Developmental Biology 454 (2019) 21–28

Contents lists available at ScienceDirect

Developmental Biology

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

Short Communication

Induction of interferon-stimulated genes and cellular stress pathways by morpholinos in zebrafish

Jason K.H. Lai [a,1,2], Kristina K. Gagalova [a,1,3], Carsten Kuenne [b], Mohamed A. El-Brolosy [a], Didier Y.R. Stainier [a,*]

[a] Max Planck Institute for Heart and Lung Research, Department of Developmental Genetics, 61231 Bad Nauheim, Germany
[b] ECCPS Bioinformatics and Deep Sequencing Platform, Max Planck Institute for Heart and Lung Research, 61231 Bad Nauheim, Germany

ARTICLE INFO

Keywords:
Morpholino
Zebrafish
ISG
p53
Knockdown
Mutant

ABSTRACT

The phenotypes caused by morpholino-mediated interference of gene function in zebrafish are often not observed in the corresponding mutant(s). We took advantage of the availability of a relatively large collection of transcriptomic datasets to identify common signatures that characterize morpholino-injected animals (morphants). In addition to the previously reported activation of *tp53* expression, we observed increased expression of the interferon-stimulated genes (ISGs), *isg15* and *isg20*, the cell death pathway gene *casp8*, and other cellular stress response genes including *phlda3*, *mdm2* and *gadd45aa*. Studies involving segmentation stage embryos were more likely to show upregulation of these genes. We also found that the expression of these genes could be upregulated by increasing doses of an *egfl7* morpholino, or even high doses of the standard control morpholino. Thus, these data show that morpholinos can induce the expression of ISGs in zebrafish embryos and further our understanding of morpholino effects.

1. Introduction

Morpholinos are anti-sense reagents capable of knocking down gene function by inhibiting post-transcriptional regulation of RNAs (Summerton, 1999). They are synthetic nucleic acid analogs which can base-pair with complementary RNA molecules to inhibit their processing. Morpholinos inhibit mRNA function by blocking their splicing or translation (Nasevicius and Ekker, 2000; Draper et al., 2001), and they can inhibit miRNA function as well (Choi et al., 2007).

The ease of morpholino use in zebrafish and its effectiveness in gene knockdown enabled scientists to transiently inhibit gene function to study various biological processes (Nasevicius and Ekker, 2000). Initial testing of morpholinos indicated that morphants could largely phenocopy known mutants (Ekker, 2000; Nasevicius and Ekker, 2000). However, the examples were restricted to genes with well characterized mutant phenotypes. With the recent advent of gene editing, scientists have now been able to test more extensively whether reported morphant phenotypes were similar to those observed in mutants, and they have found substantial discrepancy (Schulte-Merker and Stainier, 2014; Stainier et al., 2015). For example, a study by Kok et al. (2015) generated mutations in more than 20 genes and observed that most mutants appear to develop normally while the corresponding morphants for ten of these genes display a developmental phenotype, suggesting additional effects introduced by morpholinos (Kok et al., 2015). This hypothesis was supported by the observation that the hydrocephaly phenotype in *megamind* morphants could be reproduced by injecting a *megamind* morpholino into animals devoid of the *megamind* locus, albeit at a very high dose (20 ng) (Kok et al., 2015). In another study, 5 ng of a *stat3* morpholino further aggravated the convergence extension phenotype in *MZstat3* mutants (Liu et al., 2017). On the other hand, Rossi, Kontarakis and colleagues proposed that genetic compensation in mutants but not in morphants could help explain this phenotypic discrepancy (Rossi et al., 2015).

One well characterized morpholino off-target effect is the activation of the p53 signaling pathway (Robu et al., 2007). Additionally, a recent study in *Xenopus* found that morpholinos could activate innate immunity genes including *c3ar1*, *tlr5*, *il1b*, *il1r2* and *il8*, as well as induce off-target

* Corresponding author.
E-mail address: didier.stainier@mpi-bn.mpg.de (D.Y.R. Stainier).
[1] These authors contributed equally to this work.
[2] Present Address: Mechanobiology Institute Singapore, National University of Singapore, 117411 Singapore.
[3] Present Address: Canada Michael Smith Genome Sciences Center, Vancouver BC V5Z 4S6, Canada.

https://doi.org/10.1016/j.ydbio.2019.06.008
Received 19 December 2018; Received in revised form 11 June 2019; Accepted 12 June 2019
Available online 12 June 2019

mis-splicing of pre-mRNA molecules (Gentsch et al., 2018). The latter observation was also documented in zebrafish morphants (Joris et al., 2017). These studies raise questions about the effects of morpholinos in a cell, and the need to fully understand them for future applications. In this report, we sought to further characterize the effects of morpholinos by studying various microarray gene expression datasets from studies in zebrafish obtained from the Gene Expression Omnibus (GEO) database. We found that the expression of interferon-stimulated genes, ISGs (*isg15* and *isg20*), and other cellular stress pathway genes (*tp53*, *phlda3, casp8, gadd45aa* and *mdm2*) was upregulated in morphants, especially in segmentation stage embryos. Furthermore, the expression of these genes was upregulated in response to increasing doses of an *egfl7* morpholino, or high doses of the standard control morpholino. In contrast, the commonly used *vegfaa* morpholino did not induce the expression of these genes. Altogether, these data show that morpholinos can induce the expression of ISGs in zebrafish embryos, and indicate the need for additional studies to fully understand morpholino effects.

2. Materials and Methods

2.1. Ethics statement

All zebrafish husbandry was performed under standard conditions in accordance with institutional (MPG) and national ethical and animal welfare guidelines.

2.2. Microarray data collection and curation

Microarray gene expression arrays of various zebrafish morpholino studies were obtained from the GEO database. As of the 19th of March 2015, keyword search for "zebrafish morpholino" or "zebrafish knockdown" yielded 63 unique *Danio rerio* data series. The embedded metadata and raw intensities for all studies were retrieved with the R package GEOquery (Davis and Meltzer, 2007). We manually curated all datasets utilizing GPL1319, a platform manufactured by Affymetrix, to annotate the publication, morpholino sequence, morpholino dose, sample source and developmental stage. Datasets from Agilent arrays were set aside for meta-analysis. All annotated information can be found in File S1 – Section 1.

2.3. Affymetrix microarray data normalization and processing

8 GSE data series were rejected because they were not morpholino studies, were a duplicate, or were not linked to a publication. The remaining data series were pooled together for normalization. The raw data were read with the ReadAffy function from affy package in R (Gautier et al., 2004), followed by implementation of Single Channel Array Normalization (SCAN) (Piccolo et al., 2012) to normalize the pooled dataset. Default parameters were used. Integrating microarray datasets from various studies will inevitably introduce batch effects. Thus, we corrected the normalized data for batch effects using ComBat (Johnson et al., 2007), and defined the unique GSE ID as the primary source of batch effects.

The data were then filtered by removing control probes as well as averaging all probesets that interrogate the same gene. We also filtered out genes that had little variation in gene expression across all pooled samples using the varFilter from the geneFilter package. These data were used to generate the Principal Component Analysis (PCA) plots.

Finally, we grouped genes that exhibited high correlation in the microarray dataset using a 0.95 score cutoff. This step was carried out with the findCorrelation from the caret R package (Kuhn, 2008). Details and quality control plots can be found in File S1 – Section 2.

2.4. Feature selection

We tested various models with built-in feature selection on our dataset, and we also tested the final predictive accuracy. These models include: "glmnet", "pam", "gausssprLinear", "svmLinear", "svmPoly", "svmRadial", "nb", "rf", "knn" and "J48". We optimized the parameters for each method through the train function of the caret R package (Kuhn, 2008), using Leave One Out Cross Validation (LOOCV) and the accuracy as the readout scores. The models were further tested against scrambled-labels datasets. Relevant genes were ranked with *t*-test statistic as the linear models performed best. We analyzed this set of genes by Gene Set Enrichment Analysis (GSEA) with GOrilla (Eden et al., 2009) as well as pathway analysis with Ingenuity Pathway Analysis (Qiagen).

The features were shortlisted through recursive feature elimination and backward selection with the rfe caret function. The models were evaluated with 4-folds cross-validation, iterated 100 times. The model accuracy increases logarithmically with the number of features. To avoid overfitting the model, we selected the number of features at the point of diminishing returns, which is 5% below the maximum accuracy score.

2.5. Two-way ANOVA analysis

We conducted a standard two-way ANOVA analysis stratifying for developmental stage and treatment (i.e. morphants and controls). Both variables were converted into categorical variables. To determine developmental stage, we followed ZFIN's guidelines.

2.6. Agilent microarray data processing and analysis

A total of 34 microarray studies were carried out on a variety of Agilent microarray platforms. Among these studies, 12 were rejected due to the same reasons mentioned above for the Affymetrix studies. A further 5 raw datasets were not useable for a standard pipeline because they were uploaded to GEO in a wrong file format or were not readable. One other dataset could not be analyzed as the experimental design was insufficient to run statistical tests. The R package limma (Ritchie et al., 2015) was used to process the raw intensity files as well as downstream analyses.

Briefly, the raw files were read and compiled into a standard "eset" object for a standard normalization pipeline and statistical analyses (Smyth and Speed, 2003). Both dual- and single-channel arrays were first background corrected using the "normexp" method (Ritchie et al., 2007), followed by within-array normalization by the "loess" method to correct for dye biases in dual-channel arrays (Smyth and Speed, 2003). Finally, between-array normalization was performed for both microarray configurations by the "quantile" method (Bolstad et al., 2003). Control probes were then removed prior to applying empirical Bayesian statistical analyses (Smyth, 2005). We extracted the average $\log_2$-transformed fold-change (logFC) and standard errors for all probes interrogating *isg15*, *isg20*, *casp8*, *mdm2*, *gadd45aa* and *tp53*. As some genes were interrogated by multiple probes, we took the probe that detected the maximum absolute logFC.

2.7. Zebrafish

Transgenic line used in this study:
Tg(kdrl:EGFP)^{s843Tg} (Jin et al., 2005).

2.8. Zebrafish micro-injection and real-time PCR (qPCR) assay

The three morpholinos used in this study are *egfl7* (5′-CAGGTGTGTCTGACAGCAGAAAGAG-3′) (Parker et al., 2004), *vegfaa* (5′-GTATCAAATAAACAACCAAGTTCAT-3′) (Nasevicius et al., 2000) and the standard control morpholino (5′-CCTCTTACCTCAGTTACAATTTATA-3′) (Gene Tools). Each morpholino was injected into 1-cell stage *Tg(kdrl:EGFP)* zebrafish embryos (1 nL volume). The morpholino doses used are as follows: *egfl7* - 0.0625, 0.125, 0.25, 0.5, 1 and 2 ng; *vegfaa* - 0.5, 1, 2, 4 and 8 ng; control morpholino - 0.5, 1, 2, 8 and 16 ng. We also injected buffer alone to serve as a blank control. Each sample was

collected at 32 h post-fertilization (hpf). 5 embryos were pooled into one biological replicate, and a total of 4 biological replicates was obtained for each sample. We extracted RNA with TRIzol® (Ambion). The aqueous phase was passed through an RNA purification column (Zymogen) and treated with RQ1 DNase (Promega). Samples were re-extracted by acid phenol:chloroform (Ambion) followed by chloroform back-extraction. Biological replicates with poor RNA quality or insufficient RNA amount were discarded. Nevertheless, each group had at least 3 biological replicates. 150 ng of RNA from each sample was converted to cDNA by High Capacity RNA-to-cDNA Kit (Applied Biosystems). Primer sequences and optimized cycling conditions for *rpl13*, *gapdh*, *isg20*, *isg15*, *casp8*, *phlda3* and *tp53Δ113* can be found in Table S1.

2.9. Statistical analysis of qPCR assays

All statistical analyses were performed using the R statistical software. All C_T values were normalized to *gapdh* and compared to uninjected samples ($-\Delta\Delta C_T$). Morpholino dose was transformed to a log-scale (base 2). A dose-response curve test was applied using the *LL.4* function from the drc package in R to calculate the upper limit, lower limit and slope. Otherwise, linear plots were tested using linear regression to estimate the slope. For plotting the graphs, both the in-built plot function of R as well as the ggplot2 package (Wickham, 2016) were utilized.

2.10. Whole mount in situ hybridization (WISH)

We injected 0.5 ng of *egfl7* or control morpholino into 1-cell stage WT embryos (ABs). Embryos were collected at 32 hpf and fixed with 4% PFA in 1X PBST (1X PBS, 0.1% Tween-20), followed by standard WISH protocol (Thisse and Thisse, 2008). Primers used to generate the *isg20* template to synthesize the *in situ* probe are: 5′-GGAGAATCATGGGAAGTGGA-3′, and 5′-GGCATTGAGGTTGGCAGTAT-3’.

2.11. RNAseq data mining and analysis

We searched for zebrafish morpholino datasets in the GEO database using a combined query ‘((danio rerio[Organism]) AND (“2015/03/19"[Publication Date]: “3000"[Publication Date])) AND morpholino’, returning 41 results. These results were assessed for quality, adapter content and duplication rates with FastQC (Andrews S. 2010, FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc), and filtered for high quality protein coding transcriptome sequencing including replicate samples. The remaining seven experiments (GSE102662, GSE111152, GSE98888, GSE104855, GSE84601, GSE71573, GSE74929) were subjected to Trimmomatic version 0.38 to trim low quality sequence below a mean of Q15 in a window of 5 nucleotides leaving only reads of at least 15 nucleotides (Bolger et al., 2014). Trimmed and filtered reads were aligned versus the Ensembl zebrafish genome version danRer11 (GRCz11) using STAR 2.6.0c with the parameters “–outFilterMismatchNoverLmax 0.1 –alignMatesGapMax 2000 –alignIntronMax 200000 –outFilterMultimapNmax 1” (Dobin et al., 2013). The number of reads aligning to genes was counted with Subread featureCounts 1.6.0 (Liao et al., 2014). Only reads mapping at least partially inside exons were admitted and aggregated per gene. Reads overlapping multiple genes or aligning to multiple regions were excluded. Differentially expressed genes were identified using DESeq2 version 1.18.1 (Love et al., 2014). Genes were classified as significantly differentially expressed at Benjamini-Hochberg corrected $P < 0.05$ and $-0.59 \leq \text{Log2FC} \geq +0.59$.

2.12. JunctionSeq analysis

To investigate which *tp53* isoforms are upregulated in *srsf5a* morphants (Joris et al., 2017), we used the JunctionSeq package (Hartley and Mullikin, 2016) in R to quantify the normalized (sequencing depth) abundance of reads per exon in the zebrafish *tp53* locus.

3. Results

3.1. Integration of affymetrix microarray gene expression datasets

We retrieved 63 microarray gene expression datasets from Gene Expression Omnibus (GEO) by keyword search. The majority of microarrays used were manufactured by Agilent (n = 34) or Affymetrix (n = 19). Duplicated datasets, datasets with no accompanying publications and datasets in incorrect file formats (Agilent, n = 12; Affymetrix, n = 8) were removed from further analyses. We initially focused on the datasets generated using the Affymetrix microarray platform (n = 11), and applied algorithms to eliminate batch effects and other technical noise in order to integrate them into one large dataset. Principal component analysis (PCA) of this dataset revealed that developmental stage and tissue source were the main factors driving the clustering in the principal components (PC) 1 and 2, respectively (Fig. S1).

3.2. Feature selection uncovers activation of ISGs and the p53 signaling pathway

To identify genes differentially expressed between the control group (consisting of uninjected embryos, buffer-injected embryos and control morphants) and morphants, we fed the entire dataset into a supervised machine learning algorithm. The model assigns a higher score to a gene if it performs better at predicting morphants. Among the various models of feature selection, we found that the linear feature selection model tended to perform best (Table S2). We then tested the accuracy of the model and short-listed the optimal number of genes that successfully distinguished the morphants from the control group (see Materials and Methods). We refer to this set of genes as “top performing genes” (Table S3).

We detected *tp53* as one of these top performing genes upregulated in morphants compared to the control group, corroborating previous findings (Robu et al., 2007). Surprisingly, we found another gene, *isg20*, which scored higher than *tp53*, raising the possibility that interferon signaling was also activated in morphants. Furthermore, we found biological processes such as apoptosis and various cell stress responses enriched when further analyzing the top performing genes (Fig. 1).

3.3. ISG and p53 signaling activation in morphants is most pronounced at the segmentation stage

As mentioned above, the samples can be separated by developmental stage along the first PC, as expected by the dynamic nature of gene expression in developing embryos (Mathavan et al., 2005; Vesterlund et al., 2011). Thus, we proceeded to analyze our dataset by stratifying based on developmental stage. Using a two-way ANOVA approach, we identified 157 genes that were significantly ($P < 0.001$) differentially expressed between morphants and uninjected or buffer-injected embryos (collectively referred to as WT), depending on developmental stage (Table S4). Independent of the morpholino target gene, we found that *isg15*, *isg20*, *casp8*, *mdm2*, *gadd45aa*, *phlda3* and *tp53* were upregulated during segmentation stages in morphants compared to WT (Fig. 2). Notably, Robu and colleagues similarly reported the activation of *isg20* (this gene was originally referred to as *ISG20L2*, but the probe sequence given by its accession number now aligns to zebrafish *isg20*), *casp8* and *phlda3* in their microarray assay (Fig 9C in (Robu et al., 2007)). Genes that were differentially expressed at other developmental stages can be found in File S1 – Section 3.

3.4. Analysis of additional microarray and RNAseq datasets confirms the activation of ISGs and the p53 signaling pathway in morphants

The Agilent microarray datasets were composed of different platform configurations which makes data integration challenging and could

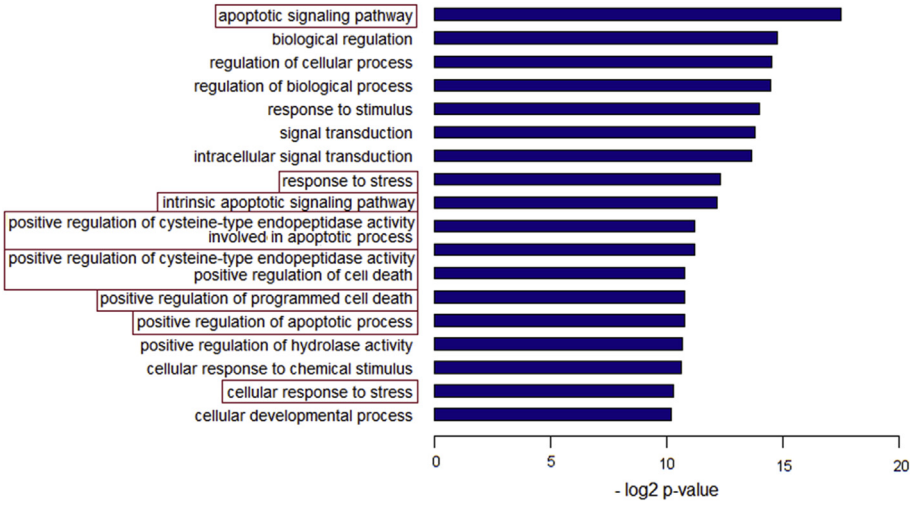

Fig. 1. Gene set enrichment analysis. Biological processes such as apoptosis and response to stress are enriched in the top performing genes (Table S3).

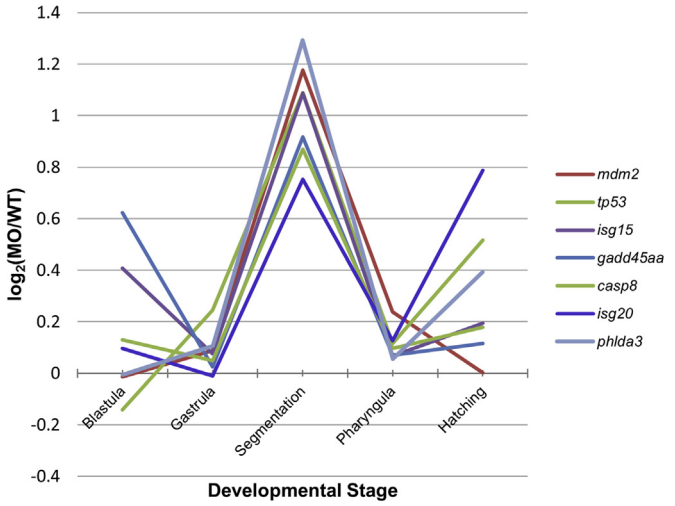

Fig. 2. Relative fold change of gene expression plotted against developmental stage. Genes from the p53 signaling pathway (*tp53*, *mdm2*, *gadd45aa*, *phlda3*) or the caspase cascade (*casp8*), as well as interferon stimulated genes (*isg15*, *isg20*) are upregulated in morphants at the segmentation stage. Morphants (MO) were compared to uninjected or buffer-injected embryos (WT).

potentially introduce artefacts into the post-processed data. Therefore, we individually normalized each Agilent microarray dataset. Some datasets compare morphants to uninjected or buffer-injected embryos (WT in Fig. 3), while other datasets compare morphants to control morphants (controlMO in Fig. 3). Analysis of these datasets reveals that *isg15*, *isg20*, *casp8*, *mdm2*, *gadd45aa* and *tp53* were similarly upregulated in some morphants, and many of these morphants were in fact assessed at segmentation and pharyngula stages (Fig. 3). However, we did not observe a strong correlation between the induction level of these genes and morpholino dose (Fig. S2).

In addition to our observations on the microarray datasets, we found that these genes were also upregulated in published RNAseq datasets (5 of 8 datasets). Specifically, *qkic*, *srsf5a*, *prep1/pbx2/4*, *tut4/7* and *rad21* morphants exhibited increased expression of *isg15*, *isg20*, *casp8*, *mdm2*, *gadd45aa* and *tp53* compared to the control group (data not shown). Notably, *prep1/pbx2/4*, *tut4/7* and *rad21* morphants exhibited upregulation of these genes at 6 hpf (*tut4/7* morphants) or after 10 hpf (*prep1/pbx2/4* and *rad21* morphants) (data not shown).

A previous report found mis-splicing of pre-mRNAs in *srsf5a* zebrafish morphants (Joris et al., 2017), which was also observed in *t/t2 Xenopus* morphants (Gentsch et al., 2018). However, the former did not investigate the induction of ISGs. The *srsf5a* morphant RNAseq dataset mentioned earlier is also derived from the work by Joris et al. (2017). Thus, this particular *srsf5a* morpholino concomitantly activates ISGs and leads to pre-mRNA mis-splicing. Moreover, we found in this dataset that the expression of the *tp53Δ113* isoform (Robu et al., 2007) is specifically upregulated in *srsf5a* morphants compared to standard control morphants (Fig. S3).

3.5. ISG expression is upregulated by increasing doses of some morpholinos

We injected increasing doses of *egfl7*, *vegfaa* and standard control morpholinos, followed by qPCR analysis of *isg20, isg15, tp53Δ113* (Robu et al., 2007)*, phlda3* and *casp8* mRNA levels. We found upregulation of these genes by increasing doses of *egfl7* morpholino, and for some genes observed a sigmoid curve (Fig. 4). In contrast, no dose-dependent upregulation was observed following injection of increasing doses of the commonly used *vegfaa* morpholino (Fig. 4). When we injected 8 or 16 ng of the standard control morpholino, not only did the injected embryos exhibit upregulation of the aforementioned genes (Fig. 4), but they also exhibited a delay in angiogenesis (Fig. S4). In addition, we performed whole-mount *in situ* hybridization in *egfl7* morphants and observed that *isg20* appeared to be upregulated throughout the embryo (Fig. S5), and not restricted to the endothelium which specifically expresses *egfl7* (Parker et al., 2004).

4. Discussion

Morpholinos are widely used as a tool to inactivate the function of genes and non-coding RNAs. However, we do not yet fully understand the responses of a cell to morpholinos. Indeed, several reported side-effects of morpholinos, including *tp53Δ113* upregulation (Robu et al., 2007), mis-splicing of nascent transcripts (Joris et al., 2017; Gentsch et al., 2018), and induction of innate immune response genes (Gentsch et al., 2018), could complicate the interpretation of morpholino studies. Similarly, our current study identified ISGs as upregulated genes across multiple zebrafish morphants as well as zebrafish embryos injected with an excess of the standard control morpholino.

In addition to the p53 signaling pathway (Robu et al., 2007), we identified the activation of ISGs by morpholino injections. *isg20* expression was previously shown to be induced by the presence of dsRNA (Espert et al., 2004; Horio et al., 2004). dsRNAs can be recognized by receptors including Toll-Like Receptor 3 (TLR3), RIG1 and MDA5, which trigger downstream interferon type-I (IFN–I) signaling (Alexopoulou et al., 2001; Kang et al., 2002; Yoneyama et al., 2004). Interestingly, we

GSE	Morpholino	Control	Developmental Stage	*isg20*	*isg15*	*tp53*	*casp8*	*phlda3*	*gadd45 aa*	*mdm2*
GSE21539	*ovol1*	WT	Segmentation							
GSE27569	*esco2*	WT	Segmentation							
GSE27569	*esco2*	WT	Pharyngula							
GSE34930	*ptpn6*	WT	Pharyngula							
GSE37332	*mitfa*	WT	Hatching							
GSE50376	*nr3c1*	controlMO	Segmentation							
GSE50376	*nr3c1*	controlMO	Pharyngula							
GSE13157	*esr2a*	WT	Hatching							
GSE13158	*esr2a*	WT	Hatching							
GSE38441	*grna*	WT	Segmentation							
GSE38441	*grna*	WT	Segmentation							
GSE38441	*grna*	WT	Pharyngula							
GSE38441	*grna*	WT	Hatching							
GSE42070	*pcsk7+tp53*	controlMO	Gastrula							
GSE42070	*pcsk7+tp53*	controlMO	Segmentation							
GSE64426	*tbx5a*	controlMO	Pharyngula							
GSE64426	*tbx5a*	controlMO	Segmentation							
GSE61625c	*dio3b*	controlMO	Hatching							
GSE61625a	*dio3b*	controlMO	Hatching							
GSE61625b	*dio3b*	controlMO	Hatching							
GSE61625a	*dio1/2*	controlMO	Hatching							
GSE61625b	*dio1/2*	controlMO	Hatching							
GSE61625c	*dio1/2*	controlMO	Hatching							
GSE17773	*arrb1/2*	controlMO	Segmentation							
GSE57836	*slc2a2*	controlMO	Hatching							
GSE16740	*tnnt2*	controlMO	Pharyngula							
GSE16740	*tnnt2*	controlMO	Pharyngula							
GSE16740	*tnnt2*	controlMO	Hatching							
GSE54754	*gba*	controlMO	Pharyngula							
GSE51541	*atg5*	WT	Hatching							
GSE51541	*becn1*	WT	Hatching							
GSE51541	*atg7*	WT	Hatching							
GSE18830	*sox2/3/19a/19b*	WT	Blastula							
GSE18830	*sox2/3/19a/19b*	WT	Gastrula							
GSE18830	*sox2/3/19a/19b*	WT	Gastrula							
GSE32594	*nfe2l2a*	controlMO	Hatching							
GSE32594	*nfe2l2b*	controlMO	Hatching							
GSE32594	*nfe2l2a+nfe2l2b*	controlMO	Hatching							
GSE19206	*spi1b*	WT	Pharyngula							
GSE45011	*klf2a*	controlMO	Blastula							
GSE45011	*klf2b*	controlMO	Blastula							
GSE52229	*ahrra*	WT	Hatching							
GSE58683	*aplnra/b*	WT	Blastula							
GSE52229	*ahrrb*	WT	Hatching							
GSE12012	*miR-126_MO1*	WT	Pharyngula							
GSE12012	*miR-126_MO2*	WT	Pharyngula							
GSE46844	*miR-34b*	controlMO	Hatching							
GSE34508	*slc2a10*	controlMO	Pharyngula							
GSE32914	*sox31*	WT	Blastula							
GSE20179	*tnnt2*	controlMO	Larval							
GSE63360	*nr3c1_MO1*	controlMO	Pharyngula							
GSE63360	*nr3c1_MO2*	controlMO	Pharyngula							

Fig. 3. Analysis of *isg20, isg15, tp53, casp8, phlda3, gadd45aa* and *mdm2* mRNA levels from Agilent microarray datasets. Differential expression of *isg20, isg15, tp53, casp8, phlda3, gadd45aa* and *mdm2* between the "Morpholino" group and "Control" group. Red indicates significant upregulation ($P < 0.05$); green indicates significant downregulation ($P < 0.05$). The darker shade of red/green indicates logFC >1. WT – buffer-injected and uninjected embryos; controlMO – control morphants (standard control morpholino and mismatch control morpholinos).

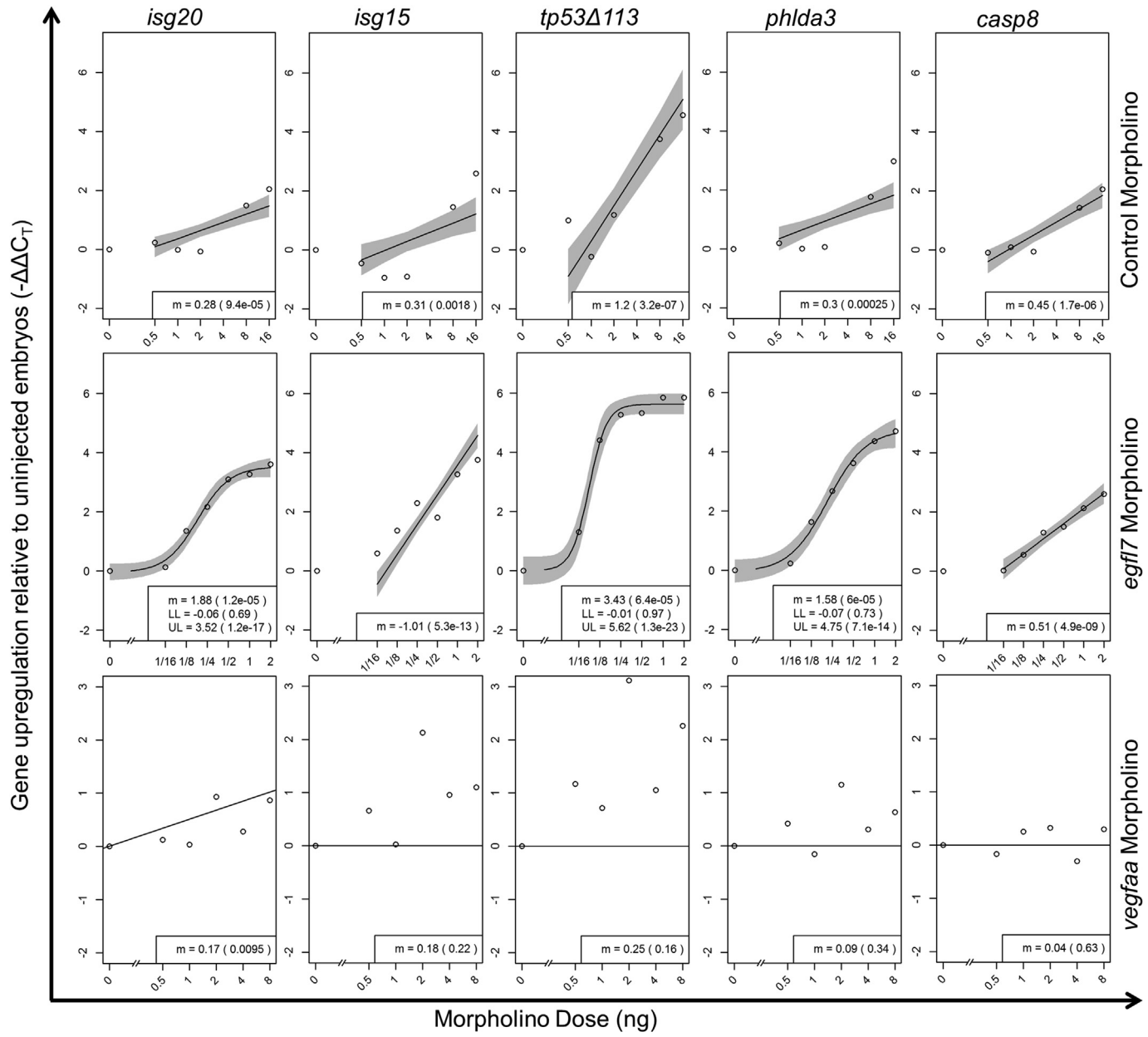

Fig. 4. ISGs and stress response genes are upregulated in a morpholino dose-dependent manner. qPCR analysis of *isg20, isg15, tp53Δ113, phlda3* and *casp8* mRNA levels in 32 hpf zebrafish embryos injected with increasing doses of the standard control morpholino, an *egfl7* morpholino or the commonly used *vegfaa* morpholino. N = 4 biological replicates per morpholino per dose. m - slope; LL - lower limit of sigmoid curve; UL - upper limit of sigmoid curve. *P*-values for individual parameters in parentheses. Grey shaded region reflects the 95% confidence band of the curve.

also detected activation of the IL-8 signaling pathway in our top upregulated genes (data not shown), suggesting the activation of TLRs and IFN-I, possibly by morpholino-RNA duplexes.

Thus far, two studies have reported that some morpholinos (including the standard control morpholino) cause mis-splicing of pre-mRNAs in zebrafish and *Xenopus*. This aberrant splicing appears to arise from the binding of morpholinos to RNA transcripts, and ~11 consecutive base-pairing is sufficient for it (Joris et al., 2017). While some morpholinos induce ISGs in zebrafish, we could not ascertain the same in *Xenopus* as the orthologs of *isg15* and *isg20* are not well annotated in the current *Xenopus* genome version. Although the IL-8 signaling pathway is enriched in the top performing genes in our study, none of the upregulated innate immune response genes reported by Gentsch et al. (2018) are present in our list. Whether both ISGs and innate immune response genes are part of a network of responses to morpholinos needs further investigation. In addition, it would be interesting to study how aberrant splicing events and induction of ISGs or innate immune response genes relate to each other. Nevertheless, these morpholino side-effects can confuse the data and their interpretation, and thus morpholino studies should be thoroughly controlled (see latest guidelines in (Stainier et al., 2017)).

Large-scale gene expression studies have been used to identify gene expression signatures that correlate with tumor progression, and with thorough validation, some of these signatures have become diagnostic markers and/or drug targets (Calin and Croce, 2006; Sotiriou and Piccart, 2007). Moreover, public databases, such as Oncomine (Rhodes et al., 2004), have been established to mine out markers across multiple studies. We applied a similar strategy to multiple gene expression datasets from zebrafish morpholino studies in order to identify gene expression signatures that are common to morpholino-injected zebrafish, and found that several ISGs and stress response genes are upregulated in

morphants. Additional work will be needed to understand the underlying biology of these genes and how they might affect morpholino experiments. A better understanding of morpholino effects should help scientists make informed decisions when conducting morpholino experiments, and also facilitate the interpretation of the data arising from these experiments.

Funding

Research in the Stainier lab is supported in part by the Max Planck Society, the European Union (ERC), the Leducq foundation, and the German Research Foundation (DFG).

Acknowledgements

We would like to thank Sven Reischauer, Oliver Stone and Hans-Martin Maischein for their insights regarding morpholinos, Mario Looso and his team for feedback and suggestions on bioinformatics, Pedro Moura for insightful discussions regarding dsRNA induction of *IFN-I*, and Andrea Rossi, Zacharias Kontarakis and Albert Wang for comments on the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ydbio.2019.06.008.

References

Ekker, S.C., 2000. Morphants: a new systematic vertebrate functional genomics approach. Yeast 17, 302–306. https://doi.org/10.1002/1097-0061(200012)17:4<302::AID-YEA53>3.0.CO;2-#.

Alexopoulou, L., Holt, A.C., Medzhitov, R., Flavell, R.A., 2001. Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. Nature 413, 732.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bolstad, B.M., Irizarry, R.A., Astrand, M., Speed, T.P., 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19, 185–193.

Calin, G.A., Croce, C.M., 2006. MicroRNA signatures in human cancers. Nat. Rev. Canc. 6, 857–866. https://doi.org/10.1038/nrc1997.

Choi, W.-Y., Giraldez, A.J., Schier, A.F., 2007. Target protectors reveal dampening and balancing of Nodal agonist and antagonist by miR-430. Science 318, 271–274. https://doi.org/10.1126/science.1147535.

Davis, S., Meltzer, P.S., 2007. GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor. Bioinformatics 23, 1846–1847. https://doi.org/10.1093/bioinformatics/btm254.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Draper, B.W., Morcos, P.A., Kimmel, C.B., 2001. Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholino oligos: a quantifiable method for gene knockdown. Genesis 30, 154–156. https://doi.org/10.1002/gene.1053.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., Yakhini, Z., 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinf. 10, 48. https://doi.org/10.1186/1471-2105-10-48.

Espert, L., Rey, C., Gonzalez, L., Degols, G., Chelbi-Alix, M.K., Mechti, N., Gongora, C., 2004. The exonuclease ISG20 is directly induced by synthetic dsRNA via NF-κB and IRF1 activation. Oncogene 23, 4636–4640. https://doi.org/10.1038/sj.onc.1207586.

Gautier, L., Cope, L., Bolstad, B.M., Irizarry, R.A., 2004. affy–analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20, 307–315. https://doi.org/10.1093/bioinformatics/btg405.

Gentsch, G.E., Spruce, T., Monteiro, R.S., Owens, N.D.L., Martin, S.R., Smith, J.C., 2018. Innate immune response and off-target mis-splicing are common morpholino-induced side effects in Xenopus. Dev. Cell 44, 597–610 e10. https://doi.org/10.1016/j.devcel.2018.01.022.

Hartley, S.W., Mullikin, J.C., 2016. Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. Nucleic Acids Res. 44, e127. https://doi.org/10.1093/nar/gkw501.

Horio, T., Murai, M., Inoue, T., Hamasaki, T., Tanaka, T., Ohgi, T., 2004. Crystal structure of human ISG20, an interferon-induced antiviral ribonuclease. FEBS Lett. 577, 111–116. https://doi.org/10.1016/j.febslet.2004.09.074.

Jin, S.-W.S.-W., Beis, D., Mitchell, T., Chen, J.-N., Stainier, D.Y.R., 2005. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. Development 132, 5199–5209. https://doi.org/10.1242/dev.02087.

Johnson, W.E., Li, C., Rabinovic, A., 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8, 118–127. https://doi.org/10.1093/biostatistics/kxj037.

Joris, M., Schloesser, M., Baurain, D., Hanikenne, M., Muller, M., Motte, P., 2017. Number of inadvertent RNA targets for morpholino knockdown in Danio rerio is largely underestimated: evidence from the study of Ser/Arg-rich splicing factors. Nucleic Acids Res. 45, 9547–9557. https://doi.org/10.1093/nar/gkx638.

Kang, D., Gopalkrishnan, R.V., Wu, Q., Jankowsky, E., Pyle, A.M., Fisher, P.B., 2002. mda-5: an interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties. Proc. Natl. Acad. Sci. U.S.A. 99, 637–642. https://doi.org/10.1073/pnas.022637199.

Kok, F.O., Shin, M., Ni, C.W., Gupta, A., Grosse, A.S., vanImpel, A., Kirchmaier, B.C., Peterson-Maduro, J., Kourkoulis, G., Male, I., DeSantis, D.F., Sheppard-Tindell, S., Ebarasi, L., Betsholtz, C., Schulte-Merker, S., Wolfe, S.A., Lawson, N.D., 2015. Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. Dev. Cell 32, 97–108. https://doi.org/10.1016/j.devcel.2014.11.018.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28, 1–27. https://doi.org/10.18637/jss.v028.i05.

Liao, Y., Smyth, G.K., Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

Liu, Y., Sepich, D.S., Solnica-Krezel, L., 2017. Stat3/Cdc25a-dependent cell proliferation promotes embryonic axis extension during zebrafish gastrulation. PLoS Genet. 13, e1006564. https://doi.org/10.1371/journal.pgen.1006564.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8.

Mathavan, S., Lee, S.G.P., Mak, A., Miller, L.D., Murthy, K.R.K., Govindarajan, K.R., Tong, Y., Wu, Y.L., Lam, S.H., Yang, H., Ruan, Y., Korzh, V., Gong, Z., Liu, E.T., Lufkin, T., 2005. Transcriptome analysis of zebrafish embryogenesis using microarrays. PLoS Genet. 1, 260–276. https://doi.org/10.1371/journal.pgen.0010029.

Nasevicius, A., Ekker, S.C., 2000. Effective targeted gene "knockdown" in zebrafish. Nat. Genet. 26, 216–220. https://doi.org/10.1038/79951.

Nasevicius, A., Larson, J., Ekker, S.C., 2000. Distinct requirements for zebrafish angiogenesis revealed by a VEGF-A morphant. Yeast 17, 294–301. https://doi.org/10.1002/1097-0061(200012)17:4<294::AID-YEA54>3.0.CO;2-5.

Parker, L.H., Schmidt, M., Jin, S.-W., Gray, A.M., Beis, D., Pham, T., Frantz, G., Palmieri, S., Hillan, K., Stainier, D.Y.R., De Sauvage, F.J., Ye, W., 2004. The endothelial-cell-derived secreted factor Egfl7 regulates vascular tube formation. Nature 428, 754–758. https://doi.org/10.1038/nature02416.

Piccolo, S.R., Sun, Y., Campbell, J.D., Lenburg, M.E., Bild, A.H., Johnson, W.E., 2012. A single-sample microarray normalization method to facilitate personalized-medicine workflows. Genomics 100, 337–344. https://doi.org/10.1016/j.ygeno.2012.08.003.

Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pander, A., Chinnaiyan, A.M., 2004. ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia 6, 1–6. https://doi.org/10.1016/S1476-5586(04)80047-2.

Ritchie, M.E., Silver, J., Oshlack, A., Holmes, M., Diyagama, D., Holloway, A., Smyth, G.K., 2007. A comparison of background correction methods for two-colour microarrays. Bioinformatics 23, 2700–2707. https://doi.org/10.1093/bioinformatics/btm412.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K., 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47. https://doi.org/10.1093/nar/gkv007.

Robu, M.E., Larson, J.D., Nasevicius, A., Beiraghi, S., Brenner, C., Farber, S.A., Ekker, S.C., 2007. P53 activation by knockdown technologies. PLoS Genet. 3, 787–801. https://doi.org/10.1371/journal.pgen.0030078.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M., Stainier, D.Y.R., 2015. Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature 524, 230–233. https://doi.org/10.1038/nature14580.

Schulte-Merker, S., Stainier, D.Y.R., 2014. Out with the old, in with the new: reassessing morpholino knockdowns in light of genome editing technology. Development 141, 3103–3104. https://doi.org/10.1242/dev.112003.

Smyth, G.K., 2005. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. 3, 1–25. https://doi.org/10.2202/1544-6115.1027.

Smyth, G.K., Speed, T., 2003. Normalization of cDNA microarray data. Methods 31, 265–273. https://doi.org/10.1016/S1046-2023(03)00155-5.

Sotiriou, C., Piccart, M.J., 2007. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? Nat. Rev. Canc. 7, 545–553. https://doi.org/10.1038/nrc2173.

Stainier, D.Y.R., Kontarakis, Z., Rossi, A., 2015. Making sense of anti-sense data. Dev. Cell 32, 7–8. https://doi.org/10.1016/j.devcel.2014.12.012.

Stainier, D.Y.R., Raz, E., Lawson, N.D., Ekker, S.C., Burdine, R.D., Eisen, J.S., Ingham, P.W., Schulte-Merker, S., Yelon, D., Weinstein, B.M., Mullins, M.C., Wilson, S.W., Ramakrishnan, L., Amacher, S.L., Neuhauss, S.C.F., Meng, A., Mochizuki, N., Panula, P., Moens, C.B., 2017. Guidelines for morpholino use in zebrafish. PLoS Genet. 13, 6–10. https://doi.org/10.1371/journal.pgen.1007000.

Summerton, J., 1999. Morpholino antisense oligomers: the case for an RNase H-independent structural type. Biochim. Biophys. Acta 1489, 141–158.

Thisse, C., Thisse, B., 2008. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat. Protoc. 3, 59–69. https://doi.org/10.1038/nprot.2007.514.

Vesterlund, L., Jiao, H., Unneberg, P., Hovatta, O., Kere, J., 2011. The zebrafish transcriptome during early development. BMC Dev. Biol. 11, 30. https://doi.org/10.1186/1471-213X-11-30.
Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.
Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., Fujita, T., 2004. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. Nat. Immunol. 5, 730–737. https://doi.org/10.1038/ni1087.