Comparative analysis of gene expression in virulent and attenuated strains of infectious bronchitis virus at sub-codon resolution

Adam Dinan[1], Sarah Keep[2], Erica Bickerton[2], Paul Britton[2], Andrew E. Firth[1] and Ian Brierley[1†]

[1]Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QP, U.K. [2]The Pirbright Institute, Woking, Surrey, GU24 0NF, U.K.

Running Title: Ribosome profiling of coronavirus IBV infection.

Abstract word count: 250

Importance word count: 140

[†]Corresponding author: IB: Tel: +44 1223 336914, Fax: +44 1223 336926

Electronic Mail: ib103@cam.ac.uk

Keywords: avian coronavirus, ribosome profiling, RNA virus, RNASeq, translation, differential gene expression

ABSTRACT

Like all coronaviruses, avian infectious bronchitis virus (IBV) possesses a long, single-stranded, positive-sense RNA genome (~27 kb) and has a complex replication strategy that includes the production of a nested set of sub-genomic mRNAs (sgmRNAs). Here, we used RNA sequencing (RNASeq) and ribosome profiling (RiboSeq) to delineate gene expression in the IBV M41-CK and Beau-CK strains at sub-codon resolution. RNASeq facilitated a comparative analysis of viral RNA synthesis and revealed two novel transcription junction sites in the attenuated Beau-CK strain, one of which would generate a sgmRNA encoding a ribosomally occupied ORF (dORF) located downstream of the nucleocapsid coding region. RiboSeq permitted quantification of the translational efficiency of virus gene expression and identified, for the first time, sites of ribosomal pausing on the genome. Quantification of reads flanking the programmed ribosomal frameshifting (PRF) signal at the genomic RNA ORF1a/ORF1b junction revealed that PRF in IBV is highly efficient (33–40%). Triplet phasing of RiboSeq data allowed precise determination of reading frames and revealed the translation of two ORFs (4b and 4c on sgmRNA IR), which are widely conserved across IBV isolates. Analysis of differential gene expression in infected primary chick kidney cells indicated that the host cell response to IBV occurs primarily at the level of transcription, with global up-regulation of immune-related mRNA transcripts following infection, and comparatively modest changes in the translation efficiencies of host genes. Cellular genes and gene networks differentially expressed during virus infection were also identified, giving insights into the host cell response to IBV infection.

IMPORTANCE

IBV is a major avian pathogen and presents a substantial economic burden to the poultry industry. Improved vaccination strategies are urgently needed to curb the global spread of this virus, and the development of suitable vaccine candidates will be aided by an improved understanding of IBV molecular biology. Our high-resolution data have enabled a precise study of transcription and translation in cells infected with both pathogenic and attenuated forms of IBV, and expand our understanding of gammacoronaviral gene expression. We demonstrate that gene expression shows considerable intra-species variation, with single nucleotide polymorphisms associated with altered production of sgmRNA transcripts, and our RiboSeq data sets enabled us to uncover novel ribosomally occupied ORFs in both strains. The numerous cellular genes and gene networks found to be differentially expressed during virus infection provide insights into the host cell response to IBV infection.

INTRODUCTION

Avian infectious bronchitis virus (IBV) is a member of the genus *Gammacoronavirus* (family *Coronaviridae,* order *Nidovirales*) and a pathogen of the domestic fowl (1). IBV infects primarily the epithelial cells of upper and lower respiratory tract tissues, though infections can also spread to the alimentary canal, as well as to the kidneys, testes and oviduct (2). The monopartite, polycistronic genomic RNA (gRNA) of IBV is approximately 27 kb in length, and – like those of other coronaviruses – it is 5′-methyl-capped and 3′-polyadenylated (3). Two large open reading frames (ORFs) – ORF1a and ORF1b – are situated within the 5′-proximal two-thirds of the genome. Translation of the former yields a *ca.* 3,950-aa polyprotein (pp1a); whereas translation of the latter requires −1 programmed ribosomal frameshifting (PRF) (4, 5), giving rise to a *ca.* 6,630-aa polyprotein (pp1ab). These polyproteins are cleaved to yield the components of the membrane-bound replication-transcription complex (RTC) (6-8). A feature of coronavirus replication is the synthesis of a nested, 3′-coterminal set of subgenomic mRNAs (sgmRNAs) encoding the viral structural and accessory proteins. The 5′ end of each sgmRNA comprises a 56-nt sequence derived from the 5’ end of the genome, the so-called leader sequence (9, 10). Incorporation of the leader occurs as a result of "polymerase hopping" – or discontinuous transcription – during negative-strand synthesis. When the RTC encounters specific “body transcription regulatory sequences” (TRS-Bs), the nascent negative strand can re-pair with a closely homologous leader TRS (TRS-L) at the 3′ end of the leader, after which the viral polymerase completes negative-strand synthesis using the leader as template (**Fig. 1A;** diamond symbols) (8-14). Subsequently, the RTC synthesises positive-strand copies of the negative-strand genomic and sgmRNAs.

Amongst the best-characterised strains of IBV are those belonging to the Massachusetts serotype, which includes the virulent Massachusetts 41 (M41-CK; Ref. 15) isolate and the laboratory-attenuated Beau-CK variant (16). Whilst M41-CK is restricted to growth in primary chicken cells,

Beau-CK is capable of replicating in both avian and non-avian cell lines; including Vero (African green monkey kidney-derived) and baby hamster kidney cells (17-20). Polymorphisms in the spike (S) glycoprotein subunit 2 (S2), which spans the viral membrane, have been shown to be responsible for this variation in host cell tropism (21). Moreover, the S protein of M41-CK – but not that of Beau-CK – elicits an immunoprotective response *in vivo*; although recombinant transfer of the protein from the former to the latter does not restore pathogenicity (22). The extent to which these strains diverge in terms of virus gene expression, or in terms of host cell gene expression in response to infection, has not been investigated in detail.

The advent of high-throughput sequencing techniques offers a means to monitor viral gene expression at unprecedented resolution (23-27). Here, we performed deep sequencing of ribosome-protected fragments (RPFs) – known as RiboSeq – in tandem with whole transcriptome sequencing (RNASeq), on total RNA extracts from primary chicken kidney (CK) cells infected with Beau-CK and M41-CK strains of IBV.

RESULTS

RiboSeq and RNASeq Data Quality

RiboSeq and RNASeq libraries were prepared from two biological repeats each of Beau-CK-infected, M41-CK-infected, and mock-infected cells. Infections were at high multiplicity (MOI = ~3) and cells were processed at 24 hours post-infection (h p.i.). An average of 1,156,819 RPFs and 1,727,024 RNASeq reads were mapped to viral gRNA in the virus-infected RiboSeq libraries (**Supp. Table S1**). The RNASeq read coverage in the library derived from the second biological repeat of M41-CK-infected cells was lower than that of other libraries due to technical losses. However, 106,741 reads were mapped to the forward strand of the viral gRNA in this case –

corresponding to a coverage of approximately 3.8-fold – and these reads were generally evenly distributed along the gRNA; hence, the sequencing depth in this replicate was deemed sufficient for further analysis. The vast majority of RPFs mapping to viral and host protein-coding regions were between 27 and 29 nt in length (**Supp. Fig. S1**), consistent with the size of the RNA fragment protected by translating eukaryotic ribosomes from digestion by RNase I (28). The length distributions of RNASeq reads were much broader, in line with the size of the gel slice excised for sequencing of fragmented RNA. RPF length was strongly related to the RPF phase relative to the reading frame of the associated coding region: 27-nt RPFs were primarily in the +1 phase; whereas 28- and 29-nt RPFs were primarily in the 0 phase (**Supp. Fig. S2**). As expected, RNASeq reads were far more evenly split over the three phases, with a slight bias towards phase 0 (**Supp. Fig. S3**), which may reflect codon usage bias – such as a preference for the use of RNY codons (25, 29, 30) – compounded with adaptor-ligation bias during library preparation. A meta-analysis of host mRNA coding regions showed that the depth of coverage of RiboSeq 5′ read ends increased substantially 12 nt upstream of the AUG (initiation) codon for RPFs in phase 0 (generally 28- and 29-nt RPFs), and 11-nt upstream of the AUG codon for RPFs in phase +1 (generally 27-nt RPFs) (**Supp. Figs. S4 and S5**). This indicates that the ribosomal P-site is situated at an offset of 11 and 12 nt from the 5′ ends of RPFs for 27-nt and 28-/29-nt reads, respectively (23). Peaks in RNASeq 5′-read end coverage were seen at the A of initiation (AUG) codons and at the middle nucleotide of termination (UNN) codons, respectively (**Supp. Figs. S6 and S7**), and are considered an artefact of ligation bias.

Fig. 1 illustrates the RiboSeq (red) and RNASeq (green) read coverage of the Beau-CK **(panel A)** and M41-CK **(panel B)** genomes. In both cases, the density of RPFs was considerably higher towards the 3′ ends of the gRNA, consistent with production of the 3′ co-terminal nested set of

sgmRNAs. In contrast, RiboSeq coverage of the ORF1a and ORF1b coding sequences was relatively low; reflecting the fact that a substantial proportion of newly synthesised gRNA (but not sgmRNA) transcripts are likely to be destined for packaging rather than translation (31). On average, negative-sense RNASeq reads were present at 0.28% of the level of positive-sense reads on average, indicating a ratio of positive:negative stranded RNA of ~350:1 at 24 h p.i., a ratio similar to that seen in ribosome profiling studies of the betacoronavirus mouse hepatitis virus (MHV) (25). Negative-sense RPFs, which may represent contamination from ribonucleoprotein complexes (25), were observed, but at low abundance (0.03% of the level of positive-sense RPFs).

Virus transcription: sequence divergence associated with IBV strain-specific TRS usage

The density of RNASeq reads mapping to a given sgmRNA represents the cumulative sum of reads derived from the gRNA and those derived from the overlapping portions of other subgenomic transcripts (**Fig. 1**). Therefore, to estimate the abundance of individual sgmRNAs, two independent approaches were used. First, we "decumulated" the raw RNASeq read densities mapping to inter-TRS regions, by subtracting the density of the 5′-adjacent inter-TRS region in each case (25). Secondly, the abundances of chimeric RNASeq reads spanning TRS junctions were quantified, by identifying unmapped reads containing an 11-nt sequence derived from the leader region, 5′-adjacent to the TRS-L (UAGAUUUUUAA, nt 46 – 56 in Beau-CK; UAGAUUUCCAA, nt 46 – 56 in M41-CK), and including at least 16 nt 3′ of this query. Chimeric reads were assigned to specific genomic loci based on the sequences 3′ of the TRS in each case (**Supp. Table S2; Fig. 2A**). Overall, the chimeric read abundances for sgmRNAs were significantly correlated with the corresponding decumulated RNASeq densities ($P < 0.01$ in both cases) (**Supp. Fig. S8**). The sequence logos in **Fig. 2B** and **Fig. 2C** illustrate the diversity of nucleotides found at TRS-B sites identified in this

study (including the novel sites discussed below) in Beau-CK and M41-CK, respectively. The core region of similarity to the TRS-L motif (CUUAACAA) is typically flanked by a 3′ adenine (A) or uracil (U) residue, and a preference for A/U residues is also seen immediately upstream of the core sequence. These flanking residues may facilitate template switching by lowering the free energy of anti-TRS-B/TRS-B duplex disassociation, since the TRS-L is also located in an AU-rich region (14).

Notably, the A nucleotides at positions four and seven of the core motif are the only invariant residues. In both Beau-CK and M41-CK, the TRS-B sequences associated with the S gene contain G residues at the third positions (CUGAACAA); in contrast to the TRS-L, which has a U at this position (CUUAACAA) (**Supp. Table S2**). Chimeric reads assigned to this gene were found to contain either a U or a G residue at position three (denoted "S [U3]" and "S [G3]", respectively, in **Fig. 2A; Supp. Table S2**); with a G being more common in M41-CK (7.5% of reads compared with 5.8%, on average) and a U being more common in Beau-CK (2.1% of reads compared with 1.5%, on average) (**Supp. Table S2**). These data indicate that the exact position at which discontinuous transcription occurs within a given TRS is subject to some variation, with either the TRS-L or the TRS-B templating the third residue. Similarly, the TRS-B for the 3a/3b/E genes diverges at the third position between Beau-CK (CUGAACAA; nt 23825 – 23832) and M41-CK (CUUAACAA; nt 23832 – 23839), with the latter matching the TRS-L sequence exactly. In this case, we found that Beau-CK-derived chimeric reads could contain either a U (denoted "3/E [U3]") or a G (denoted "3/E [G3]"), with the G residue being slightly more common (1.6% versus 1.2%, respectively, on average); whereas M41-CK-derived reads contained only the U residue (**Fig. 2A; Supp. Table S2**).

In agreement with a previous report (32), we found that the 3′-most of two adjacent canonical TRS-B sequences (both CUUAACAA; nt 25,460 – 25,467 and nt 25,471 – 25,478, labelled "5a/5b TRS 1" and "5a/5b TRS 2", respectively, in **Fig. 2A**) within the 30-nt region upstream of genes 5a/5b

was preferentially utilised in IBV Beau-CK; accounting for 18.8% of chimeric reads on average, compared with 1.5% for the 5′-most TRS-B. Interestingly, more chimeric reads were assigned to the non-canonical TRS-B associated with genes 4b/4c (33) – which has a low homology to the TRS-L (**Supp. Table S2**) – than to the first of these 5a/5b-associated TRSs in IBV Beau-CK; emphasising the importance of the genomic context in facilitating discontinuous transcription (**Fig. 2B** and **Fig. 2C**) (14). Only one of the two 5a/5b TRSs (TRS 2) is found in the IBV M41-CK genome (**Fig. 2A**).

Novel TRS in IBV Beau-CK

Two additional non-canonical leader/body chimeras were identified, both specific to the Beau-CK strain (**Supp. Table S2**). The more abundant of these (0.6% of chimeric reads) mapped to a position immediately downstream of the IBV Beau-CK N gene termination codon, within the 3′ "untranslated" region (UTR). Chimeric reads derived from this site contained the sequence CUUAACAU; the last six nt of which could have been templated by the genomic (TRS-B) sequence (UAACAU, nt 27104 – 27109). There is an AUG-initiated downstream ORF (dORF) in Beau-CK beginning two nt 3′ of this TRS, which comprises 11 codons (nt 27111 – nt 27143). Inspection of our RiboSeq libraries shows that the dORF is ribosomally occupied (**Fig. 2D**). Such AUG-initiated dORFs are present immediately 3′ of the N genes in most IBV strains and related avian coronaviruses, including TCoV, goose coronavirus (34, 35) and pigeon coronavirus (34), but this region appears to have been deleted in the IBV M41-CK lineage; and M41-CK also lacks the TRS-B downstream of the N gene (UAAAAU, nt 27156 – 27161).

The second novel chimeric sequence identified in RNASeq libraries maps to a TRS-B (CUUACCAA) within the coding region of the S gene in Beau-CK (nt 21242 – 21249). This is consistent with the previous detection of a sgmRNA of appropriate length via Northern blot analysis

(33). Whilst the core sequence of the TRS-B in this case is conserved in M41-CK, there is a single nucleotide (A to C) polymorphism located four nt downstream in the 3′ flanking region, which may contribute to its lack of utilisation in this strain (**Supp. Table S2**).

Virus translation: direct measurement of –1 PRF between ORF1a and ORF1b

Ribosome profiling of eukaryotic systems typically has the characteristic that mappings of the 5′ end positions of RPFs to coding sequences reflect the triplet periodicity of genetic decoding. A clear phase transition is evident in the RiboSeq libraries at the junction of ORF1a and ORF1b; where frameshifting of a proportion of ribosomes from the former ORF into the latter occurs (**Fig. 3A and 3B**). The mean normalised ratios of ORF1b to ORF1a RiboSeq density were 0.32 and 0.37 in IBV Beau-CK and IBV M41-CK, respectively; while the corresponding RNASeq ratios were 0.97 and 0.94, respectively (**Fig. 3C**). Thus on average, 33% of ribosomes in Beau-CK and 39% in M41-CK undergo –1 PRF prior to reaching the ORF1a termination codon (**Fig. 3D**). These values are very similar to those measured in *in vitro* PRF assays (4, 5) and alongside related profiling studies of MHV (25), this indicates that coronaviruses exhibit highly efficient PRF both *in vitro* and in the context of the infected cell.

Ribosomal occupancy of ORF4b and ORF4c

Situated between the M and 5a genes in Beau-CK and M41-CK is a >300-nt ostensibly “intergenic" region (IR) (**Fig. 1**). No protein-coding genes are annotated here but two putative AUG-initiated ORFs are present in each virus, referred to as ORF4b and ORF4c, after their homologs in turkey

coronavirus [TCoV] (36, 37) and in the genomes of most IBV isolates (38). The putative ORF4b genes of Beau-CK and M41-CK are encoded by nt 25,183 – 25,335 (50 codons) and nt 25,190 – 25,474 (94 codons), respectively, of the gRNA; whereas the ORF4c genes are encoded by nt 25,339 – 25,422 (27 codons) and nt 25,395 – 25,457 (20 codons), respectively (**Supp. Fig. S9)**. Thus, in Beau-CK, the two ORFs are separated by a 3-nt spacer region and in the same reading frame (**Fig. 4A**); whereas in M41-CK, ORF4c is located entirely within the ORF4b gene and in the +1 phase (**Fig. 4B**). Inspection of the ribosomal profiling datasets reveals substantial RPF coverage of both ORF4b and ORF4c, providing the first clear illustration that ORFs 4b and 4c are ribosomally occupied **(Fig. 4)**. Visualisation of ORF4c translation in M41-CK was facilitated by good phasing in the datasets, allowing expression of both ORF4b and ORF4c to be visualised (as both blue and orange RPF peaks in the overlap region). Previous work (33) has shown that a non-canonical TRS-B sequence – situated approximately 100 nt upstream of the M gene termination codon – facilitates production of a sgmRNA (IR) that harbours ORF4b at its 5′ end, and this TRS-B was also identified in our RNASeq data.

Translation efficiencies of IBV genes

To estimate the translational efficiency (TE) of virus genes, we summed RPFs whose 5′ end mapped in-phase between the first nucleotide of the initiation codon and 30 nt 5′ of the termination codon; thereby excluding RPFs derived from ribosomes paused during initiation or termination (the raw ribosome footprint data is provided in **Supp. Table S3**). The TE of each ORF was measured as the quotient of the RPF density and the abundance of the corresponding mRNA; with separate calculations performed using TRS chimeric reads counts and the decumulated RNASeq densities (**Fig. 5A** and **Supp. Fig. S10**, respectively). In the case of ORF4b and ORF4c, transcript abundance

could not be accurately deduced via the RNASeq decumulation procedure, because the significantly lower level of expression of the 4b/4c transcript relative to that of the 5′-adjacent M gene (**Fig. 1**) was associated with a proportionate increase in the level of noise. Similarly, as a result of the high abundance of gRNA relative to the sgmRNA encoding S, the decumulated RNASeq density for the latter is likely to be poorly estimated, and therefore the TE value for S calculated using the chimeric read count is likely to be more accurate. From this analysis, it was observed that the 4b gene is more efficiently translated than the 4c gene; a trend also observed for the accessory genes 3a/3b and 5a/5b (**Fig. 5A**). This is consistent with the likely requirement for leaky scanning to access the downstream ORF on each sgmRNA (see Discussion). Surprisingly, despite the fact that the nucleocapsid (N) protein is an abundant viral protein, it was not found to be efficiently translated relative to the other structural proteins, regardless of the approach used to estimate transcript abundance (**Fig. 5A; Supp. Fig. S10**). In the case of the ORF1a and ORF1b genes, a large proportion of the genomic RNA is expected to be destined for packaging rather than translation, as mentioned above, and this probably explains the low TE values calculated for these genes (**Fig. 5A; Supp. Fig. S10**). Additionally, the short length of the dORF precluded an accurate assessment of its translation efficiency. **Fig. 5B** compares the translation efficiencies of virus and host CDSs, with the former calculated using decumulated RNASeq densities. The latter are calculated on a per gene (rather than per transcript) basis, using RNASeq and RiboSeq reads contained entirely within annotated CDS regions (i.e. excluding 5′ and 3′ UTRs and also RPFs accumulating at or near to initiation or termination sites), and, like the virus values, are expressed relative to the mean levels for the cell (due to normalization by library size). The analysis shows that the virus translation efficiencies fall within the general range of those of host genes and indicates that virus transcripts are not preferentially translated during virus infection. Instead, the massive production of virus proteins (in particular the N protein) is achieved through high levels of transcription.

Ribosomal pauses during IBV genome translation

Inspection of the profiling datasets revealed a number of genomic locations where RPFs accumulated to a much higher level than at neighbouring sites, indicative of ribosomal pausing. As such pauses may have biological significance, we first sought to discount those that may have arisen artefactually. The known translation initiation sites in the virus genome generally showed high ribosome occupancy, but as the infected cells were treated with cycloheximide (CHX) prior to lysis to "freeze" ribosomes onto the mRNA, these pauses are likely to be over-represented, as ribosomes can accumulate at start codons during the CHX treatment period (39). Fluctuations in RPF density can also occur as a result of nuclease, ligation, and PCR biases during library construction. As the latter two biases can also occur during RNASeq library generation, we also discounted any pauses that had an obvious counterpart in RNASeq datasets. With these criteria, we identified five obvious sites of ribosomal pausing conserved in Beau-CK and M41-CK, one in the 5′ UTR and four within the coding region (indicated in **Fig. 1**, purple triangles; see **Table 1**). Pauses in 5′ UTRs can represent ribosomes initiating at upstream ORFs (uORFs), although in both Beau-CK and M41-CK, the P-site of the ribosome paused over bases 28–56 in the 5′ UTR of the genome is on a non-AUG codon (UUG) in a weak Kozak initiation consensus. As this pause is located upstream of the TRS_L, it reflects the sum of pausing on all sgmRNAs. To view the extent of the pause in context, we remapped reads to the most abundant sgmRNA, i.e. that of the N gene (**Fig. 6**). As can be seen, the "Leader pause" remains clearly evident (as is a smaller pause three codons downstream), albeit smaller in magnitude than those pauses seen at an N uORF (see below) and the authentic AUG codon of the N protein. Initiation at the UUG codon would result in translation of solely a dipeptide and thus the pause, if biologically relevant, may act as a regulator of downstream initiation events rather than through the encoded product. We note that an equivalent Leader pause is seen in MHV

(UUG codon, 1-codon ORF; Ref. 25). It is possible that pausing at this codon is potentiated by queueing of initiating ribosomes on sgmRNAs. The origin of the pauses within the coding region are enigmatic. The two adjacent pauses referred to collectively as Pause 2 in **Table 1** correspond to translation of a region of non-structural protein 4 (nsP4) downstream of the membrane spanning domains (40, 41). It is feasible that ribosomes pause here whilst the nascent peptide is being folded into membranes. Pauses 3 and 4 are noticeably large and correspond to ribosomes pausing soon after initiation of the S and M proteins, respectively. In the case of the former, the pause is unlikely to be linked to an interaction of the signal sequence at the N-terminus of the S protein with membranes, as this peptide would still be within the exit tunnel of paused ribosomes. Pause 5 corresponds to a potential non-AUG uORF (AUU, in a reasonable context) within the N mRNA (**Fig. 6**).

It is noteworthy that in our analysis of ribosomal pause sites, we did not see pausing at the AUG of the previously described 11 amino acid uORF of the genomic mRNA (AUG at nt 131–133; Ref. 42), and indeed there were few reads on the uORF itself, indicating that it is not heavily translated. Further, no pausing was observed at the PRF site at the ORF1a/ORF1b overlap.

Differential expression of host genes in response to IBV infection

We investigated the differential transcription and translation of host genes in response to IBV infection by comparing RNA and RPF densities per coding region for infected samples and mocks (see Materials and Methods). Details of the genes found to be differentially expressed (DEGs; FDR < 0.05 with multiple testing correction using the Benjamini-Hochberg method) at the level of transcription (4,266 genes) or translation (3,627 genes) respectively, are provided in **Supp. Data Sets S1 and S2.** Overall, the patterns of change in host cell gene expression in response to infection

were broadly similar for Beau-CK and M41-CK, with positive inter-strain correlations in the log2 fold changes (log2FC) in transcript abundance and TE (R^2 = 0.95 and R^2 = 0.85, respectively, *P* values both < 2.2×10^{-16}; **Fig. 7A**). Notably, the majority of differentially transcribed genes were up-regulated rather than down-regulated (i.e. log2FC > 0) for both strains (**Fig. 7A; left panel**), with 2.1-fold and 3.5-fold more up-regulated compared with down-regulated transcripts (FDR <0.05; see Materials and Methods) detected in Beau-CK-infected cells and M41-CK-infected cells, respectively (**Supp. Data Set S1**). This effect was not seen at the level of translation, where there were fewer differentially expressed genes overall, and the logFC values of those genes were more evenly distributed around 0, with slight skewing towards negative values (i.e. reduced TE) (**Fig. 7A; right panel** and **Supp. Data Set S2**). The core host transcriptional response to the two strains involved 579 commonly up-regulated and 132 commonly down-regulated genes, while the core translational response consisted of 34 commonly up-regulated and 79 commonly down-regulated genes. Gene ontology (GO) term enrichment revealed that numerous immune-related pathways were among the most significantly enriched terms in the core response sets (**Fig. 7B** and **Fig. 7C**). There was also evidence of integration and coordination of responses at the transcriptional and translational levels. For example, the GO term "positive regulation of NF-kappaB transcription factor activity" (GO:0051092) was enriched among transcriptionally up-regulated genes; whereas "negative regulation of NF-kappaB transcription factor activity" (GO:0032088) was enriched among translationally down-regulated genes. In a direct inter-strain comparison of statistically significant DEGs we identified 51 differentially transcribed genes, 45 of which were more highly expressed in Beau-CK-infected samples, and six of which were more highly expressed in M41-CK-infected samples (**Supp. Data Set S1**). The most significantly enriched GO term in the former set was "regulation of signalling receptor activity" (GO:0010469); while pro-proliferative and anti-apoptotic GO terms were also enriched (**Supp. Table S4**). The latter set included three heat shock

protein-encoding genes, and consequently the top enriched GO terms were related to "protein refolding" (**Supp. Table S5**). Just one gene (ENSGALG00000015358 [MYH15], encoding myosin heavy chain 15) had a significantly higher translation efficiency in M41-CK-infected samples compared with Beau-CK-infected samples.

In comparisons of host gene expression between Beau-CK-, M41-CK- and mock-infected cells, the significantly differentially expressed genes (FDR <0.05) were ranked by log2FC (**Supp. Data Set S3**) and the top 100 DEGs (or fewer) within each category were subjected to STRING analysis (43) to identify potential protein-protein interaction pathways (**Fig. 8 and Supp. Fig. S11)**. A selection of the key pathways proposed and examples of the associated genes are shown in **Table 2**. Clear patterns of host response to virus infection were present that are discussed below. Note in inter-strain comparisons of M41-CK versus Beau-CK, only the transcriptionally downregulated category had sufficient gene candidates for STRING analysis; the other three categories had a total of only seven DEGs (thus no plots are shown in the **Supp**. **Info.** for these).

Discussion

Here, we describe the first high-resolution study of gammacoronaviral gene expression during infection of primary chick kidney cells. Analysis of RNASeq data sets through chimeric read analysis or decumulation allowed us to quantify the relative levels of viral genomic and subgenomic mRNAs and to define the sequence diversity of strain-specific TRS utilisation. The predominant sgmRNA in both strains was that encoding the N protein, and between strains, the M transcript was relatively more abundant in M41-CK. In Beau-CK, two novel TRS were identified, one in the viral 3′ UTR immediately downstream of the N gene termination codon, and one mapping to a TRS-B within the S gene. In the former, a short ORF (dORF) – initiated two nt 3′ of the TRS – is present

and ribosomally occupied. The potential biological relevance of this ORF remains to be determined; such dORFs are present in most IBV strains and other avian gammacoronaviruses, but it is lacking in M41-CK (as is the TRS-B). A recent report has described the same sgmRNA (initiating at the identical TRS) as a novel non-coding RNA of IBV (44). The S gene TRS-B, proposed earlier (33), was also identified.

RiboSeq analysis, in conjunction with RNASeq, revealed that the N protein is not more efficiently translated than other structural proteins, despite being a structural component of IBV virions. It is possible that N expression may be regulated by a putative uORF whose initiation codon is located some 50 nt upstream of the N AUG codon (**Fig. 6**).

The efficiency of PRF at the IBV ORF1a/ORF1b overlap in natural infection was found to be 33–40%. This range is in close agreement with previous *in vitro* measurements of IBV frameshifting efficiency carried out using reporter constructs (5) and is consistent with the notion that coronaviruses are among the more efficient examples of canonical eukaryotic −1 PRF signals that have been studied to date (25, 45). Whether the modest difference in –1 PRF efficiency measured for Beau-CK and M41-CK has biological significance is uncertain, and it may represent experimental variation. The frameshift signal of M41-CK differs from Beau-CK in only three of 81 nucleotide positions, all of which are located in loop 3 of the stimulatory pseudoknot and not expected to affect pseudoknot function or stability (46). As also described for MHV-infected cells (25), there was no evidence that the frameshift-stimulatory pseudoknot induced ribosomal pausing on the slippery sequence. Thus pausing may not be a component of the frameshifting mechanism, or the pause may be too short-lived to be visualised by the profiling technique.

A meta-analysis of host genes revealed highly specific phasing of the RiboSeq data sets, enabling the accurate determination of the reading frame of translation for individual RPFs. Good phasing in the datasets and substantial read depth also allowed us to examine translation of viral accessory ORFs. It was evident that both 4b and 4c are efficiently translated, at levels comparable to those of the 5a/5b accessory protein-encoding genes. The mechanism by which ribosomes might access ORF4c, however, is not clear. Given the absence of AUG codons within the regions between the 5′ ends of ORF4b and ORF4c in both Beau-CK and M41-CK (**Supp. Fig. S9**), and the weak initiation context of the ORF4b start codon, it is possible that a proportion of ribosomes might bypass the ORF4b initiation codon and instead translate ORF4c via "leaky scanning" (47), although it should be noted that intervening AUG codons do exist in some other IBV strains (**Supp. Fig. S9**).

The relevance to virus gene expression of the sites of significant ribosomal pausing identified in the genome remains to be investigated experimentally. Two of these pause sites appear to correspond to uORFs initiated at non-AUG initiation codons, one in the 5′ UTR and one upstream of the N gene. In each case, ribosomes initiating on the main ORF AUG (ORF1a and N respectively) could potentiate initiation on the uORFs through stacking of scanning ribosomes, and this could be artefactually increased by the cycloheximide pre-treatment used during sample preparation. Two of the other pause sites correspond to ribosomes paused post-initiation early in the coding regions of the S and M genes. A biological explanation for this is lacking at present. We are aware that the treatment of yeast cells with cycloheximide can lead to an early block in elongation in stressed cells (48, 49), but meta-analysis of host genes in our infected cells does not reveal an obvious elongation block. Further, the S and M genes show deep ribosome coverage along their lengths, inconsistent

413 with a block in elongation. The remaining pause site (in fact a doublet) appears during translation of
414 the nsP4 region of the polyprotein close to the C-terminus of nsP4. Coronaviral nsP4 proteins are
415 important for the membrane rearrangements required for viral RNA synthesis and contain multiple
416 membrane spanning domains (41). A possible explanation for the ribosomal pauses seen here is that
417 translation is paused to permit the correct folding of nsP4 into membranes.

418

419 In general, the pauses we discern during translation of the IBV genome are discrete, substantial in
420 terms of read counts, and reproducible. As mentioned above, their origin is uncertain, but it seems
421 unrelated to the identity of the P-site tRNA. Recent studies have shown that P-site prolyl-tRNA is a
422 strong determinant of ribosomal pausing, partly due to the slow rate of peptide bond formation with
423 this amino acid (50, 51). However, none of the stall sites identified here have proline tRNA in the P-
424 site. It is possible that the nascent peptide engenders pausing through interactions with the ribosome
425 exit tunnel or chaperones as clusters of positively-charged amino acid residues have been
426 documented to induce ribosome pausing (51, 52). However, such clusters are not evident at the IBV
427 pause sites documented here. The recent developments of methodologies and algorithms to identify
428 and characterise ribosomal pause sites may clarify the situation in future (53, 54).

429

430 Our data indicate that the host response to IBV is mediated primarily at the level of transcription,
431 with the up-regulation of hundreds of genes, many of which have immune-related functions.
432 Changes in translational efficiency were modest, with more genes showing decreased rather than
433 increased translation in response to IBV infection. Many of the transcriptionally upregulated genes
434 identified reflect the host response to virus infection, as seen previously with IBV infection of
435 chickens (55, 56) and with other RNA viruses (57). Some of the protein pathways identified have

not been associated with coronavirus infection previously and warrant experimental follow up, for example, the potential downregulation of transcription of genes linked to FAM20C, a kinase that generates the majority of the extracellular phosphoproteome (58). Also of interest is the translational upregulation of ribosomal protein synthesis in infected cells for both Beau-CK and M41-CK. The basis of the attenuated phenotype of Beau-CK is not fully understood (reviewed in Ref. 59), but could potentially involve differential host cell binding (60) or features of the replicase genes (61). The direct comparison of DEGs in M41-CK and Beau-CK-infected primary chick kidney cells here did not identify any pre-eminent pathways that might reflect their differential pathogenesis, although several cytokines were expressed at a lower level in M41-CK-infected cells (IL6, IL8L1, IFN-Beta (**Supp. Fig. S11**). Overall, these data contribute towards a substantially improved understanding of the early innate immune response to IBV infection, including distinct features of the transcriptional and translational responses.

MATERIALS AND METHODS

Virus and cells: The apathogenic molecular clone of IBV, Beau-R, has been described previously (62) and was used to generate virus Beau-CK. The pathogenic isolate M41-CK (GenBank accession number MK728875.1) has been described previously (63). The two strains have an average nucleotide identity of 93% (assessed in a 1 kb window, with a step size of 200 nt). Most of the variation between them occurs as single nucleotide polymorphisms; however, a single large (185 nt) region is present in the 3' untranslated region (UTR) of Beau-CK which is absent from the genome of M41-CK. The Beau-CK sequence is identical to Beau-R (GenBank: AJ311317.1) excepting two point mutations, one in nsp16 (C19666U; Ser to Leu) and one in N (A27087G, synonymous). Primary chick kidney (CK) cells were produced from 2–3 week-old specific pathogen free (SPF)

459 Rhode Island Red chickens (64). CK cells (0.8×10^6 cells/ml) were plated in 10-cm dishes and upon
460 reaching 100% confluence (two days post-seeding) were washed once with PBS and infected with
461 9.6×10^6 PFU Beau-CK or M41-CK (MOI = ~3). After 1 hour incubation at 37 °C, 5% CO_2, the
462 inoculum was removed and replaced with 10 ml fresh 1x BES (1X minimal essential Eagle's
463 medium [MEM], 0.3% tryptose phosphate broth, 0.2% bovine serum albumin, 20 mM N,N-Bis(2-
464 hydroxyethyl)-2-aminoethanesulfonic acid (BES), 0.21% sodium bicarbonate, 2 mM L-glutamine,
465 250 U/ml nystatin, 100 U/ml penicillin, and 100 U/ml streptomycin). Cells were harvested at 24
466 hours post-infection when clear regions of cytopathic effect (CPE) were visible.
467
468 **Drug treatment, cell harvesting and lysis:** Cycloheximide (CHX; Sigma-Aldrich) was added
469 directly to the growth medium (to 100 µg/ml) and the cells incubated for 2 min at 37 °C before
470 rinsing with 5 ml of ice-cold PBS containing CHX (100 µg/ml). Subsequently, dishes were
471 incubated on ice and 400 µl of lysis buffer [20 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$,
472 1 mM DTT, 1% Triton X-100, 100 µg/ml cycloheximide and 25 U/ml TURBO™ DNase (Life
473 Technologies)] dripped onto the cells. The cells were scraped extensively to ensure lysis, collected
474 and triturated with a 26-G needle ten times. Lysates were clarified by centrifugation for 20 min at
475 13,000 g at 4 °C, the supernatants recovered and stored at −80 °C.
476
477 **Ribosomal profiling and RNASeq:** Cell lysates were subjected to RiboSeq and RNASeq. The
478 methodologies employed were based on the original protocols of Ingolia and colleagues (23, 65),
479 except ribosomal RNA contamination was removed using a commercial RiboZero Gold magnetic
480 kit (Illumina) and library amplicons were constructed using a small RNA cloning strategy (66)
481 adapted to Illumina smallRNA v2 to allow multiplexing. The methods used were as described by
482 Chung et al. (67), except the 5′ and 3′ adapters included seven consecutive randomised bases at the

3′ and 5′ ends (respectively). This facilitated removal of reads duplicated during polymerase chain reaction (PCR) amplification of cDNA libraries (68) and reduced ligation bias. Amplicon libraries were deep sequenced using an Illumina NextSeq platform (Department of Pathology, University of Cambridge).

Computational analysis of RiboSeq and RNASeq data

Adaptor sequences were trimmed using FASTX-Toolkit (hannonlab.cshl.edu/fastx_toolkit), and reads shorter than 25 nt following adaptor trimming were discarded. Mapping was performed using Bowtie version 1 (69) with parameters -v 2 --best (i.e. maximum 2 mismatches, report best match). Adaptor-trimmed, de-duplicated reads were mapped sequentially to host (*Gallus gallus*) ribosomal RNA (rRNA); IBV Beau-CK (GenBank accession: NC_001451.1) or IBV M41-CK (GenBank accession: DQ834384.1) gRNA; Ensembl host non-coding RNA (ncRNA); NCBI RefSeq host mRNA; and to the host genome. The order of mapping was tested to check that virus-derived reads were not lost accidentally due to mis-mapping to host RNA, or *vice versa*. When performing analyses of viral and host gene expression, only 28- and 29-nt RiboSeq reads (corresponding to RPFs mapping primarily in phase 0) and only ≥ 40 nt RNASeq reads were used. A 12-nt offset was applied to the 5′ mapping positions of RPFs, to approximate the P-site position of the ribosome (see **Supp. Fig. S4** and Ref. 25). To normalize for different library sizes, reads per million mapped reads (RPM) values were calculated using the sum of total virus RNA plus total host RefSeq mRNA (positive sense reads only) as the denominator.

Host mRNA RiboSeq and RNASeq phasing distributions were derived from reads mapping internally to the coding regions of ORFs; specifically, the 5′ end of the read had to map between the first nucleotide of the initiation codon and 30 nt 5′ of the last nucleotide of the termination codon,

thus, in general, excluding RPFs of initiating or terminating ribosomes. Histograms of 5′ end positions of host mRNA reads relative to initiation and termination codons (Supp. Figs. 4 – 7) were derived from reads mapping to RefSeq mRNAs with annotated CDSs at least 450 nt in length and annotated 5′ and 3′ UTRs at least 60 nt in length. When calculating the translation efficiencies of viral genes, only in-phase (i.e. phase 0 with respect to the ORF in question) RiboSeq reads were counted.

For host differential expression analyses, non-ribosomal, non-viral reads in each library were mapped to the *Gallus gallus* 5.0 assembly (December 2015) using STAR (70), with gene annotations from Ensembl release 94 (71). A maximum of two mismatches were allowed when mapping. Read counts per gene (protein-coding genes only) were obtained using HTSeq (72), with a requirement that reads map entirely within the forward strand coding sequence (htseq-count parameters: -m intersection-strict -s yes -t CDS). For each comparison of experimental groups, only genes with an average of at least 50 mapped reads were included in differential expression analyses. GO term enrichment analysis was carried out using the topGO package in R (73) and Fisher's exact test was used to assess the enrichment of individual GO terms in specific gene lists. Protein-protein interaction networks were constructed using the Search Tool for Retrieval of Interacting Genes (STRING) database (43).

Data availability

Sequencing data have been deposited in ArrayExpress (http://www.ebi.ac.uk/arrayexpress) under the accession number E-MTAB-7849.

ACKNOWLEDGEMENTS

This work was supported by UK Medical Research Council [MR/M011747/1] and Wellcome Trust [202797/Z/16/Z] grants to I.B. and Wellcome Trust grant [106207] and European Research Council (ERC) European Union's Horizon 2020 research and innovation programme grant [646891] to A.E.F.

REFERENCES

1. Cavanagh, D. (2005). Coronaviruses in poultry and other birds. Avian Pathol. 34: 439-448.

2. Cavanagh, D. (2007). Coronavirus avian infectious bronchitis virus. Vet. Res. 38: 281-297.

3. Gorbalenya, A.E., Enjuanes, L., Ziebuhr, J. and Snijder, E.J. (2006). Nidovirales: evolving the largest RNA virus genome. Virus Res. 117:17-37.

4. Brierley, I., Boursnell, M.E., Binns, M.M., Bilimoria, B., Blok, V.C., Brown, T.D. and Inglis, S.C. (1987). An efficient ribosomal frameshifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO J. 6:3779-3785.

5. Brierley, I, Digard, P. and Inglis, S.C. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell 57:537-547.

554 6. Liu, D.X., Xu, H.Y. and Brown, T.D. (1997). Proteolytic processing of the coronavirus infectious
555 bronchitis virus 1a polyprotein: identification of a 10-kilodalton polypeptide and determination of
556 its cleavage sites. J Virol. 71:1814-1820.
557
558 7. Brockway, S.M, Clay, C.T., Lu, X.T. and Denison, M.R. (2003). Characterization of the
559 expression, intracellular localization, and replication complex association of the putative mouse
560 hepatitis virus RNA-dependent RNA polymerase. J. Virol. 77:10515-10527.
561
562 8. Sawicki, S.G., Sawicki, D.L. and Siddell, S.G. (2007). A contemporary view of coronavirus
563 transcription. J. Virol. 81:20-29.
564
565 9. Brown, T.D., Boursnell, M.E., Binns, M.M. and Tomley, F.M. (1986). Cloning and sequencing of
566 5' terminal sequences from avian infectious bronchitis virus genomic RNA. J. Gen. Virol. 67:221-
567 228.
568
569 10. Sawicki, S.G. and Sawicki, D.L. (1995). Coronaviruses use discontinuous extension for
570 synthesis of subgenome-length negative strands. Adv. Exp. Med. Biol. 380:499-506.
571
572 11. Sawicki, S.G. and Sawicki, D.L. (1998). A new model for coronavirus transcription. Adv. Exp.
573 Med. Biol. 440:215-219.
574
575 12. Pasternak, A.O., van den Born, E., Spaan, W.J. and Snijder, E.J. (2001). Sequence requirements
576 for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. EMBO J.
577 20:7220-7228.

13. Zúñiga, S., Sola, I., Alonso, S. and Enjuanes, L. (2004). Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. J. Virol. 78:980-994.

14. Sola, I., Moreno, J.L., Zúñiga, S., Alonso, S. and Enjuanes, L. (2005). Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. J. Virol. 79:2506-2516.

15. Van Roekel, H., Clarke, M.K., Bullis, K.L., Olesiuk, O.M. and Sperling, F.G. (1951). Infectious bronchitis. Am. J. Vet. Res. 12:140-146.

16. Beaudette, F. R. and Hudson C. R. (1937). Cultivation of the virus of infectious bronchitis. J. Am. Vet. Med. Assoc. 90:51–60.

17. Cunningham, C.H., Spring, M.P. and Nazerian, K. (1972). Replication of avian infectious bronchitis virus in African green monkey kidney cell line VERO. J. Gen. Virol. 16:423-427.

18. Otsuki, K., Noro, K., Yamamoto, H. and Tsubokura, M. (1979). Studies on avian infectious bronchitis virus (IBV). II. Propagation of IBV in several cultured cells. Arch. Virol. 60:115-122.

19. Alonso-Caplen, F.V., Matsuoka, Y., Wilcox, G.E. and Compans, R.W. (1984). Replication and morphogenesis of avian coronavirus in Vero cells and their inhibition by monensin. Virus Res. 1:153-167.

602 20. Casais, R., Dove, B., Cavanagh D. and Britton, P. (2003). Recombinant avian infectious
603 bronchitis virus expressing a heterologous spike gene demonstrates that the spike protein is a
604 determinant of cell tropism. J. Virol. 77:9084-9089.
605
606 21. Bickerton, E., Maier, HJ., Stevenson-Leggett, P., Armesto, M. and Britton, P. (2018). The S2
607 Subunit of Infectious Bronchitis Virus Beaudette Is a Determinant of Cellular Tropism. J. Virol. 92.
608 pii: e01044-18.
609
610 22. Hodgson, T., Casais, R., Dove, B., Britton, P. and Cavanagh, D. (2014). Recombinant infectious
611 bronchitis coronavirus Beaudette with the spike protein gene of the pathogenic M41 strain remains
612 attenuated but induces protective immunity. J. Virol. 78:13804-13811.
613
614 23. Ingolia, N.T., Ghaemmaghami, S., Newman, J.R. and Weissman, J.S. (2009). Genome-wide
615 analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. Science.
616 324:218-223.
617
618 24. Stern-Ginossar, N. (2015). Decoding viral infection by ribosome profiling. J. Virol. 89:6164-
619 6166.
620
621 25. Irigoyen, N., Firth, A.E., Jones, J.D., Chung, B.Y., Siddell, S.G and Brierley, I. (2016). High-
622 resolution analysis of coronavirus gene expression by RNA sequencing and ribosome profiling.
623 PLoS Pathog. 12:e1005473.
624

625 26. Irigoyen, N., Dinan, A.M., Brierley, I. and Firth, A.E. (2018). Ribosome profiling of the
626 retrovirus murine leukemia virus. Retrovirology 15: 10.
627
628 27. Stewart, H., Brown, K., Dinan, A.M., Irigoyen, N., Snijder, E.J. and Firth, A.E. (2018).
629 Transcriptional and translational landscape of equine torovirus. J. Virol. 92 pii: e00589-18.
630
631 28. Wolin, S.L. and Walter, P. (1988). Ribosome pausing and stacking during translation of a
632 eukaryotic mRNA. EMBO J. 7:3559-3569.
633
634 29. Jukes, T.H. (1996). On the prevalence of certain codons ("RNY") in genes for proteins. J. Mol.
635 Evol. 42:377-381.
636
637 30. Fuchs, R.T., Sun, Z., Zhuang, F. and Robb, G.B. (2015). Bias in ligation-based small RNA
638 sequencing library construction is determined by adaptor and RNA structure. PLoS One.
639 10:e0126049.
640
641 31. Kuo, L. and Masters, P.S. (2013). Functional analysis of the murine coronavirus genomic RNA
642 packaging signal. J. Virol. 87:5182-5192.
643
644 32. Stirrups, K., Shaw, K., Evans, S., Dalton, K., Casais, R., Cavanagh, D. and Britton, P. (2000).
645 Expression of reporter genes from the defective RNA CD-61 of the coronavirus infectious
646 bronchitis virus. J. Gen. Virol. 81:1687-1698.
647

33. Bentley, K., Keep, S.M., Armesto, M. and Britton, P. (2013). Identification of a noncanonically transcribed subgenomic mRNA of infectious bronchitis virus and other gammacoronaviruses. J. Virol. 87:2128-2136.

34. Jonassen, C.M., Kofstad, T., Larsen, I.L., Løvland, A., Handeland, K., Follestad, A. and Lillehaug, A. (2005). Molecular identification and characterization of novel coronaviruses infecting graylag geese (Anser anser), feral pigeons (Columbia livia) and mallards (Anas platyrhynchos). J. Gen. Virol. 86:1597-607.

35. Papineau, A., Berhane, Y., Wylie, T.N., Wylie, K.M., Sharpe, S. and Lung, O. (2019). Genome organization of Canada goose coronavirus, a novel species identified in a mass die-off of Canada geese. Sci. Rep. 9:5954.

36. Gomaa, M.H., Barta, J.R., Ojkic, D. and Yoo, D. (2008). Complete genomic sequence of turkey coronavirus. Virus Res. 135:237-246.

37. Cao, J., Wu, C.C. and Lin, T.L. (2008). Complete nucleotide sequence of polyprotein gene 1 and genome organization of turkey coronavirus. Virus Res. 136:43-49.

38. Reddy, V.R., Theuns, S., Roukaerts, I.D., Zeller, M., Matthijnssens, J. and Nauwynck, H.J. (2015). Genetic characterization of the Belgian nephropathogenic infectious bronchitis virus (NIBV) reference strain B1648. Viruses 7:4488-4506.

671 39. Ingolia, N.T., Lareau, L.F. and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic
672 stem cells reveals the complexity and dynamics of mammalian proteomes. Cell. 147:789-802.
673
674 40. Oostra, M., te Lintelo, E.G., Deijs, M., Verheije, M.H., Rottier, P.J. and de Haan, C.A. (2007).
675 Localization and membrane topology of coronavirus nonstructural protein 4: involvement of the
676 early secretory pathway in replication. J. Virol. 81:12323-12336.
677
678 41. Doyle, N., Neuman, B.W., Simpson, J., Hawes, P.C., Mantell, J., Verkade, P., Alrashedi, H. and
679 Maier, H.J. (2018). Infectious bronchitis virus nonstructural protein 4 alone induces membrane
680 pairing. Viruses 10 pii: E477.
681
682 42. Boursnell, M.E., Brown, T.D., Foulds, I.J., Green, P.F., Tomley, F.M., Binns, M.M. (1987).
683 Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J.
684 Gen. Virol. 68:57-77.
685
686 43. Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A.,
687 Doncheva, N.T., Roth, A., Bork, P., Jensen, L.J. and von Mering, C. (2017). The STRING database
688 in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic
689 Acids Res. 45(D1):D362-D368.
690
691 44. An, H., Cai, Z., Yang ,Y., Wang, Z., Liu, D.X. and Fang, S. (2019). Identification and formation
692 mechanism of a novel noncoding RNA produced by avian infectious bronchitis virus. Virology
693 528:176-180.
694

45. Atkins, J.F., Loughran, G., Bhatt, P.R., Firth, A.E. and Baranov, P.V. (2016). Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. Nucleic Acids Res. 44:7007-7078.

46. Brierley I. and Pennell, S. (2001). Structure and function of the stimulatory RNAs involved in programmed eukaryotic-1 ribosomal frameshifting. Cold Spring Harb. Symp. Quant. Biol. 66:233-248.

47. Firth, A.E. and Brierley, I. (2012). Non-canonical translation in RNA viruses. J. Gen. Virol. 93:1385-1409.

48. Gerashchenko, M.V. and Gladyshev, V.N. (2014). Translation inhibitors cause abnormalities in ribosome profiling experiments. Nucleic Acids Res. 42:e134.

49. Duncan, C.D.S. and Mata, J. (2017). Effects of cycloheximide on the interpretation of ribosome profiling experiments in Schizosaccharomyces pombe. Sci. Rep. 7:10331.

50. Sabi, R. and Tuller T. (2015). A comparative genomics study on the effect of individual amino acids on ribosome stalling. BMC Genomics 16 Suppl. 10:S5.

51. Sabi, R. and Tuller, T. (2017). Computational analysis of nascent peptides that induce ribosome stalling and their proteomic distribution in *Saccharomyces cerevisiae*. RNA 23:983-994.

718 52. Ingolia, N.T. (2014). Ribosome profiling: new views of translation, from single codons to
719 genome scale. Nat. Rev. Genet. 15:205-213.
720
721 53. Chadani, Y., Niwa, T., Chiba, S., Taguchi, H. and Ito, K. (2016). Integrated *in vivo* and *in vitro*
722 nascent chain profiling reveals widespread translational pausing. Proc. Natl. Acad. Sci. U.S.A.
723 113:E829-38.
724
725 54. Kumari, R., Michel, A.M. and Baranov, P.V. (2018). PausePred and Rfeet: webtools for
726 inferring ribosome pauses and visualizing footprint density from ribosome profiling data. RNA
727 24:1297-1304.
728
729 55. Cong, F., Liu, X., Han Z., Shao, Y., Kong, X. and Liu, S. (2013). Transcriptome analysis of
730 chicken kidney tissues following coronavirus avian infectious bronchitis virus infection. BMC
731 Genomics 14:743.
732
733 56. Chhabra, R., Ball, C., Chantrey, J. and Ganapathy, K. (2018). Differential innate immune
734 responses induced by classical and variant infectious bronchitis viruses in specific pathogen free
735 chicks. Dev. Comp. Immunol. 87:16-23.
736
737 57. Zhang, J., Kaiser, M.G., Deist, M.S., Gallardo, R.A., Bunn, D.A., Kelly, T.R., Dekkers, J.C.M.,
738 Zhou, H. and Lamont, S.J. (2018). Transcriptome analysis in spleen reveals differential regulation
739 of response to Newcastle disease virus in two chicken lines. Sci. Rep. 8:1278.
740

741 58. Tagliabracci, V.S., Wiley, S.E., Guo, X., Kinch, L.N., Durrant, E., Wen, J., Xiao, J., Cui, J.,
742 Nguyen, K.B., Engel, J.L., Coon, J.J., Grishin, N., Pinna, L.A., Pagliarini, D.J. and Dixon, J.E.
743 (2015). A single kinase generates the majority of the secreted phosphoproteome. Cell 161:1619-
744 1632.
745
746 59. Lin, S.Y. and Chen, H.W. (2017). Infectious Bronchitis Virus Variants: Molecular Analysis and
747 Pathogenicity Investigation. Int. J. Mol. Sci. 18 pii: E2030.
748
749 60. Wickramasinghe, I.N., de Vries, R.P., Gröne, A., de Haan, C.A. and Verheije, M.H. (2011).
750 Binding of avian coronavirus spike proteins to host factors reflects virus tropism and pathogenicity.
751 J. Virol. 85:8903-8912.
752
753 61. Armesto, M., Cavanagh, D. and Britton, P. (2009). The replicase gene of avian coronavirus
754 infectious bronchitis virus is a determinant of pathogenicity. PLoS One 4:e7384.
755
756 62. Casais, R., Thiel, V., Siddell, S.G., Cavanagh, D. and Britton, P. (2001). Reverse genetics
757 system for the avian coronavirus infectious bronchitis virus. J. Virol. 75;12359-12369.
758
759 63. Kottier, S.A., Cavanagh, D. and Britton, P. (1995). Experimental evidence of recombination in
760 coronavirus infectious bronchitis virus. Virology. 213:569-580.
761
762 64. Hennion, R. and Hill, G. (2015). In *Coronaviruses* Vol. 1282 *Methods in Molecular Biology*
763 (eds Helena Jane Maier, Erica Bickerton, & Paul Britton) Ch. 6, 57–62. (Springer New York).
764

765 65. Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M. and Weissman, J.S. (2012). The
766 ribosome profiling strategy for monitoring translation *in vivo* by deep sequencing of ribosome-
767 protected mRNA fragments. Nat. Protoc. 7:1534-1550.
768
769 66. Guo, H., Ingolia, N.T., Weissman, J.S. and Bartel, D.P. (2010). Mammalian microRNAs
770 predominantly act to decrease target mRNA levels. Nature 466:835-840.
771
772 67. Chung, B.Y., Hardcastle, T.J., Jones, J.D., Irigoyen, N., Firth, A.E., Baulcombe, D.C. and
773 Brierley, I. (2015). The use of duplex-specific nuclease in ribosome profiling and a user-friendly
774 software package for Ribo-seq data analysis. RNA 21:1731-1745.
775
776 68. Aird, D., Ross, M.G., Chen, W.S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D.B., Nusbaum,
777 C. and Gnirke, A. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing
778 libraries. Genome Biol. 12: R18.
779
780 69. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009). Ultrafast and memory-efficient
781 alignment of short DNA sequences to the human genome. Genome Biol. 10:R25.
782
783 70. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M.
784 and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21.
785
786 71. Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M.R., Armean, I.M., Bennett, R.,
787 Bhai, J., Billis, K., Boddu, S. et al. (2019). Ensembl 2019. Nucleic Acids Res. 47:D745-D751.
788

72.. Anders, S., Pyl, P.T. and Huber, W. (2015). HTSeq - a Python framework to work with high-throughput sequencing data. Bioinformatics 31:166-169.

73. Alexa, A. and Rahnenfuhrer, J (2018). *topGO: Enrichment Analysis for Gene Ontology*. R package version 2.34.0.

FIGURE LEGENDS

Figure 1. Structure and read coverage of the (A) IBV Beau-CK and (B) IBV M41-CK genomes. Coverage in the RiboSeq (red) and RNASeq (green) libraries is plotted on a logarithmic scale, with negative sense reads in blue. RPM is reads per million mapped reads. Histograms show the positions of the 5′ ends of reads. The 5′ two-thirds of the IBV gRNA contains two large ORFs (1a and 1b) encoding pp1a and pp1b, respectively. Translation of the latter requires −1 programmed ribosomal frameshifting (PRF) at the indicated site. A nested set of 3′-coterminal sgmRNAs is produced during infection. Diamond symbols show the locations of canonical TRSs at which discontinuous transcription occurs (TRS-L in orange and TRS-B in green). Downward arrows indicate the positions of non-canonical TRSs discussed in this work. Purple triangles indicate sites of ribosomal pausing (see text).

Figure 2. (A) Proportion of chimeric reads assigned to each of the indicated TRS junctions. Novel TRS identified in this study are indicated with asterisks. Note that the 5a/5b TRS 1 and the dORF TRS are present in IBV Beau-CK only. **(B)** Sequence logo depicting nucleotides surrounding the identified TRS-B sites in IBV Beau-CK. **(C)** Equivalent sequence logo for IBV M41-CK. **(D)** RiboSeq and RNASeq coverage of the IBV Beau-CK dORF. A +12 nt offset is applied to the 5′ ends of all reads, to approximate the position of the ribosomal P-site in RiboSeq libraries (and to make the RNASeq coverage directly comparable). Reads whose 5′ ends map to the first, second or third positions of codons are indicated in blue, purple, and orange, respectively and ORFs are coloured according to the frame in which they are encoded. The location of the novel TRS-B sequence, which begins 2 nt 3′ of the N gene termination codon, is indicated with an arrow.

Figure 3. RiboSeq and RNASeq coverage proximal to the junction between ORF1a and ORF1b for IBV Beau-CK **(A)** and IBV M41-CK **(B)**. The last 2,500 nt of ORF1a and first 2,500 nt of ORF1b are shown. Coverage is normalised to reads per million mapped reads (RPM), using the sum of total virus RNA plus total host RefSeq mRNA (positive sense reads only) as the denominator, and smoothed with a 121-codon sliding window. Reads in phase 0, +1, and +2 relative to ORF1a are shown in blue, purple, and orange, respectively; and ORFs are coloured according to the frame in which they are encoded. **(C)** Ratios of ORF1b to ORF1a read density expressed as reads per kilobase per million mapped reads (RPKM). RPKM values exclude the 150-nt regions downstream of the ORF1a initiation codon, upstream of the ORF1b termination codon, and either side of the frameshift site. **(D)** Frameshifting efficiencies calculated using the values plotted in **(C)**.

Figure 4. RiboSeq and RNASeq coverage of ORF4b and ORF4c in **(A)** IBV Beau-CK and **(B)** IBV M41-CK. Coverage is expressed as reads per million mapped reads (RPM). Reads in phase 0, +1, and +2 relative to ORF4b are shown in blue, purple, and orange, respectively; and ORFs are coloured according to the frame in which they are encoded.

Figure 5. (A) Translation efficiencies (TEs) of virus ORFs, as calculated using the relative abundances (reads per million, RPM) of chimeric TRS-spanning RNASeq reads. Values shown are relative to the mean efficiency per TRS (TEs of virus ORFs calculated using decumulated RNASeq densities are shown in **Supp. Fig. S10). (B)** Comparison of host and virus translation efficiencies. TEs of virus ORFs were calculated using decumulated RNASeq densities, as described in the text. Host mRNA TEs are based on the ratio (after normalization for library size) of all RiboSeq or RNASeq reads mapping to any annotated coding region of any splice form of a given gene. Host

data are shown only for genes with ≥ 50 RNASeq coding-region reads on average across samples (prior to normalization for library size). Horizontal dashed lines indicate the mean values for host cell genes. Note, the points for 1a and 1b overlap.

Figure 6. RiboSeq and RNASeq coverage of sgmRNA N in IBV Beau-CK. Coverage is expressed as reads per million mapped reads (RPM). Reads in phase 0, +1, and +2 relative to N are shown in purple, orange and blue, respectively; and ORFs are coloured according to the frame in which they are encoded.

Figure 7. (A) Log2 fold changes (log2FC) in host transcript abundance and translation efficiency in infected cells relative to mocks. In both cases, a high degree of correlation was observed between logFC values in Beau-CK-infected samples (*x*-axes) and M41-CK-infected samples (*y*-axes), with transcript abundances skewed towards positive log2FC values. **(B)** The ten most significantly enriched GO terms among commonly up-regulated (left panel) and commonly down-regulated (right panel) genes at the level of transcription. **(C)** The ten most significantly enriched GO terms among commonly up-regulated (left panel) and commonly down-regulated (right panel) genes at the level of translation efficiency.

Figure 8. STRING analysis of the relationships between differentially expressed transcripts in comparisons of IBV M41-CK- and mock-infected cells. **(A)** Downregulated genes **(B)** Upregulated genes. The network nodes represent the proteins encoded by the differentially expressed genes. Seven different coloured lines link a number of nodes and represent seven types of evidence used in predicting associations. A red line indicates the presence of fusion evidence; a green line represents

neighbourhood evidence; a blue line represents co-occurrence evidence; a purple line represents experimental evidence; a yellow line represents text-mining evidence; a light blue line represents database evidence; and a black line represents co-expression evidence.

Table 1. Ribosomal pause sites within the IBV genome.

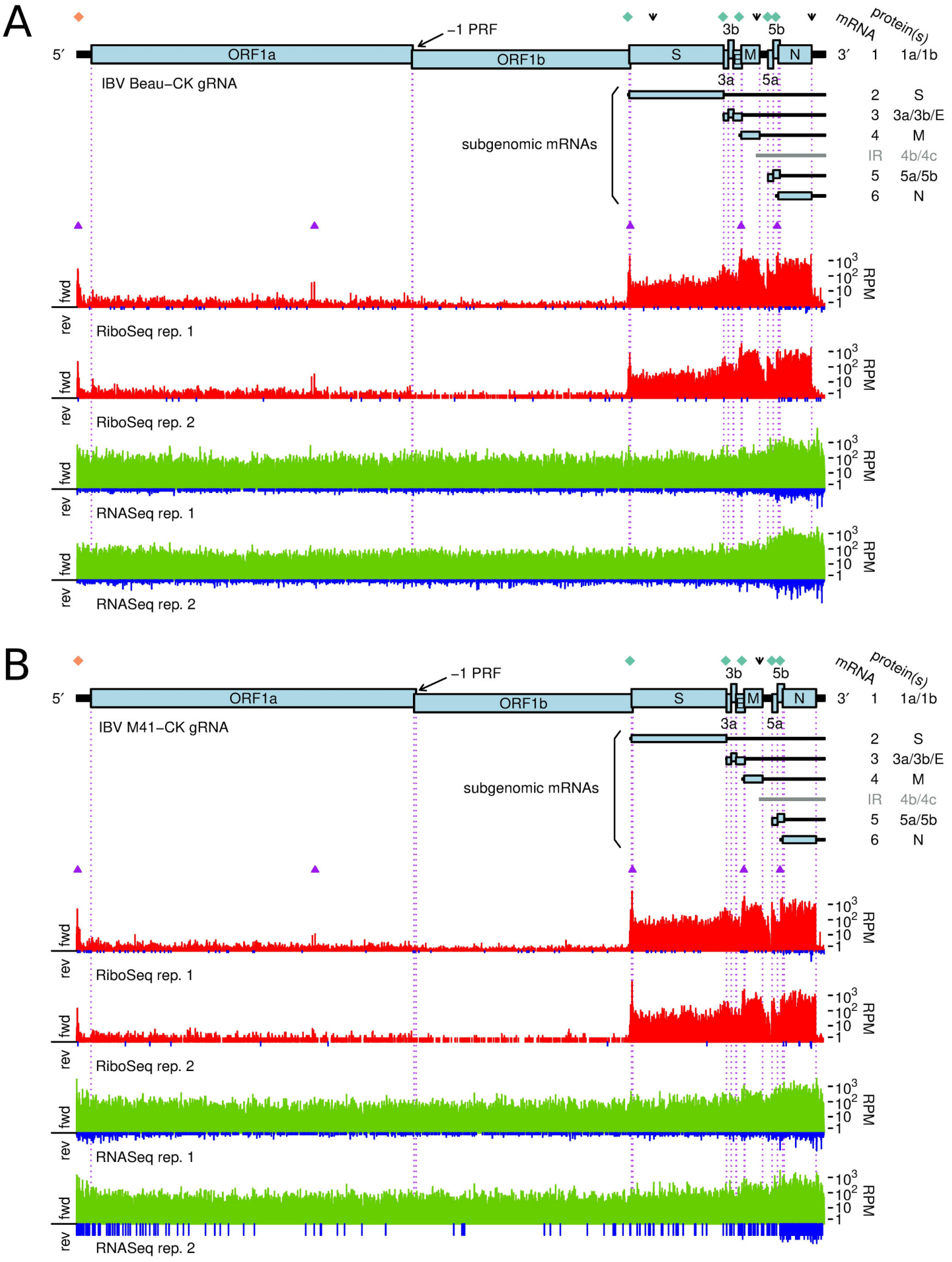
Pause	Genomic location and RPF sequence	Nascent peptide
Pause 1 (5′NCR)	5′ end of genome (bases 28–56) near TRS_L 5′ AUUACACUAGCCUUGCGCUAGAUUUUUA 3′	*YISITLALR*
Pause 2 (nsP4)	Two adjacent peaks within nsP4 coding region (~8660 and 8760) 5′ UUUGUUAAGCUUACUAAUGAGAUAGGU 3′ 5′ UUGCAAGCUUGUCGUGCAUGGUUAGCU 3′	YDGNEFVGNYDLAAKSTFVIRGSEFVKLTN KFEAYLSAYARLKYYSGTGSEQDYLQACRA
Pause 3 (S)	Large pause downstream of initiation codon of S protein (~20,410) 5′ CUAGUGACUCUUUUGUGUGCACUAUGU 3′ (Beau)	†MLVTPLLLVTLLC (Beau)
Pause 4 (M)	Very large pause immediately downstream of initiation codon of the M protein (~24,500) 5′ (AUG)CCCAACGAGACAAAUUGUACUCUUGACU 3′.	†MPNETNC
Pause 5 (N)	Broad pause peak centred on YLSSIPREN near end of 5b ORF (~25,830), just upstream of N start codon ribosome stack 5′ UACCUCUCUAGUAUUCCAAGGGAAAACU 3′	QSRTSRALSRVYLSSIP(RENL*)

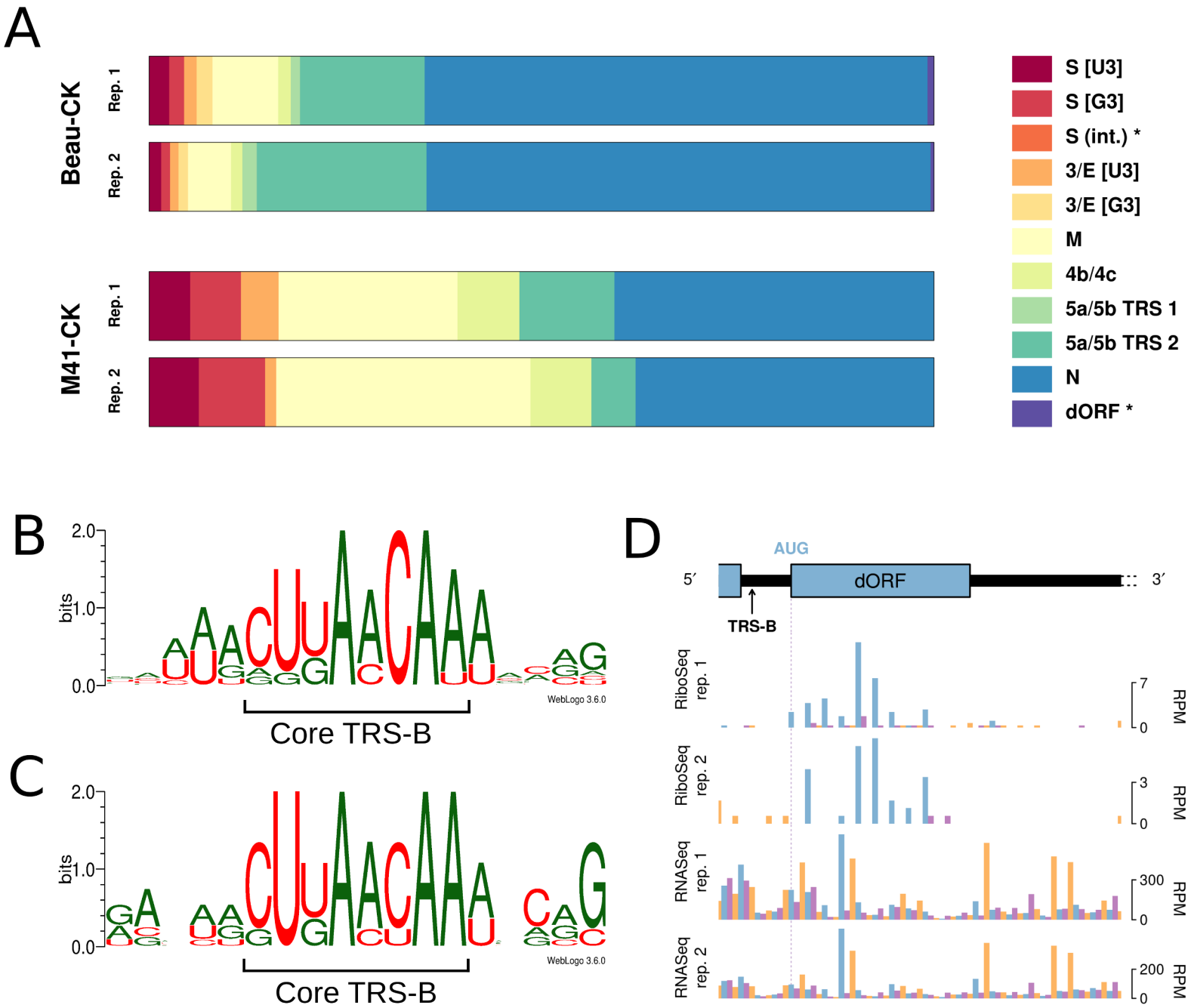
Underlined characters signify codon/amino acid of the ribosomal P-site tRNA.
* in-frame stop codon.
† initiator methionine

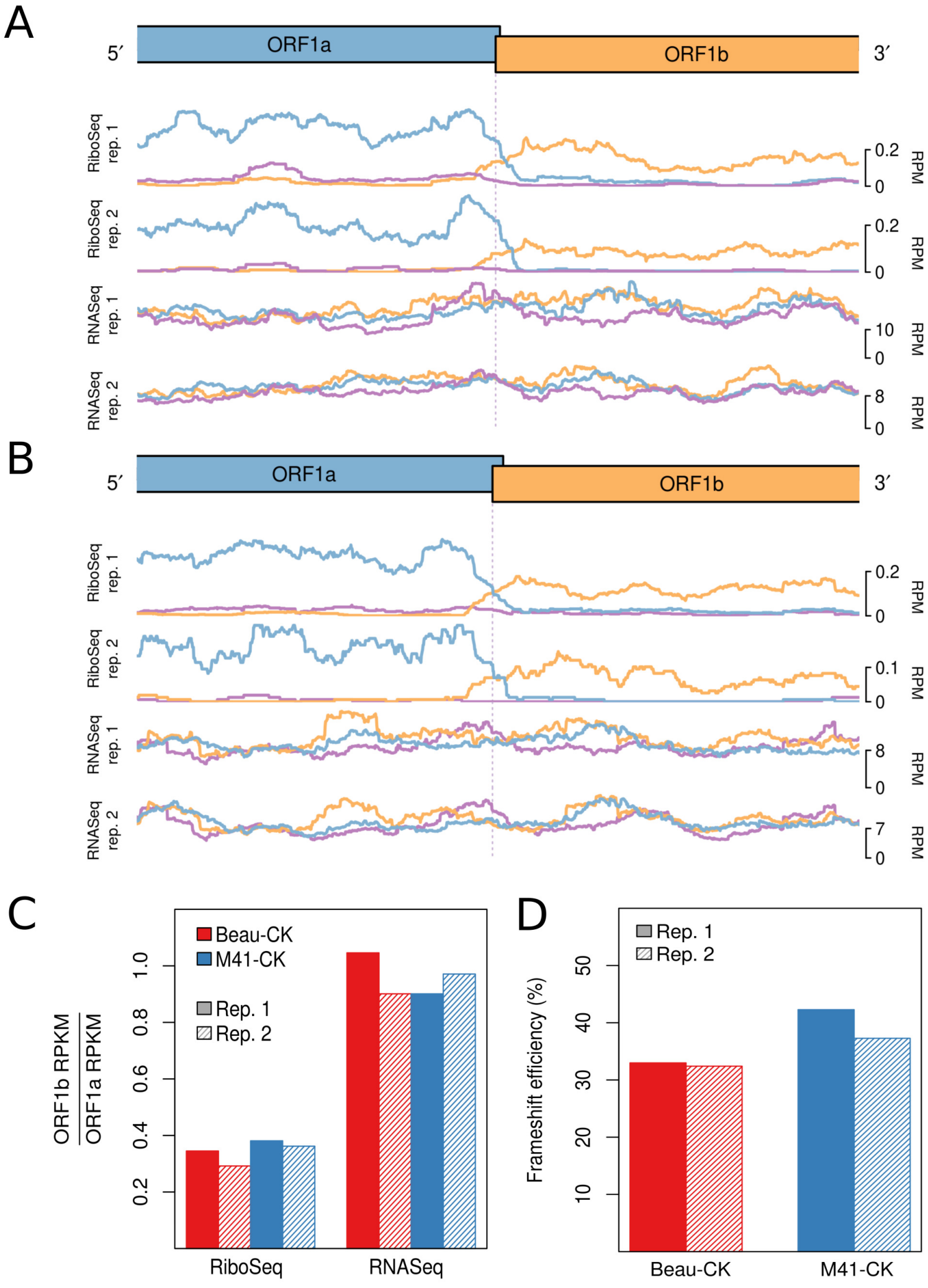
Table 2. STRING analysis of differential gene expression.

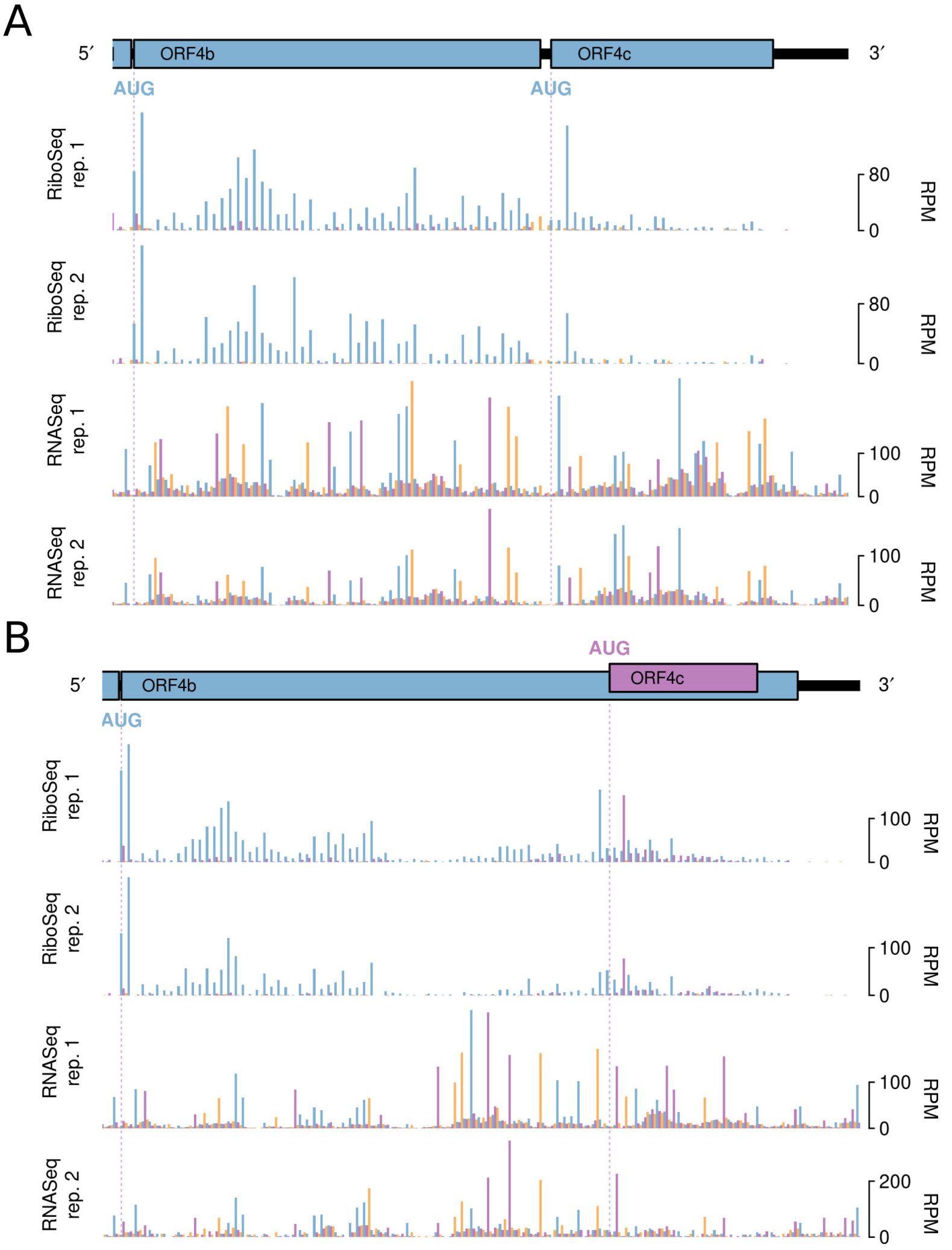
Comparison	Parameter	Main pathway(s)	Examples in pathway(s) (FDR; log2fc)
Transcription			
Beau-CK vs Mock	Downregulated	FAM20C substrates	SPP1 (secreted phosphoprotein 1) (7.31×10^{-15}; –2.72). TF (transferrin) (1.83×10^{-13}; –3.17).
M41-CK vs Mock	Downregulated	FAM20C substrates	SPP1 (secreted phosphoprotein 1) (6.13×10^{-10}; –3.177). CHGB (chromogranin B) (7.42×10^{-8}; –2.09).
Beau-CK vs Mock	Upregulated	Antiviral state, receptor signalling, cytokine interactions	RSAD2 (viperin) (3.92×10^{-155}; 9.49). IFIT5 (interferon induced protein with tetratricopeptide repeats 5) (4.48×10^{-142}; 8.66).
M41-CK vs Mock	Upregulated	Antiviral state, receptor signalling, cytokine interactions	RSAD2 (viperin) (1.80×10^{-140}; 9.04). IFIT5 (interferon induced protein with tetratricopeptide repeats 5) (3.01×10^{-119}; 7.96).
M41-CK vs Beau-CK	Downregulated	Cytokines, cytokine-receptor interactions	IL6 (interleukin 6) (4.38×10^{-4}; –1.51). IL8L1 (interleukin 8-like 1) (1.38×10^{-2};–1.44).
M41-CK vs Beau-CK	Uregulated	* Heat shock family members	HSPA5 (heat shock 70kDa protein 5) (2.28×10^{-5}; 1.74). HSP90AA1(heat shock protein 90 alpha family class A member 1) (4.38×10^{-4}; 1.44).
Translation			
Beau-CK vs Mock	Downregulated	No obvious pathways identified (top two hits shown to right)	TIPARP (TCDD inducible poly(ADP-ribose) polymerase) (8.42×10^{-23}; –4.32). ADAMTS1 (ADAM metallopeptidase with thrombospondin type 1 motif 1) (1.0×10^{-12}; –3.57).
M41-CK vs Mock	Downregulated	No obvious pathways identified (top two hits shown to right)	TIPARP (TCDD inducible poly(ADP-ribose) polymerase) (8.42×10^{-23}; –4.32). PFKFB3 (6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3) (2.47×10^{-13}; –3.49).
Beau-CK vs Mock	Upregulated	Antiviral response, translation, 80S ribosome, RACK1	OASL (2'-5'-oligoadenylate synthetase like) (5.08×10^{-6}; 2.24). RPSL37 (ribosomal protein L37) (5.34×10^{-5}; 1.78).
M41-CK vs Mock	Upregulated	Antiviral response, 80S ribosome	OASL (2'-5'-oligoadenylate synthetase-like) (1.16×10^{-3}; 1.79). RPS8 (ribosomal protein S8) (1.03×10^{-2}; 1.35).
M41-CK vs Beau-CK	Downregulated	No pathways identified	No significant genes identified.
M41-CK vs Beau-CK	Upregulated	No pathways identified	Only one significant gene identified, MYH15 (myosin heavy chain 15) (0.032; 1.77)

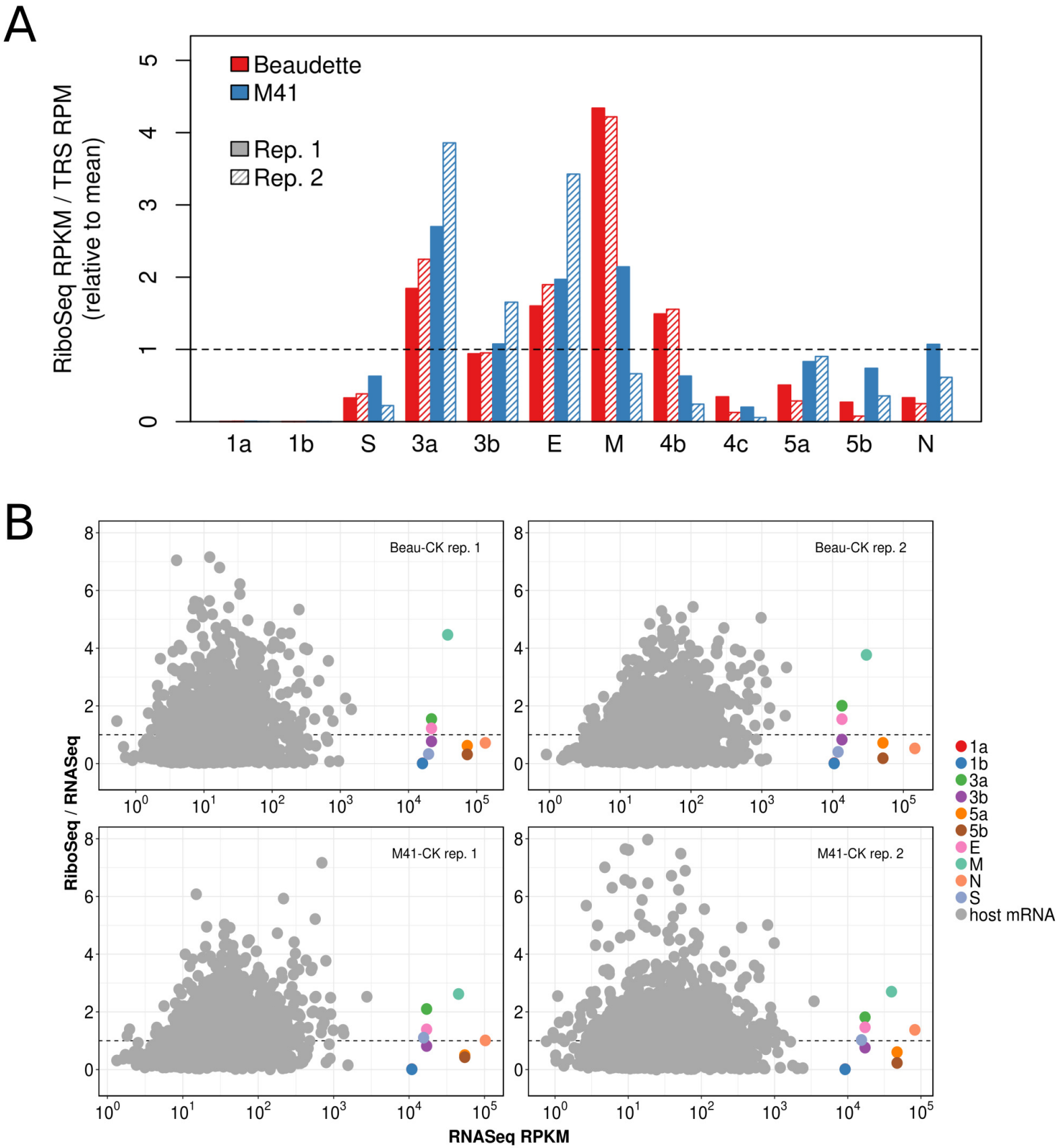
*Only six of the top 100 DEGs were significant in this category (see Supplementary Data S3).

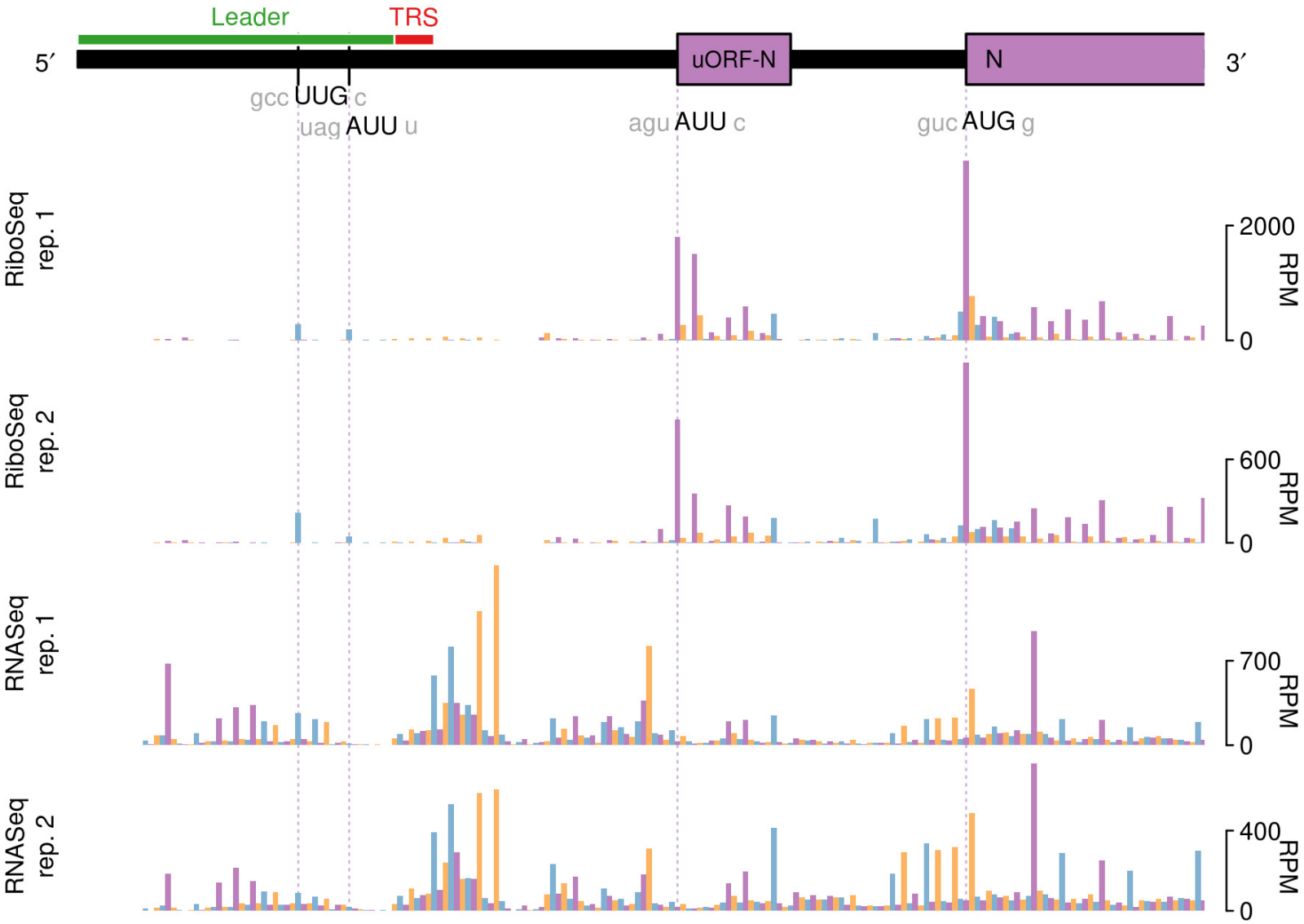

A
5′
ORF1a
−1 PRF
ORF1b
S
3b
M
5b
N
3′
mRNA
protein(s)
1
1a/1b
IBV Beau−CK gRNA
3a
5a
subgenomic mRNAs
2
S
3
3a/3b/E
4
M
IR
4b/4c
5
5a/5b
6
N
fwd
rev
RPM
RiboSeq rep. 1
RiboSeq rep. 2
RNASeq rep. 1
RNASeq rep. 2
B
IBV M41−CK gRNA

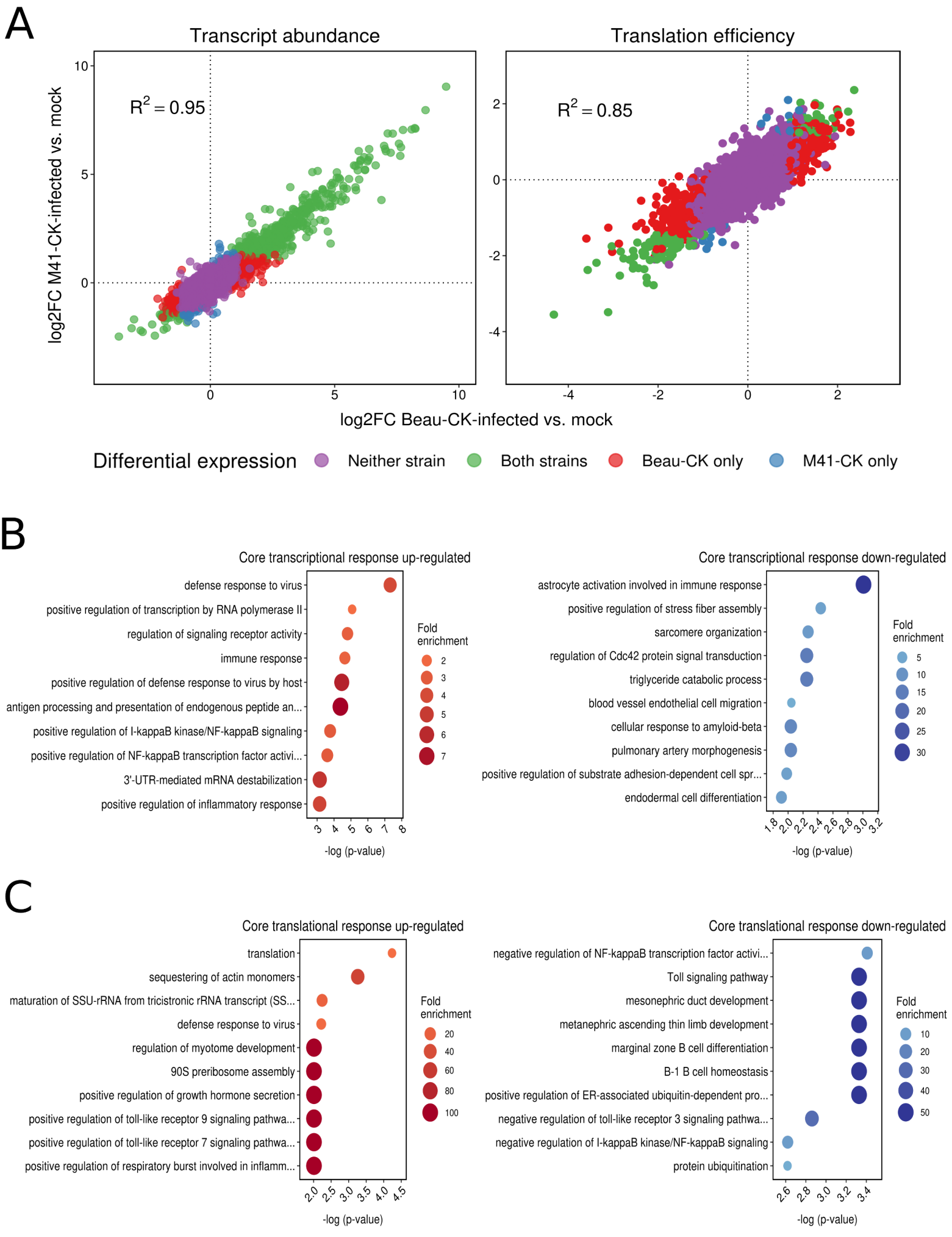

A
Beau-CK
Rep. 1
Rep. 2
M41-CK
Rep. 1
Rep. 2
S [U3]
S [G3]
S (int.) *
3/E [U3]
3/E [G3]
M
4b/4c
5a/5b TRS 1
5a/5b TRS 2
N
dORF *
B
bits
2.0
1.0
0.0
WebLogo 3.6.0
Core TRS-B
C
bits
2.0
1.0
0.0
WebLogo 3.6.0
Core TRS-B
D
AUG
5′
dORF
3′
TRS-B
RiboSeq rep. 1
RiboSeq rep. 2
RNASeq rep. 1
RNASeq rep. 2
RPM
7
0
3
0
300
0
200
0

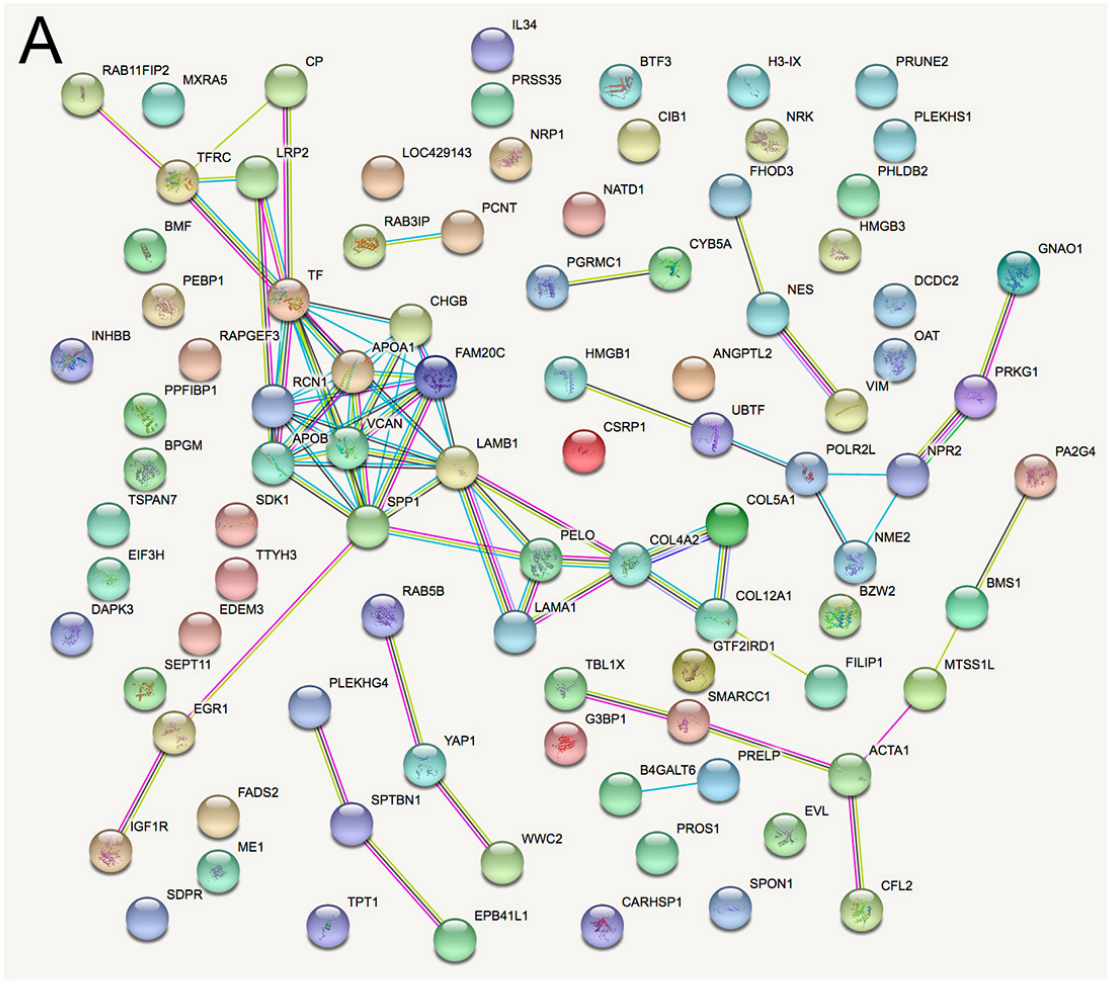

A
5′
ORF1a
ORF1b
3′
RiboSeq rep. 1
RiboSeq rep. 2
RNASeq rep. 1
RNASeq rep. 2
0.2
0
10
8
RPM
B
5′
ORF1a
ORF1b
3′
RiboSeq rep. 1
RiboSeq rep. 2
RNASeq rep. 1
RNASeq rep. 2
0.2
0.1
8
7
RPM
C
Beau-CK
M41-CK
Rep. 1
Rep. 2
ORF1b RPKM / ORF1a RPKM
0.2
0.4
0.6
0.8
1.0
RiboSeq
RNASeq
D
Rep. 1
Rep. 2
Frameshift efficiency (%)
10
20
30
40
50
Beau-CK
M41-CK


A
5′
ORF4b
ORF4c
3′
AUG
AUG
RiboSeq rep. 1
80
0
RPM
RiboSeq rep. 2
80
0
RPM
RNASeq rep. 1
100
0
RPM
RNASeq rep. 2
100
0
RPM
B
AUG
5′
ORF4b
ORF4c
3′
AUG
RiboSeq rep. 1
100
0
RPM
RiboSeq rep. 2
100
0
RPM
RNASeq rep. 1
100
0
RPM
RNASeq rep. 2
200
0
RPM


A
RiboSeq RPKM / TRS RPM
(relative to mean)
Beaudette
M41
Rep. 1
Rep. 2
0
1
2
3
4
5
1a
1b
S
3a
3b
E
M
4b
4c
5a
5b
N
B
Beau-CK rep. 1
Beau-CK rep. 2
M41-CK rep. 1
M41-CK rep. 2
RiboSeq / RNASeq
RNASeq RPKM
1a
1b
3a
3b
5a
5b
E
M
N
S
host mRNA


Leader
TRS
5′
uORF-N
N
3′
gcc UUG c
uag AUU u
agu AUU c
guc AUG g
RiboSeq rep. 1
2000
RPM
0
RiboSeq rep. 2
600
RPM
0
RNASeq rep. 1
700
RPM
0
RNASeq rep. 2
400
RPM
0


A
Transcript abundance
Translation efficiency
$R^2 = 0.95$
$R^2 = 0.85$
log2FC M41-CK-infected vs. mock
log2FC Beau-CK-infected vs. mock
Differential expression
Neither strain
Both strains
Beau-CK only
M41-CK only
B
Core transcriptional response up-regulated
defense response to virus
positive regulation of transcription by RNA polymerase II
regulation of signaling receptor activity
immune response
positive regulation of defense response to virus by host
antigen processing and presentation of endogenous peptide an...
positive regulation of I-kappaB kinase/NF-kappaB signaling
positive regulation of NF-kappaB transcription factor activi...
3'-UTR-mediated mRNA destabilization
positive regulation of inflammatory response
Fold enrichment
-log (p-value)
Core transcriptional response down-regulated
astrocyte activation involved in immune response
positive regulation of stress fiber assembly
sarcomere organization
regulation of Cdc42 protein signal transduction
triglyceride catabolic process
blood vessel endothelial cell migration
cellular response to amyloid-beta
pulmonary artery morphogenesis
positive regulation of substrate adhesion-dependent cell spr...
endodermal cell differentiation
Fold enrichment
-log (p-value)
C
Core translational response up-regulated
translation
sequestering of actin monomers
maturation of SSU-rRNA from tricistronic rRNA transcript (SS...
defense response to virus
regulation of myotome development
90S preribosome assembly
positive regulation of growth hormone secretion
positive regulation of toll-like receptor 9 signaling pathwa...
positive regulation of toll-like receptor 7 signaling pathwa...
positive regulation of respiratory burst involved in inflamm...
Fold enrichment
-log (p-value)
Core translational response down-regulated
negative regulation of NF-kappaB transcription factor activi...
Toll signaling pathway
mesonephric duct development
metanephric ascending thin limb development
marginal zone B cell differentiation
B-1 B cell homeostasis
positive regulation of ER-associated ubiquitin-dependent pro...
negative regulation of toll-like receptor 3 signaling pathwa...
negative regulation of I-kappaB kinase/NF-kappaB signaling
protein ubiquitination
Fold enrichment
-log (p-value)


A
IL34
RAB11FIP2
MXRA5
CP
PRSS35
BTF3
H3-IX
PRUNE2
CIB1
NRK
PLEKHS1
TFRC
LRP2
LOC429143
NRP1
FHOD3
PHLDB2
NATD1
BMF
RAB3IP
PCNT
HMGB3
CYB5A
GNAO1
PGRMC1
PEBP1
TF
NES
DCDC2
CHGB
INHBB
RAPGEF3
APOA1
FAM20C
OAT
HMGB1
ANGPTL2
PPFIBP1
RCN1
VIM
PRKG1
UBTF
VCAN
CSRP1
APOB
LAMB1
POLR2L
NPR2
PA2G4
BPGM
TSPAN7
SDK1
SPP1
COL5A1
PELO
COL4A2
NME2
EIF3H
TTYH3
DAPK3
EDEM3
RAB5B
LAMA1
COL12A1
BZW2
BMS1
SEPT11
GTF2IRD1
FILIP1
MTSS1L
TBL1X
PLEKHG4
SMARCC1
EGR1
G3BP1
YAP1
PRELP
ACTA1
B4GALT6
FADS2
SPTBN1
EVL
IGF1R
PROS1
WWC2
ME1
SPON1
CFL2
SDPR
TPT1
CARHSP1
EPB41L1
B
OGFR
CMTR1
TIPARP
SPTBN5
LRP3
SNX10
PLACL2
B3GNT7
TRANK1
LOC770718
TRAFD1
RASSF5
PPM1J
KLHL10
TAP1
LY6E
ADAMTS1
ZC3HAV1
MOV10
HAS2
GBP
ZNFX1
BCMO1
HELZ2
GDPD4
RNF213
ZFAND2A
PARP14
PMLL
ISG12-2
DDX60
IFI27L2
CAMKK2
NLRC5
EPSTI1
OASL
USP18
EIF2AK2
PHF11
LOC417973
TRIM25
MX1
RSAD2
PARP9
NR4A2
SLC2A6
IRF7
FOXS1
IFIT5
CH25H
SOCS1
IFIH1
PARP12
CMPK1
NT5C3L
STAT1
PDK4
GADD45B
TLR3
TNIP2
ATF3
IL13RA1
STAT2
ELF3
IFNB
TMEM173
CASZ1
CD274
IKBKE
PIGA
IL6
TNFAIP3
FAM149A
IL12A
IL8L1
GVINP1
SQSTM1
GATA3
PFKFB3
LOC417192
BIRC2
GGT5
ISL1
ANGPT1
CD83
GGCL1
ZBTB49
EPB41L3
SCEL