubley, R., and Green, P. (2015). *RepeatMasker Open-1.0*. Available online at: <http://www.repeatmasker.org> (accessed June 13, 2021).
- Smukowski Heil, C., Patterson, K., Hickey, A. S.-M., Alcantara, E., and Dunham, M. J. (2021). Transposable element mobilization in interspecific yeast hybrids. *Genome Biol. Evol.* 13. doi: 10.1093/gbe/evab033
- Steensels, J., Gallone, B., and Verstrepen, K. J. (2021). Interspecific hybridization as a driver of fungal evolution and adaptation. *Nat. Rev. Microbiol.* 19, 485–500. doi: 10.1038/s41579-021-00537-4
- Tandonnet, S., and Torres, T. T. (2017). Traditional versus 3' RNA-seq in a non-model species. *Genomics Data* 11, 9–16. doi: 10.1016/j.gdata.2016.11.002
- Tirosh, I., Reikhav, S., Levy, A. A., and Barkai, N. (2009). A yeast hybrid provides insight into the evolution of gene expression regulation. *Science*. 324, 659–662. doi: 10.1126/science.1169766
- Van Rossum, G., and Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.
- Yoon, O. K., and Brem, R. B. (2010). Noncanonical transcript forms in yeast and their regulation during environmental stress. *RNA* 16, 1256–1267. doi: 10.1261/rna.2038810
- Yue, J.-X., Li, J., Aigrain, L., Hallin, J., Persson, K., Oliver, K., et al. (2017). Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat. Genet.* 49, 913–924. doi: 10.1038/ng.3847
- Zhu, A., Ibrahim, J. G., and Love, M. I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinforma. Oxf. Engl.* 35, 2084–2092. doi: 10.1093/bioinformatics/bty895

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Drouin, Hénault, Hallin and Landry. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.