METHODS AND SYSTEMS FOR HETEROLOGOUS RNA PAYLOAD PRODUCTION IN VIRUS-LIKE PARTICLES IN THE YEAST GENUS YARROWIA

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority under 35 U.S.C. §119(e) to U.S. Provisional Application No. 63/683,077, filed August 14, 2024. The content of the prior application considered part of and is hereby incorporated by reference in its entirety.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING

[0002] The material in the accompanying sequence listing is hereby incorporated by reference into this application. The accompanying sequence listing xml file, name SBIO1120-1WO, Sequence Listing ST26.xml, was created on February 4, 2025, and is 59,474 bytes.

FIELD OF DISCLSOURE

[0003] The present disclosure generally relates to methods and systems for producing RNA payloads in living cells in the yeast *Yarrowia sp.* by encapsulating the RNA in Tobacco Mosaic virus capsid proteins to form viral like particles (VLPs) to increase the stability, yield and/or translational efficiency of the RNA that is extracted from those VLPs as compared to other yeasts.

BACKGROUND

[0004] The growing understanding of RNA functions and their critical role in health and disease has promoted the development of various commercial and research applications of RNA. Currently, the only method used to produce therapeutic-grade mRNA is in vitro transcription (IVT). IVT requires multiple, separate reactions and extensive purification steps resulting in a costly and complex GMP manufacturing process. Moreover, during the DNA template transcription step in IVT, immunogenic double-stranded RNA side products are produced that reduce the efficacy of the therapeutic RNA produced and evoke a proinflammatory response. In addition, purified RNA obtained using in vitro transcription methods tends to be unstable and vulnerable to degradation.

[0005] One potential alternative to producing therapeutic mRNA in IVT is through the generation of virus-like particles (VLPs) containing the desired mRNA in an appropriate host cell. United States patent publication US20200010866 discloses the use of yeast lacking an

endogenous Ty retrotransposon that is transformed with a first DNA sequence that encodes the full-length Gag protein of a yeast retrotransposon and a second DNA sequence that encodes an RNA molecule comprising an RNA sequence of interest linked to an RNA sequence encoding a fragment of the Gag protein to produce RNA. The RNA sequence encoding a fragment of the Gag protein associates with the expressed full-length Gag protein, the latter of which forms a VLP containing the encoded RNA molecule. This work was done in *Saccharomyces paradoxus* and in *Saccharomyces cerevisiae*.

[0006] VLPs made up of capsid proteins from different viruses have been used to produce vaccines and immunogens. In some cases, yeast strains have been used to produce these VLPs. The predominant yeast strains used in VLP-based biomanufacturing are *Saccharomyces cerevisiae*, *Pichia pastoris,* and *Hansenula polymorpha*. S. Brachelente et al., Appl. Microbiol. 2023, 3, 805-825.

[0007] One of the most studied forms of VLPs are those comprising the Tobacco Mosaic Virus (TMV) capsid protein. In TMV, a packaging sequence termed an Origin of Assembly Sequence (OAS), is located approximately 1 kb from the 3' end of its genomic RNA. The OAS sequence forms a hairpin structure capable of inserting into the disk-shaped rings of the TMV capsid protein (TMV CP) that initiates the assembly of the TMV virus like particles (VLPs) into a helical filament-capsid structure. P.J.G. Butler, Philosophical Transactions of the Royal Society of London, 1999, 354: 537-550; see also P.J.G. Butler, J Gen Virol, 1984, 65 (Pt 2): 253-279. There are relatively few reports of using TMV VLPs in yeast to produce mRNA. However, there is one report on encapsidating heterologous mRNA containing an OAS in TMV VLPs produced in the yeast *Schizosaccharomyces pombe*. A. Kadri et al, J Virol Meth, 2013, 189(2), 328-340.

[0008] The yeast genus *Yarrowia*, which includes *Y. lipolytica*, *Y. divulgata,* and *Y. keelungensis*, is often considered a nonconventional yeast due to its distinctive genome structure and its relatively large phylogenetic distance to other yeasts while sharing common properties with higher eukaryotes. *Y. lipolytica* is used in industry to produce lipids, organic acids such as citric acid, terpenoids and some proteins (e.g., lipase, phospholipase and alpha-glucosidase). YK Park and R Ledesma-Amaro, 2023, Trends in Biotechnology, 41(2), 242-54; M. Garvey, J. Fungi, 2022, 8, 1179. What little work has been done on producing VLPs in *Yarrowia* is limited to producing vaccines, i.e., a VLP expressing a viral antigen on its surface. VT Luu et al, J Microbiol, 2017, 55(8), 655-64. We are unaware of any reports using *Yarrowia* to produce VLPs for RNA production.

[0009] Based on the above, there remains a need for systems that produce therapeutic quality RNA in sufficient quantities that would bypass the aforementioned limitations of the IVT and which can take advantage of the inherent production advantages arising from the properties of virus-like particles.

SUMMARY OF THE INVENTION

[0010] Applicants herein provide a solution to this problem by producing TMV VLPs containing an RNA molecule of interest (e.g., a therapeutic RNA molecule or a therapeutic mRNA molecule) in *Yarrowia*. The amounts of desired RNA produced in this system were 20-fold greater than using a similar TMV VLP system in *S. cerevisiae*. In addition, when the RNA was isolated from these VLPs and used to transfect THP-1 derived macrophages, the RNA from the Yarrowia-produced VLPs produced a statistically significant greater expression of the encoded protein as compared to the RNA from *S. cerevisiae*-derived VLPs. The RNA from the *Yarrowia*-produced VLPs also showed substantially greater expression in THP-1 derived macrophages as compared to a similar IVT produced RNA. In primary human dendritic cells, the *Yarrowia* VLP-derived RNA caused less induction of various inflammatory cell surface markers than the similar IVT-produced RNA. These results suggest that more RNA is produced in *Yarrowia* VLPs than in corresponding *S. cerevisiae* VLPs and that immunostimulatory properties of the *Yarrowia* VLP-derived RNA are measurably lower than a corresponding IVT-produced RNA. The quantitative and qualitative superiority of the *Yarrowia* VLP-derived RNA over *S. cerevisiae* VLP- or IVT-derived RNA was both surprising and unexpected.

[0011] In a first aspect, the disclosure provides a *Yarrowia sp.* cell comprising: a first promoter operatively linked to a DNA sequence capable of being transcribed into a heterologous RNA molecule that comprises, in a 5' to 3' order, a 5'-UTR, a heterologous RNA payload, and a 3'-UTR, wherein the 3'-UTR comprises one or more tobacco mosaic virus (TMV) packaging signal; and a second promoter operatively linked to a DNA sequence encoding a TMV viral capsid protein.

[0012] In some embodiments of the first aspect, the *Yarrowia sp.* cell harbors a first vector comprising the DNA sequence capable of being transcribed into the heterologous RNA molecule; and a second vector comprising the DNA sequence encoding a TMV viral capsid protein. In alternate embodiments of the first aspect, the *Yarrowia sp.* cell harbors a vector comprising both the DNA sequence capable of being transcribed into the heterologous RNA molecule; and the DNA sequence encoding a TMV viral capsid protein.

[0013] In some embodiments of the first aspect, the heterologous RNA molecule contains only one TMV packaging signal. In alternate embodiments of the first aspect, the heterologous RNA molecule contains only two TMV packaging signals. In other alternate embodiments of the first aspect, the heterologous RNA molecule contains more than two TMV packaging signals.

[0014] In some embodiments of the first aspect, the heterologous RNA payload encodes a polypeptide or protein. In some embodiments of the first aspect, the heterologous RNA payload encodes a mammalian polypeptide or protein. In some embodiments of the first aspect, the heterologous RNA payload encodes a therapeutically useful mammalian polypeptide or protein. In other embodiments of the first aspect, the heterologous RNA payload is non-coding (e.g., it does not encode a polypeptide sequence).

[0015] In a second aspect, the disclosure provides a method of producing a heterologous RNA molecule comprising the step of growing the *Yarrowia sp.* cell of any of the embodiments of the first aspect under conditions in which the heterologous RNA molecule and the TMV capsid protein are expressed and, through the presence of a TMV packaging signal, the heterologous RNA molecule is bound to and encapsidated by the TMV viral capsid protein in a virus-like particle, thereby producing the heterologous mRNA molecule.

[0016] In some embodiments of the second aspect, the method comprises the further step of isolating the virus-like particle from the *Yarrowia sp.* cell.

[0017] In some further embodiments of the second aspect, the method further comprises the step of treating the isolated virus-like particle with one or more RNases or nucleases that digest double-stranded RNA.

[0018] In some further embodiments of the second aspect, the method further comprises the step of purifying the heterologous RNA from the isolated virus-like particle.

[0019] In a third aspect, the disclosure provides an isolated virus-like particle produced by the method of the second aspect.

[0020] In a fourth aspect, the disclosure provides an isolated heterologous RNA molecule produced by the method of the second aspect.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] FIG. 1 depicts a plasmid map of SC-GFP UTR-1, an expression vector used to express EGFP in *S. cerevisiae* and described in Table 3.

[0022] FIG. 2 depicts a plasmid map of SC-Luc OAS, an expression vector used to express firefly luciferase in *S. cerevisiae* and described in Table 3.

[0023] FIG. 3 depicts a plasmid map of SC-Luc UTR-4, an expression vector used to express firefly luciferase in *S. cerevisiae* and described in Table 3.

[0024] FIG. 4 depicts a plasmid map of SC-EPO UTR-4, an expression vector used to express erythropoietin in *S. cerevisiae* and described in Table 3.

[0025] FIG. 5 depicts a plasmid map of TMV-CP, an expression vector used to express TMV capsid protein in *S. cerevisiae* and described in Table 3.

[0026] FIG. 6 depicts a plasmid map of YL-TMV-GFP UTR-1, an expression vector used to express both EGFP and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0027] FIG. 7 depicts a plasmid map of YL-TMV-Luc OAS, an expression vector used to express both firefly luciferase and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0028] FIG. 8 depicts a plasmid map of YL-TMV-Luc UTR-4, an expression vector used to express both firefly luciferase and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0029] FIG. 9 depicts a plasmid map of YL-TMV-EPO UTR-4, an expression vector used to express both erythropoietin and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0030] FIG. 10 depicts a plasmid map of YL-Luc OAS, an expression vector used to express firefly luciferase in *Y. lipolytica* and described in Table 3.

[0031] FIG. 11 depicts a plasmid map of YL-Luc UTR-4, an expression vector used to express firefly luciferase in *Y. lipolytica* and described in Table 3.

[0032] FIG. 12 depicts a plasmid map of YL-TMV-CP, an expression vector used to express TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0033] FIG. 13A-FIG. 13B are bar graphs depicting the total amount of TMV capsid protein (FIG. 13A) and heterologous RNA (FIG. 13B) obtained from TMV VLPs produced in *S. cerevisiae* cells co-transformed with TMV-CP and one of SC-GFP UTR-1, SC-Luc UTR-4, or SC-EPO UTR-4, and in *Y. lipolytica* transformed YL-TMV-GFP UTR-1, or YL-TMV-EPO UTR-4, or co-transformed with YL -Luc UTR-4 and YL-TMV-CP. One asterisk represents a p-value of <0.05 between the indicated samples using the Welch's t test. Two asterisks represent a p-value of <0.01 between the indicated samples in the same test.

[0034] FIG. 14 depicts capillary gel electropherograms of RNA extracted from TMV VLPs produced in *S. cerevisiae* co-transformed with SC-GFP UTR-1 and TMV-CP (FIG. 14A) and *Y. lipolytica* transformed with YL-TMV-GFP UTR-1 (FIG. 14B).

[0035] FIG. 15 is a bar graph depicting the level of GFP expression in THP-1 derived macrophage cells transfected with GFP-encoding IVT-transcribed RNA from a commercial source, IVT transcribed RNA from SEQ ID NO:35, or VLP-derived RNA from *S. cerevisiae*

cells co-transformed with TMV-CP and SC-GFP UTR-1, or from *Y. lipolytica* cells transformed with YL-TMV-GFP UTR-1. Three asterisks indicate a p-value of <0.001 using the Welch's t test between the indicated samples.

[0036] FIG. 16A, is a bar graph depicting the level of firefly luciferase expression in THP-1 derived macrophage cells transfected with firefly luciferase-encoding IVT-transcribed RNA from a commercial source, IVT transcribed RNA from SEQ ID NO:33 ("IVT Luc UTR-4"), or VLP-derived RNA from *S. cerevisiae* cells co-transformed with TMV-CP and SC-Luc UTR-4, *Y. lipolytica* cells transformed with YL-TMV-Luc UTR-4 or *Y. lipolytica* cells co-transformed with TMV-CP and YP-Luc UTR-4. FIG 16B, is a bar graph depicting the level of firefly luciferase expression in THP-1 derived macrophage cells transfected with IVT transcribed RNA from SEQ ID NO:33 ("IVT Luc UTR-4"), IVT transcribed RNA from SEQ ID NO:34 ("IVT Luc-OAS"), or VLP-derived RNA from *Y. lipolytica* cells co-transformed with YL-TMV-CP and either YL-Luc UTR-4 or YL-Luc OAS. Four asterisks indicate a p-value of <0.0001 between the indicated samples using the Welch's t test.

[0037] FIG. 17 is a bar graph depicting the level of erythropoietin expression in THP-1 derived macrophage cells transfected with VLP-derived RNA from *S. cerevisiae* cells co-transformed with TMV-CP and SC-EPO UTR-4, or from *Y. lipolytica* cells transformed with YL-TMV-EPO UTR-4. One asterisk represents a p-value of <0.05 between the indicated samples using the Welch's t test.

[0038] FIG. 18 is a bar graph depicting the level of interferon regulatory factor-induced Lucia luciferase in THP-1 derived macrophage cells transfected with GFP-encoding IVT-transcribed RNA from a commercial source, IVT transcribed RNA from SEQ ID NO:35, or VLP-derived RNA from *S. cerevisiae* cells co-transformed with TMV-CP and SC-GFP UTR-1, or from *Y. lipolytica* cells transformed with YL-TMV-GFP UTR-1, as compared to a positive control ("Poly (dA:dT)"). Three asterisks indicate a p-value of <0.001 between the indicated samples using the Welch's t test.

[0039] FIG. 19 is a bar graph depicting the level of interferon-induced Lucia luciferase in THP-1 derived macrophage cells transfected with firefly luciferase-encoding IVT-transcribed RNA from a commercial source, IVT transcribed RNA from SEQ ID NO:33 ("IVT Luc UTR-4"), or VLP-derived RNA from *S. cerevisiae* cells co-transformed with TMV-CP and SC-Luc UTR-4, *Y. lipolytica* cells transformed with YL-TMV-Luc UTR-4, or *Y. lipolytica* cells co-transformed with TMV-CP and YP-Luc UTR-4, as compared to a positive control ("Poly (dA:dT)"). Three

asterisks indicate a p-value of <0.001 between the indicated samples using the Welch's t test. Two asterisks represent a p-value of <0.01 between the indicated samples in the same test.

[0040] FIG. 20 is a bar graph depicting the level of interferon-induced Lucia luciferase in THP-1 1 derived macrophage cells transfected with IVT transcribed RNA from SEQ ID NO:33 ("IVT Luc UTR-4"), IVT transcribed RNA from SEQ ID NO:34 ("IVT Luc-OAS"), or VLP-derived RNA from *Y. lipolytica* cells co-transformed with YL-TMV-CP and either YL-Luc UTR-4 or YL-Luc OAS, as compared to a positive control ("Poly (dA:dT)"). Three asterisks indicate a p-value of <0.001 between the indicated samples using the Welch's t test.

[0041] FIG. 21 is a bar graph depicting the level of interferon-induced Lucia luciferase in THP-1 cells transfected with VLP-derived RNA from *S. cerevisiae* cells co-transformed with TMV-CP and SC-EPO UTR-4, or from *Y. lipolytica* cells transformed with YL-TMV-EPO UTR-4, as compared to a positive control ("Poly (dA:dT)"). Two asterisks indicate a p-value of <0.01 between the indicated samples using the Welch's t test.

[0042] FIGS. 22A-22F. FIG. 22A, is a bar graph depicting the percent of GFP-expressing primary human PBMC-derived dendritic cells transfected with RNA isolated from VLPs produced from two separate batches of *Y. lipolytica* transformed with YL-TMV-GFP UTR-1 ("B1" and "B2") and a corresponding IVT-produced commercial RNA, as compared to a negative control. FIG. 22B-FIG. 22F are bar graphs depicting the amounts of CD86 (FIG. 22B), CD80 (FIG. 22C), CD54 (FIG. 22D), CD40 (FIG. 22E) and IL-6 (FIG. 22F) produced from primary human PBMC-derived dendritic cells transfected with RNA isolated from VLPs produced from two separate batches of *Y. lipolytica* transformed with YL-TMV-GFP UTR-1 ("B1" and "B2") and a corresponding IVT-produced commercial RNA, as compared to a negative control.

[0043] FIG. 23 depicts a plasmid map of plasmid 230, an expression vector used to express both luciferase and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0044] FIG. 24 depicts a plasmid map of plasmid 333, an expression vector used to express both luciferase and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0045] FIG. 25 depicts a plasmid map of plasmid 334, an expression vector used to express both luciferase and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0046] FIG. 26 depicts a plasmid map of plasmid 335, an expression vector used to express both GFP and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0047] FIG. 27 depicts a plasmid map of plasmid 336, an expression vector used to express both GFP and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0048] FIG. 28 depicts a plasmid map of plasmid 337, an expression vector used to express both GFP and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0049] FIG. 29 depicts a plasmid map of plasmid 338, an expression vector used to express both human erythropoietin and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0050] FIG. 30 depicts a plasmid map of plasmid 339, an expression vector used to express both human erythropoietin and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0051] FIG. 31 depicts a plasmid map of plasmid 340, an expression vector used to express both human erythropoietin and TMV capsid protein in *Y. lipolytica* and described in Table 3.

[0052] FIG. 32 depicts a plasmid map of plasmid 78, an expression vector used to express luciferase in *S. cerevisiae* and described in Table 3.

[0053] FIG. 33 is a bar graph depicting the level of GFP expression in THP-1 derived macrophage cells transfected with RNA isolated from VLPs produced by *Yarrowia lipolytica* transformed with Plasmid 336 ("YL_eGFP_AES") or Plasmid 335 ("YL-eGFP-mtRNR1") or VLPs produced in *Saccharomyces cerevisiae* harboring the TMV-capsid encoding SC-TMV-CP and transformed with Plasmid 341 ("SC_eGFP_mtRNR1") or Plasmid 342 ("SC_eGFP_AES"). Two asterisks indicate a p-value of <0.01 between the indicated samples using the Welch's t test.

[0054] FIG. 34 is a bar graph depicting the level of human EPO (hEPO) expression in THP-1 derived macrophage cells transfected with RNA isolated from VLPs produced by *Yarrowia lipolytica* transformed with Plasmid 339 ("YL_hEPO_AES") or Plasmid 338 ("YL-hEPO-mtRNR1") or VLPs produced in *Saccharomyces cerevisiae* harboring the TMV-capsid encoding SC-TMV-CP and transformed with Plasmid 344 ("SC_hEPO_mtRNR1") or Plasmid 345 ("SC_hEPO_AES"). Two asterisks indicate a p-value of <0.01 between the indicated samples using the Welch's t test.

[0055] Various objectives, features, and advantages of the disclosed subject matter can be more fully appreciated with reference to the following detailed description of the disclosed subject matter when considered in connection with the following drawings, in which like reference numerals identify like elements.

DETAILED DESCRIPTION

[0056] The details of one or more variations of the subject matter described herein are set forth in the accompanying drawings and the description below. Other features and advantages of the subject matter described herein will be apparent from the description and drawings, and from the

claims. The disclosed subject matter is not, however, limited to any particular embodiment disclosed.

[0057] Definitions

[0058] The term "DNA sequence capable of being transcribed to a RNA sequence" as used herein means a polydeoxyribonucleotide that serves as a template for transcription when contacted with an RNA polymerase into an RNA sequence. Such RNA polymerase may be added to the DNA sequence *in vitro* such as in in vitro translation or may be present in a host cell that harbors such DNA sequence. The RNA polymerase that performs such transcription and is present in a host cell may be naturally encoded by the host cell or may be recombinantly produced in a host cell that harbors such DNA sequence.

[0059] The term "5'-UTR" or "5'-untranslated region" as used herein means a polyribonucleotide sequence that is not translated into an amino acid sequence and is located 5' or upstream from the beginning of an RNA sequence that is the heterologous RNA payload. In embodiments when the heterologous RNA payload encodes a polypeptide, the beginning of a functional RNA sequence is a start codon and the 5'-UTR ends at the nucleotide preceding that start codon. The 5'UTR typically comprises secondary and tertiary structures and other sequence elements that regulate the stability of the transcribed RNA and, in some embodiments, the translatability of the heterologous RNA payload.

[0060] The term "3'-UTR" or "3'-untranslated region" as used herein means a polyribonucleotide sequence that is not translated into an amino acid sequence and is located 3' or downstream from the end of an RNA sequence that is the heterologous RNA payload. In some embodiments when the heterologous RNA payload encodes a polypeptide, the end of a functional RNA sequence is a termination codon and the 3'-UTR begins at the nucleotide immediately following that termination codon. The 3'-UTR typically comprises secondary and tertiary structures and other sequence elements that can impact the stability the transcribed RNA molecule and, in some embodiments, the translatability of the heterologous RNA payload. In some, but not all, embodiments, the 3'-UTR comprises a polyadenylation tail.

[0061] The term "heterologous RNA payload" as used herein means an RNA molecule that is not found naturally in Tobacco Mosaic Virus and that has activity at the RNA level itself (e.g., without being translated into a protein), or by encoding upon translation a polypeptide (e.g., as a mRNA). The heterologous mRNA payloads disclosed herein are intended for therapeutic, prophylactic, diagnostic or research use in a cell, cell lysate, *in vitro* translation system, a tissue, or an organism.

[0062] The term “TMV packaging signal” as used herein means an RNA sequence that can bind to a Tobacco Mosaic Virus capsid protein.

[0063] The term “proximal TMV packaging signals” as used herein refers to a pair of TMV packaging signals that are closest to one another in an RNA molecule. For clarity, an RNA molecule comprising three TMV packaging signals, designated as “1”, “2” and “3” in the structure: 1-2-3, wherein every two proximal TMV packaging signals are separated by less than XX nucleotides (wherein XX is an integer), means that “1” and “2” are separated from one another by less than XX nucleotides and that “2” and “3” are separated from one another by less than XX nucleotides. It does not mean or imply that “1” and “3” are separated from one another by less than XX nucleotides.

[0064] As used herein, the term “*Yarrowia sp.*” refers to any species of the genus *Yarrowia*.

[0065] As used herein, the term “operably linked” refers to a first molecule joined to a second molecule, wherein the molecules are so arranged that the first molecule affects the function of the second molecule. The two molecules may or may not be part of a single contiguous molecule and may or may not be adjacent. For example, a promoter is operably linked to a transcribable polydeoxyribonucleotide molecule if the promoter modulates transcription of the transcribable polydeoxyribonucleotide molecule payload in a cell. Additionally, two portions of a transcription regulatory element are operably linked to one another if they are joined such that the transcription-activating functionality of one portion is not adversely affected by the presence of the other portion.

[0066] Two transcription regulatory elements may be operably linked to one another by way of a linker nucleic acid (e.g., an intervening non-coding nucleic acid) or may be operably linked to one another with no intervening nucleotides present.

[0067] As used herein, the term "promoter" refers to a recognition site on DNA that is bound by an RNA polymerase. The polymerase drives transcription of the DNA to which such promoter is operably linked. The term “constitutive promoter” refers to a promoter whose ability to bind an RNA polymerase is not significantly affected by environmental or other factors.

[0068] “Percent (%) sequence identity” with respect to a reference polynucleotide or polypeptide sequence is defined as the percentage of nucleic acids or amino acids in a candidate sequence that are identical to the nucleic acids or amino acids in the reference polynucleotide or polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid or amino acid sequence identity can be achieved in various ways that are within the

capabilities of one of skill in the art, for example, using publicly available computer software such as BLAST, BLAST-2, or MegAlign software. Those skilled in the art can determine appropriate parameters for aligning sequences, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For example, percent sequence identity values may be generated using the sequence comparison computer program BLAST. As an illustration, the percent sequence identity of a given nucleic acid or amino acid sequence, A, to, with, or against a given nucleic acid or amino acid sequence, B, (which can alternatively be phrased as a given nucleic acid or amino acid sequence, A that has a certain percent sequence identity to, with, or against a given nucleic acid or amino acid sequence, B) is calculated as follows:

[0069] 100 multiplied by (the fraction X/Y)

[0070] where X is the number of nucleotides or amino acids scored as identical matches by a sequence alignment program (e.g., BLAST) in that program's alignment of A and B, and where Y is the total number of nucleic acids in B. It will be appreciated that where the length of nucleic acid or amino acid sequence A is not equal to the length of nucleic acid or amino acid sequence B, the percent sequence identity of A to B will not equal the percent sequence identity of B to A.

[0071] As used herein, the term "vector" includes a nucleic acid vector, e.g., a DNA vector, such as a plasmid, cosmid, or artificial chromosome, an RNA vector, a virus, or any other suitable replicon (e.g., viral vector). A variety of vectors have been developed for the delivery of polynucleotides encoding exogenous proteins into a prokaryotic or eukaryotic cell. Examples of such expression vectors are described in, e.g., Gellissen, Production of Recombinant Proteins: Novel Microbial and Eukaryotic Expression Systems (John Wiley & Sons, Marblehead, MA, 2006). Expression vectors suitable for use herein contain a polydeoxyribonucleotide sequence that can be transcribed into RNA, as well as, e.g., additional sequence elements used for such transcription. Certain vectors that can be used for the transcription of RNA as described herein include vectors that contain regulatory sequences, such as promoter and enhancer regions, which direct transcription. The expression vectors suitable for use herein may also contain a polynucleotide encoding a marker for selection of cells that contain such a vector. Examples of a suitable marker include genes that encode resistance to antibiotics, such as ampicillin, chloramphenicol, kanamycin, or nourseothricin.

[0072] Overview

[0073] The present disclosure provides *Yarrowia sp.* cells capable of producing heterologous RNA payloads, methods for such production, and the resulting RNA payloads. According to the methods described herein, the RNA payloads are contained within tobacco mosaic virus (TMV) viral-like particles (VLPs) produced by the *Yarrowia sp.* cells. The VLPs are composed of TMV viral capsid proteins that self-assemble, mimicking the structure of the native virus from which they are derived. VLPs are non-infectious as they lack genetic material required for integration, reverse transcription, replication and/or infection. Isolated VLPs comprising RNA payloads are also provided herein.

[0074] DNA and RNA Sequences

[0075] The *Yarrowia sp.* cell disclosed herein comprises: a first promoter operatively linked to a DNA sequence capable of being transcribed into a heterologous RNA molecule that comprises, in a 5' to 3' order, a 5'-UTR, a heterologous RNA payload, and a 3'-UTR, wherein the 3'-UTR comprises one or more tobacco mosaic virus (TMV) packaging signals; and a second promoter operatively linked to a DNA sequence encoding a TMV viral capsid protein. Applicant has discovered that using a *Yarrowia sp.* cell, and particularly a *Y. lipolytica* cell, results in unexpected improvements over similar transcribed RNA sequences produced in other yeast, such as *Saccharomyces cerevisiae*. Such improvements include one or more of: improved quantity of TMV VLPs produced, improved quantity of RNA recovered from VLPs, and improved quality of RNA recovered from VLPs (in terms of quantity of protein transcribed in cells transfected with such RNA). Also, the RNA produced in *Yarrowia sp.* described herein showed superiority to *in vitro* transcription (IVT)-produced RNA in one or both quantity of protein translated in cells transfected with such RNA and reduced immunogenicity of such RNA is transfected cells.

[0076] *5'-UTR*

[0077] As set forth above, the RNA transcripts disclosed herein comprise a 5'-UTR. The 5'-UTR sequences disclosed herein are a region of the transcribed RNA that is not translated into an amino acid sequence, nor necessarily directly involved in the activity of the heterologous RNA payload, but include regulatory elements that modulate the stability of the transcribed RNA and its translation (e.g., if the RNA payload encodes a polypeptide). The 5' UTR typically begins at the transcription start site and continues to just before the beginning of the functional RNA sequence (e.g., the start codon when the heterologous RNA payload encodes a polypeptide).

[0078] Naturally occurring RNA transcripts typically comprise a 5'-UTR. In some embodiments, the 5' UTR comprises a nucleic acid sequence that is naturally present in the transcript of the encoded heterologous RNA payload. In some embodiments, the 5' UTR comprises a nucleic acid sequence that is not naturally associated with the encoded heterologous mRNA payload. In some embodiments, the 5' UTR comprises a nucleic acid sequence that has high sequence identity (e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%) to the nucleic acid sequence that is naturally present in the transcript of the encoded heterologous RNA payload, wherein those non-identical nucleotides impart a quantitative or qualitative improvement to the transcript (e.g., in terms of stability, function and/or translatability). Modifications to naturally occurring 5'-UTRs that result in such improvements are known in the art.

[0079] As a non-limiting example, the 5'-UTR can be a human 5'-UTR and the heterologous RNA payload can be a human RNA payload. In some embodiments, the 5' UTR can be heterologous to the cell used to produce the VLPs. For example, the 5'-UTR and the RNA payload can be human whereas the cell used to produce the VLP comprising such RNA payload is a non-human cell e.g., a mouse cell, a rat cell, a hamster cell, an insect cell, a yeast cell, a bacterial cell, or a plant cell.

[0080] The 5' UTR can have a sequence or structure that alters the level of translation (e.g., increases, decreases, or prevents) of the RNA payload into a polypeptide. In some embodiments, the 5'-UTR has a sequence or structure that promotes a cell type-specific translation of the RNA payload.

[0081] The 5' UTR can include one or more ribosome binding sequences. As used herein, the term "ribosome binding sequence" refers to a nucleic acid sequence that is recognized by the ribosome or a self-cleaving ribozyme, and upon binding to the sequence, can initiate translation of the RNA payload into a polypeptide. In some embodiments, a ribosome binding sequence can be a Kozak sequence. In some embodiments, the 5' UTR includes permuted introns to mediate RNA self-circularization.

[0082] The 5'-UTR sequences disclosed herein may differ from the resulting 5'-UTR in the transcribed RNA by the 5' extension and/or modification of the encoded 5'-UTR. In some embodiments, when the DNA sequence encoding the 5'-UTR disclosed herein is placed into an expression vector for ultimate transcription in a host cell, the expression vector itself may template additional ribonucleotides that are added to the 5'-end of the encoded 5'-UTR. In some embodiments, the 5'-UTR may be capped. Capping refers to a modification of the first

(5'-most) transcribed nucleotide of an RNA transcript by enzymes present in the cell producing such transcript. Capping typically introduces a methylated guanosine onto that first transcribed nucleotide. Capping blocks 5′-3′ exonuclease-mediated degradation and recruits specific RNA processing, export and translation factors. The presence of a cap and the specific modification of the cap is dependent upon the cell in which the isolated DNA sequences disclosed herein are transcribed. In some embodiments, the cap can be a 7-methylguanosine (m7G) cap. In some embodiments the 2' ribose position of the first cap-proximal nucleotide of the 5'-UTR is methylated to form a Cap 1 structure. In some embodiments, the 5'-UTR includes a methylation in the N6-position of adenosine (also herein N6-methyladenosine or m6A). In some embodiments, the 2' ribose position of the second cap-proximal nucleotide of the 5'-UTR is methylated to form a Cap 2 structure. Capping of the 5'-UTR in the population of cells transcribing the isolated DNA sequences disclosed herein (e.g., the cells producing the VLPs) can be achieved by native enzymes in such population of cells and/or by concurrent recombinant expression of capping enzymes or polynucleotides encoding capping enzymes e.g., a 2'O methyl-transferase enzyme. Modification of the 5'-UTR cap may also be achieved after the RNA is isolated (i.e., from VLPs) through treatment with isolated methyl transferases.

[0083] In some embodiments, the 5'-UTR is an RNA sequence comprising the following nucleotide sequence:

GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCC ACC (SEQ ID NO:1).

[0084] 3'-UTR

[0085] As set forth above, the RNA transcripts disclosed herein comprise a 3'-UTR. The 3'-UTR is a region of the resulting transcribed RNA that is not translated into an amino acid sequence, nor necessarily directly involved in the activity of the heterologous RNA payload but includes regulatory elements that modulate the stability of the transcribed RNA and its translation (e.g., if the RNA payload encodes a polypeptide). The 3' UTR typically begins immediately after the sequence encoding the heterologous RNA payload (e.g., the termination codon when the heterologous RNA payload encodes a polypeptide) and continues to the end of the transcribed RNA.

[0086] Naturally occurring RNA transcripts typically comprise a 3'-UTR. In some embodiments, the 3' UTR disclosed herein comprises a nucleic acid sequence that is naturally present in the transcript of the encoded heterologous RNA payload. In some embodiments, the 3' UTR comprises a nucleic acid sequence that is not naturally associated with the encoded heterologous mRNA payload. In some embodiments, the 3' UTR comprises a nucleic acid

sequence that has high sequence identity (e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%) to the nucleic acid sequence that is naturally present in the transcript of the encoded heterologous RNA payload, wherein those non-identical nucleotides impart a quantitative or qualitative improvement to the transcript (e.g., in terms of stability, function and/or translatability). Modifications to naturally occurring 3'-UTRs that result in such improvements are known in the art. Some non-limiting naturally occurring 3'-UTR sequences (or core motifs therefrom) are set forth in Supplemental Table 3 of A.G. Orlandini von Niessen et al., Mol Ther 2019, 27(4), 724-836.

[0087] In some embodiments, a 3'-UTR includes sequence elements mediating binding to proteins involved in mRNA trafficking, and/or translation. In some embodiments, a 3'-UTR can include sequences repressing the deadenylation of mRNA. In some embodiments, a 3'-UTR can include a polyadenylation signal sequence. In some embodiments, a 3'-UTR can include a polyadenylation tail. In some embodiments, a 3'-UTR can include a nucleotide sequence repressing the deadenylation of mRNA. As a non-limiting example, a 3'-UTR can include one or more regions of, or the entire 3′-UTR of human α-globin, which corresponds to the RNA sequence:

GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCC
UCCCCUUCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
(SEQ ID NO:2)
or a variant thereof:
GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCC
UCCCCUUCCUGCACCCGAUGGGGGCACAGAAACUUAUUUCAGACUCACCCGUACC
CCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO:3)
with the one or more TMV packaging signals located at either end of and/or interspersed within such 3'-UTR and optionally containing one or more poly(A) tracts of varying lengths.

[0088] In some embodiments, a 3'-UTR can include one or more regions of, or the entire 3′-UTR of the Mitochondrial-Encoded 12S RNA gene (mtRNR1), in particular the core motif therefrom having the RNA sequence:

CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAA
ACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAAC
CCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC (SEQ ID NO:39)
with the one or more TMV packaging signals located at either end of and/or interspersed within such 3'-UTR. In some specific embodiments, the 3'-UTR comprises one TMV packaging signal located 5' to the mtRNR1 3'-UTR.

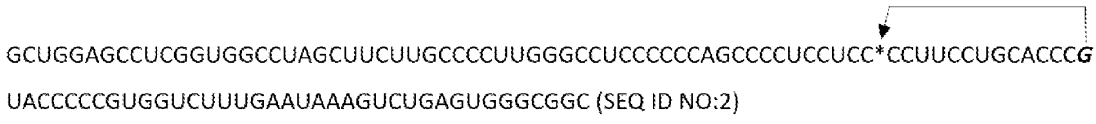
[0089] In some embodiments, a 3'-UTR can include one or more regions of, or the entire 3′-UTR of the amino-terminal enhancer of split gene (AES), in particular the core motif therefrom having the RNA sequence:

CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGA
GUCUCCCCCGACCUCGGGUCCCAGGUAUGCUCCCACCUCCACCUGCCCCACUCACC
ACCUCUGCUAGUUCCAGACACCUCC
(SEQ ID NO:40)
with the one or more TMV packaging signals located at either end of and/or interspersed within such 3'-UTR. In some specific embodiments, the 3'-UTR comprises one TMV packaging signal located 5' to the AES 3'-UTR.

[0090] In some embodiments, a 3'-UTR can include one or more regions of, or the entire 3′-UTR of the magnesium transporter MRS2 gene (MRS2), , in particular the core motif therefrom having the RNA sequence:

CUGCUGCCUGCUUCUUGCUCCAGCACCAUGGAAUGCCUGCGCAGUUUACCCUGCC
UCCUGCCCCGCGCGAUGAGACUUCCCCGGCGGACGCUGUGUGCCCUGGCCUUGGA
CGUGACCUCUGUGGGUCCUCCCGUUGCUGCCU
(SEQ ID NO:41)
with the one or more TMV packaging signals located at either end of and/or interspersed within such 3'-UTR. In some specific embodiments, the 3'-UTR comprises one TMV packaging signal located 5' to the MRS2 3'-UTR.

[0091] It will be understood that in creating DNA encoding TMV packaging signal-containing 3'-UTRs, especially where the TMV packaging signal is inserted within the parent 3'-UTR (e.g., the 3'-UTR of SEQ ID NO:2 or SEQ ID NO:3), cloning artifacts may be introduced where one nucleotide of the parent 3'-UTR is deleted or shifted in location, or where one additional nucleotide is inserted the parent 3'-UTR. For example, in SEQ ID NOs:4 and 5, below, the G nucleotide located 5' to the first TMV packaging sequence is shifted compared to the native α-globin 3'-UTR of SEQ ID NO:2, from which it is derived. The location of that G in SEQ ID NO:2 is indicated by italics and bolding below. In SEQ ID NOs:4 and 5, that G is shifted as shown by the arrow to the gap indicated with an asterisk below:



UACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (SEQ ID NO:2)

Applicant believes that this one nucleotide difference has minimal effect on 3'-UTR function.

[0092] In some embodiments, the 3'-UTR is an RNA sequence comprising any of the following nucleotide sequences in Table 1 (SEQ ID NOs: 4-15 or 42-44) or an RNA sequence having at least 90% sequence identity to any one of SEQ ID NOs:4-15 or 42-44 and comprising at least one copy of the minimal TMV packaging sequence of SEQ ID NO:16. In Table 1, bolded

nucleotides are the portions of the RNA sequence that correspond to the TMV packaging signals.

[0093] Table 1. 3'-UTR Sequences Containing TMV Packaging Signal(s)

SEQ ID NO:	Sequence
4	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGC CCCUCCUCCG**UUGUUUAUAGAAAUAAUAUAAAAUUAGGUUUGAGAGA GAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGAA GAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUCAGGC UUGCAAAGUUUCGAUCUCGAACCGGAAAAAAGAGUGAUGUCCGCAA AGGGAAAAAUA**ACCUUCCUGCACCC**UUGUUUAUAGAAAUAAUAUAAAA UUAGGUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCA UGGAACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCC UAUGUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGGAAAAAAG AGUGAUGUCCGCAAAGGGAAAAAUA**UACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGCAAAAAAAAA
5	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGC CCCUCCUCCG**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGA UGUCCCUAUGUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**C CUUCCUGCACCC**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGA GGGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAG AUGUCCCUAUGUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG** UACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGCAAAAAAAAA
6	**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUG UCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**GCUGGAGCCUCG GUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGAUGGGGGCACAGAAACUUAUUUCAGACUCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGCAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA**GUUUGAGAGAGAAGAUUACAAACGUG AGAGACGGAGGGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAAAGUUUCGAUC UCGAACCGG**AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO:	Sequence
7	**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUG UCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**GCUGGAGCCUCG GUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGAUGGGGGCACAGAAACUUAUUUCAGACUCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC**GUUUGAGAGAGA AGAUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGAAGA AGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUU GCAAAGUUUCGAUCUCGAACCGG**GCUGGAGCCUCGGUGGCCUAGCUU CUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGA UGGGGGCACAGAAACUUAUUUCAGACUCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGCAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAA
8	**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUG UCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**GCUGGAGCCUCG GUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGAUGGGGGCACAGAAACUUAUUUCAGACUCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC**GUUUGAGAGAGA AGAUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGAAGA AGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUU GCAAAGUUUCGAUCUCGAACCGG**
9	**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUG UCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**GCUGGAGCCUCG GUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGAUGGGGGCACAGAAACUUAUUUCAGACUCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC**GUUUGAGAGAGA AGAUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGAAGA AGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUU GCAAAGUUUCGAUCUCGAACCGG**AAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA
10	**UUGUUUAUAGAAAUAAUAUAAAAUUAGGUUUGAGAGAGAAGAUUAC AAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGAAGAAGUCGUU GAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAAAGU UUCGAUCUCGAACCGGAAAAAAGAGUGAUGUCCGCAAAGGGAAAAA UA**CCUUCCUGCACCC**UUGUUUAUAGAAAUAAUAUAAAAUUAGGUUUG AGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUUA CAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAU CAGGCUUGCAAAGUUUCGAUCUCGAACCGGAAAAAAGAGUGAUGUC CGCAAAGGGAAAAAUA**GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCC UUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGAUGGGGGCA CAGAAACUUAUUUCAGACUCACCCGUACCCCCGUGGUC1UUUGAAUAA AGUCUGAGUGGGCGGC

SEQ ID NO:	Sequence
11	**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUG UCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**CCUUCCUGCACC C**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGG AACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAU GUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**GCUGGAGCCU CGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCC CUUCCUGCACCCGAUGGGGGCACAGAAACUUAUUUCAGACUCACCCGU ACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
12	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA**GUUUGAG AGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUUACA GAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUCA GGCUUGCAAAGUUUCGAUCUCGAACCGG**CCUUCCUGCACCC**GUUUGA GAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUUAC AGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUC AGGCUUGCAAAGUUUCGAUCUCGAACCGG**AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAA
13	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGC CCCUCCUCCCCUUCCUGCACCCG**GUUUGAGAGAGAAGAUUACAAACGU GAGAGACGGAGGGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAG UUCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAAAGUUUCGAU CUCGAACCGG**UACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG CAAAAAAAAA
14	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGC CCCUCCUCCCCUUCCUGCACCCG**UUGUUUAUAGAAAUAAUAUAAAAUU AGGUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUG GAACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUA UGUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGGAAAAAAGAG UGAUGUCCGCAAAGGGAAAAAUA**UACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGCAAAAAAAAA
15	**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUG UCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**GCUGGAGCCUCG GUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCU UCCUGCACCCGAUGGGGGCACAGAAACUUAUUUCAGACUCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
42	**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUG UCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**CAAGCACGCAGC AAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGC AGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUA ACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC

SEQ ID NO:	Sequence
43	**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUG UCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**CUGGUACUGCA UGCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCU CCCCCGACCUCGGGUCCCAGGUAUGCUCCCACCUCCACCUGCCCCACUC ACCACCUCUGCUAGUUCCAGACACCUCC
44	**GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUG UCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG**CUGCUGCCUGCU UCUUGCUCCAGCACCAUGGAAUGCCUGCGCAGUUUACCCUGCCUCCUG CCCCGCGCGAUGAGACUUCCCCGGCGGACGCUGUGUGCCCUGGCCUUG GACGUGACCUCUGUGGGUCCUCCCGUUGCUGCCU

[0094] Heterologous RNA Payload

[0095] As set forth above, the transcribed RNA sequences disclosed herein comprise a heterologous RNA payload.

[0096] RNA payload can be of any length, for example at least 9 nucleotides in length. For example, the RNA payload length can be at least 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 10000, 15000, 20000, 25000, 30000, 35000, 40000, 45000, or 50000 nucleotides in length. In some embodiments, the RNA payload can be fewer than 50000, 45000, 40000, 35000, 30000, 25000, 20000, 15000, 10000, 5000, 4500, 4000, 3500, 3000, 2500, 2000, 1500, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100, 90, 80, 70, 60, 50, 40, 30, 20, 10 nucleotides in length. In some embodiments, the RNA payload can be 10-20, 15-25, 20-30, 25-35, 30 - 35, 35 - 40, 40 - 45, 45 - 50, 50 - 55, 55 - 60, 60 - 65, 65 - 70, 70 - 75, 75 - 80, 80 - 85, 85 - 90, 90 - 95, 95 - 100, 100 - 150, 150 - 200, 200 - 250, 250 - 300, 300 - 350, 350 - 400, 400 - 450, 450 - 500, 500 - 550, 550 - 600, 600 - 650, 650 - 700, 700 - 750, 750 - 800, 800 - 850, 850 - 900, 900 - 950, 950 - 1000, 1000 - 1500, 1500 - 2000, 2000 - 2500, 2500 - 3000, 3000 - 3500, 3500 - 4000, 4000 - 4500, 4500 - 5000, 5000 - 10000, 10000 - 15000, 15000 - 20000, 20000 - 25000, 25000 - 30000, 30000 - 35000, 35000 - 40000, 40000 - 45000, or 45000 – 50000 nucleotides in length.

[0097] The RNA payload can be derived from a naturally occurring RNA or can be a synthetic or chimeric RNA.

[0098] The RNA payload can be a eukaryotic RNA payload or a prokaryotic RNA payload. In some embodiments, the eukaryotic RNA payload can be a mammalian RNA payload. As a non-limiting example, the mammalian RNA payload can be a human RNA payload. In some

embodiments, the RNA payloads can be homologous to the population of cells used to produce the VLPs containing the RNA payload. In some embodiments, the RNA payloads can be heterologous to the population of cells used to produce the VLPs containing the RNA payloads. As used herein, "homologous" refers to two entities that are derived from the same species, whereas "heterologous," refers to two entities that are derived from different species. For example, the RNA payload can be a human RNA payload and the population of cells can be a human population of cells (e.g., a human-derived cell line). As another example, the RNA payload can be a human RNA payload and the population of cells can be a hamster population of cells (e.g., a hamster-derived cell line, e.g., CHO cells).

[0099] In some embodiments, the RNA payload can be self-amplifying RNA payload, which can include a sequence of nucleotides coding for nonstructural proteins. The nonstructural proteins can be, for example, nsP1, nsP2, nsP3 and/or nsP4 of alphaviruses.

[0100] Coding RNA payload

[0101] The RNA payload can be a coding RNA payload. As used herein a "coding RNA payload" refers to an RNA payload that is capable of being translated to an encoded protein, e.g., *in vitro*, *in vivo*, *in situ,* or *ex vivo*. A coding RNA payload will begin with a start codon and end with a stop codon sequence. In some embodiments, the coding RNA payload includes introns. Introns can aid in nuclear export of the ORF payload. In one embodiment, the ORF payload includes permuted introns for production of circular RNA.

[0102] The coding RNA payload can be of any length. In some embodiments, the coding RNA payload can be at least 30 nucleotides in length. In some embodiments, the coding RNA payload can be at least 300 nucleotides in length. In some embodiments, the coding RNA payload can be at least 4500 nucleotides in length. For example, the coding RNA payload length can be at least 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 10000, 15000, 20000, 25000, 30000, 35000, 40000, 45000, or 50000 nucleotides in length. In some embodiments, the coding RNA payload can be fewer than 50000, 45000, 40000, 35000, 30000, 25000, 20000, 15000, 10000, 5000, 4500, 4000, 3500, 3000, 2500, 2000, 1500, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100, 90, 80, 70, 60, 50, 40, or 30 nucleotides in length. In some embodiments, the coding RNA payload can be 30 - 35, 35 - 40, 40 - 45, 45 - 50, 50 - 55, 55 - 60, 60 - 65, 65 - 70, 70 - 75, 75 - 80, 80 - 85, 85 - 90, 90 - 95, 95 - 100, 100 - 150, 150 - 200, 200 - 250, 250 - 300, 300 - 350, 350 - 400, 400 - 450, 450 - 500, 500 - 550, 550 - 600, 600 - 650, 650 - 700, 700 - 750, 750 - 800, 800 - 850, 850 - 900, 900 - 950, 950 - 1000, 1000 - 1500, 1500 - 2000, 2000 - 2500, 2500 - 3000, 3000 - 3500, 3500 - 4000,

4000 - 4500, 4500 - 5000, 5000 - 10000, 10000 - 15000, 15000 - 20000, 20000 - 25000, 25000 - 30000, 30000 - 35000, 35000 - 40000, 40000 - 45000, or 45000 – 50000 nucleotides in length.

[0103] In some embodiments, the coding RNA payload is not translated or minimally translated to a polypeptide or a protein sequence in a population of cells that are used to prepare VLPs containing the coding RNA payload.

[0104] A coding RNA payload can encode, for example, a bacterial protein or polypeptide, a genome editing enzyme (e.g., Cas9, transposase, meganuclease, Base Editor or retrotransposon), a prime editor, a DNA writer, a virus protein or polypeptide, a eukaryotic transcription factor (e.g., a mammalian transcription factor), a growth factor, an antibody, a polypeptide containing one or more viral or a bacterial epitopes or antigens, an enzyme, or a therapeutically useful mammalian polypeptide (e.g., a therapeutically useful human polypeptide). In some embodiments, the coding RNA payload encodes a neo-antigen. In some embodiments, the neo-antigen is a mammalian cancer neo-antigen. In some embodiments, the coding RNA payload encodes a therapeutically useful human polypeptide. In some embodiments, the coding RNA payload encodes a therapeutically useful human enzyme. In some embodiments, the coding RNA payload encodes a therapeutically useful human enzyme that can be used to treat a disease or condition characterized by reduced activity of the native form of such enzyme.

[0105] In some embodiments, a coding RNA payload can encode a polypeptide useful in immunotherapy. In some embodiments, the polypeptide can be a chimeric antigen receptor (CAR), such as an anti-c-Met CAR or an anti-CD19 CAR. In some embodiments, the coding RNA payload can encode an immunomodulatory protein, such as a cytokine, or a costimulatory molecules polypeptide. In some embodiments, the coding RNA payload can encode a polypeptide that serves as a vaccine for immunotherapy. For example, the polypeptide can be a full-length or a portion of a protein such as p53, survivin, gp100, tyrosinase, mTRP-2, MAGE-A3, MAGE-C2, PSCA, PSMA, PSA, STEAP1, NY-ESO-1, 5T4, MAGE-C1, telomerase, WT1, hTERT, and/or pp65.

[0106] In some embodiments, a coding RNA payload can encode a protein, or a polypeptide for protein replacement in a genetic disorder, a lung disease, a hematological disease, an orthopedic disease, a neurogenic disease, a metabolic disorder, and/or a cardiovascular disease. A coding RNA payload can code for a methyl malonyl-CoA mutase, propionyl-CoA carboxylase, ornithine transcarbamylase, alpha 1-antitrypsin, CFTR, Factor VIII, Factor IX, VEGFA, and/or a region or a portion thereof.

[0107] In some embodiments, the coding RNA payload can encode a gene editing nuclease, e.g., zinc finger nuclease, transcription activator-like effector nuclease, a Base Editor, a prime editor a DNA writer, and/or the clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein nuclease system.

[0108] In some embodiments, the coding RNA payload can encode a polypeptide related to the treatment of an infectious disease. A coding RNA payload can encode a vaccine for infectious diseases. For example, a coding RNA payload can encode a polypeptide useful as a COVID-19 vaccine. Such polypeptide can contain the entire protein or a portion thereof such as, but not limited to, S-2P (and variants thereof), receptor binding domain, trimerized receptor binding domain, and/or transmembrane spike (S) protein. In some embodiments, a coding RNA payload can encode a rabies vaccine comprising Rabies virus glycoprotein. In some embodiments, a coding RNA payload can encode an influenza vaccine comprising the Hemagglutinin glycoprotein or a portion thereof. In some embodiments, a coding RNA payload can encode a HIV vaccine comprising one or more of HIV Gag, Nef, Vpr, Rev proteins or a portion thereof. In some embodiments, a coding RNA payload can encode an anti-Chikungunya antibody. As a non-limiting example, a coding RNA payload can be the SARS-CoV-2 spike protein or a region or a portion thereof.

[0109] In some embodiments, RNA payload can encode a bacterial antigen.

[0110] In some embodiments, the RNA payload can encode an immunoglobulin (IgG) for use in passive immunotherapy.

[0111] In some embodiments, the sequence of coding RNA payloads can be optimized to promote the translation of the coding RNA payload in a cell, population of cells, or organism intended to be administered the RNA once that RNA is isolated from the cells used for VLP production. In some embodiments, the sequence of coding RNA payloads can be optimized to increase the half-life of the RNA payload. In one embodiment, the coding RNA payload is codon optimized or uridine depleted to reduce immunogenicity in a cell, population of cells, or organism intended to be administered the RNA.

[0112] *Non-coding RNA payload*

[0113] RNA payloads useful herein can be non-coding RNA payloads. RNA molecules that are not translated to an encoded peptide, polypeptide, or protein are referred to herein as non-coding RNAs. The non-coding RNA payloads can be a short non-coding RNA payload or a long non-coding RNA payload. The short non-coding RNA payload can be 200, 150, 100, 75, 50, 25 nucleotides or less in length. Non-limiting examples of short non-coding RNA include micro-

RNA (miRNA), small interfering RNA (siRNA), piwi-interacting RNA (piRNA), small Cajal body-specific RNAs (scaRNAs), trans-activating CRISPR (tracr) RNA, small nuclear RNA (snRNA), U-rich snRNA (snRNA), small nucleolar RNA (snoRNA), and/or Y RNAs.

[0114] The long non-coding RNAs can be greater than 200 nucleotides in length (e.g., about 200, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,500, 2,000, 2,500, 3,000, 3,500, 4,000, 4,500 or more). Non-limiting examples of long non-coding RNA include long intergenic ncRNA (lincRNA), natural antisense transcript (NAT) RNA, and/or circular RNA (circRNA). In one embodiment, the RNA payload is RNA with permuted introns to circularize the RNA to form a circRNA.

[0115] The non-coding RNAs include but are not limited to, transfer RNAs (tRNA), ribosomal RNAs (rRNA), small nuclear RNAs (snRNA), small nucleolar RNAs (snoRNA), signal recognition particle RNA (SRP RNA), antisense RNA (aRNA), micro-RNA (miRNA), small interfering RNA (siRNA), Y RNA, telomerase RNA, and/or piwi-interacting RNA (piRNA), RNAs encoding RNA-dependent RNA polymerase complexes for self-amplification.

[0116] The RNA payload can be a single-stranded RNA payload or a double-stranded RNA payload. In one aspect, the RNA payload is single-stranded. Coding RNA payloads are typically single-stranded.

[0117] *TMV Packaging Signal*

[0118] As set forth above, the transcribed RNA sequences disclosed herein comprise one or more tobacco mosaic virus (TMV) packaging signals present in the 3'-UTR. In TMV, the packaging signal is often referred to in the prior art as the origin of assembly sequence, or OAS. The role of the TMV packaging signals in the transcribed RNA molecules herein is to bind to TMV capsid protein expressed in the same cell as the transcribed RNA, which allows that RNA to be encapsidated by TMV VLPs produced in such cells from the TMV capsid protein. It is believed that the packaging signals may also play a role in VLP assembly and formation.

[0119] Each of the one or more TMV packaging signals useful herein may independently range in size from 75 to about 500 nucleotides in length, 75 to 300 nucleotides 75-250 nucleotides, 100-250 nucleotides, or 100-200 nucleotides in length and each will comprise the nucleotides corresponding to the following RNA sequence:

UGAGAGACGGAGGGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGUUCAUGG AAGAUGUCCCUAUGUCGAUCA (SEQ ID NO:16).

[0120] SEQ ID NO:16 is a 75 ribonucleotide sequence recognized as the minimal sequence necessary to bind to TMV coat protein and allow assembly of VLPs. In some embodiments, the sequence of the TMV packaging has 100% sequence identity to a sequence present in naturally

occurring TMV. In some embodiments, each of the one or more TMV packaging signals independently corresponds to one of SEQ ID NOs:16-20 set forth in Table 2, below, or a nucleotide sequence comprising SEQ ID NO:16 and having at least 90% sequence identity to any one of SEQ ID NOs:16-20.

[0121] Table 2. TMV Packaging Sequences

SEQ ID NO.	RNA Sequence
16	UGAGAGACGGAGGGCCCAUGGAACUUACAGAAGAAGUCGUUGAU GAGUUCAUGGAAGAUGUCCCUAUGUCGAUCA
17	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUG GAACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCC UAUGUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG
18	UUGUUUAUAGAAAUAAUAUAAAAUUAGGUUUGAGAGAGAAGAU UACAAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGAAGAAG UCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUU GCAAAGUUUCGAUCUCGAACCGGAAAAAAGAGUGAUGUCCGCAA AGGGAAAAAUA
19	AUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGAAGA AGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUCAGGC UUGCAAAGUUU
20	GCAAGUUUUAGUUAAUAUUAGAAAUGUGAAGAUGUCAGCGGGUU UCUGUCCGCUUUCUCUGGAGUUUGUGUCGGUGUGUAUUGUUUAU AGAAAUAAUAUAAAAUUAGGUUUGAGAGAGAAGAUUACAAACGU GAGAGACGGAGGGCCCAUGGAACUUACAGAAGAAGUCGUUGAUG AGUUCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAAAGUUU CGAUCUCGAACCGG

[0122] In some embodiments, each of the one or more TMV packaging signals independently is one of SEQ ID NOs:16-20. In some embodiments, each of the one or more TMV packaging signals is the same and is one of SEQ ID NOs: 16-20. In some embodiments, each of the one or more TMV packaging signals is the same and is SEQ ID NO:16. In some embodiments, each of the one or more TMV packaging signals is SEQ ID NO:17. In some embodiments, each of the one or more TMV packaging signals is SEQ ID NO:18. In some embodiments, each of the one or more TMV packaging signals is SEQ ID NO:19. In some embodiments, each of the one or more TMV packaging signals is SEQ ID NO:20. In some embodiments the RNA contains only one TMV packaging signal. In some embodiments the RNA contains only two TMV packaging signals.

[0123] In some embodiments where two or more TMV packaging signals are present, every two proximal TMV packaging signals in the 3'-UTR are separated by less than 500, less than 450, less than 400, less than 350, less than 300, less than 250, less than 200, less than 150, less than 100, less than 90, less than 80, less than 70, less than 60, less than 50, less than 40, less than 30,

less than 20, or less than 15, less than 10, or less than 5 nucleotides. It will readily apparent that the maximum distance between two proximal TMV packaging signals can be no greater than the length of the 3'-UTR within which they reside.

[0124] Expression Vectors, Promoters and Terminators

[0125] The first promoter operatively linked to a DNA sequence capable of being transcribed into a heterologous RNA molecule is required to allow the *Yarrowia sp.* cell harboring that DNA sequence to transcribe RNA therefrom. The promoter can be a strong, weak, constitutive, inducible, cell type specific promoter, or a tissue specific promoter. The promoter can be a constitutive promoter or an inducible promoter. The selection of the promoter can be optimized for *Yarrowia sp.* and/or the nature of the heterologous RNA payload.

[0126] Non-limiting examples of a constitutive promoter includes, TDH3 promoter, CCW12 promoter, PGK1 promoter, HHF2 promoter, TEF1 promoter, TEF2 promoter, YEF3 promoter, RPL3 promoter, RPL15A promoter, RPL4 promoter, RPL8B promoter, PDA1 promoter, HHF1 promoter, HTB2 promoter, RPL18B promoter, ALD6 promoter, PAB1 promoter, RET2 promoter, RNR1 promoter, SAC6 promoter, RNR2 promoter, POP6 promoter, RAD27 promoter, PSP2 promoter, TPI1 promoter, ENO2 promoter, ADH1 promoter, CYC1 promoter, GAP promoter, GPD promoter, REV1 promoter, SV40 promoter, hCMV-IE promoter, mCMV-IE promoter, RSV-LTR promoter, MMTV-LTR promoter, MoMLV-LTR promoter, Ad2MLP-TPL promoter, hUBC promoter, hEF-1α promoter, mPGK promoter, β-actin promoter, TRE promoter, UAS promoter, Ac5 promoter, polyhedrin promoter, CaMKIIa promoter, GDS promoter, ADH1 promoter, Ubi promoter, or CAG promoter.

[0127] Non-limiting examples of an inducible promoter includes, GAL1 promoter, GAL10 promoter, TetO7 promoter, PhlO6 promoter, LuxO5 promoter, Met promoter, TPS1 promoter, HXT7 promoter, ADH2 promoter, SSA1 promoter, Z3V promoter, CUP1 promoter, mMT-I promoter, hMT-II promoter, hMT-IIA promoter, AlcA promoter, LexA promoter, or Ptet promoter.

[0128] The choice of promoter can determine the level of RNA transcribed from the DNA sequence and under what conditions, if any, such transcription can occur. This allows for fine tuning the level of transcribed RNA so that it is optimized for the formation of VLPs encapsulating such RNA (e.g., in a cell that co-expresses TMV coat protein).

[0129] In some embodiments, the promoter used in an expression vector to drive transcription of the encoded RNA is a GAP promoter. In some embodiments, the promoter used in an

expression vector to drive transcription of the encoded RNA is a GAP promoter comprising the DNA sequence:

TCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTC
CTAACTTTATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCC
AAAATAGGGGGCGGGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAG
TTTATTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGA
AAAAAAAAGAATCCCAGCACCAAAATATTGTTTTCTTCACCAACCATCAGTTCATA
GGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCACAAACAGGCAAAAAACG
GGCACAACCTCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCA
ATTGACCCACGCATGTATCTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTCT
CTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCTGAAATTATTCCC
CTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCT
ATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTAGTTTTAAAACA
CCAAGAACTTAGTTTCGAATAAACACACATAAACAAACAAA (SEQ ID NO:21),
or a sequence having at least 90% sequence identity to SEQ ID NO:21. In some embodiments, the GAP promoter comprises the DNA sequence of SEQ ID NO:21.

[0130] In some embodiments, the promoter used in an expression vector to drive transcription of the encoded RNA is a TEF1 promoter. In some embodiments, the promoter used in an expression vector to drive transcription of the encoded RNA is a TEF1 promoter comprising the DNA sequence:

CCACACACCATAGCTTCAAAATGTTTCTACTCCTTTTTTACTCTTCCAGATTTTCTCG
GACTCCGCGCATCGCCGTACCACTTCAAAACACCCAAGCACAGCATACTAAATTTC
CCCTCTTTCTTCCTCTAGGGTGTCGTTAATTACCCGTACTAAAGGTTTGGAAAAGAA
AAAAGACACCGCCTCGTTTCTTTTTCTTCGTCGAAAAAGGCAATAAAAATTTTTATC
ACGTTTCTTTTTCTTGAAAATTTTTTTTTTTGATTTTTTTCTCTTTCGATGACCTCCCA
TTGATATTTAAGTTAATAAACGGTCATCAATTTCTCAAGTTTCAGTTTCATTTTTCTT
GTTCTATTACAACTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAGCAATCTAAT
CTAAG (SEQ ID NO:22),
or a sequence having at least 90% sequence identity to SEQ ID NO:22. In some embodiments, the TEF1 promoter comprises the DNA sequence of SEQ ID NO:22.

[0131] In some embodiments, the promoter used in an expression vector to drive transcription of the encoded RNA is a UAS promoter.

[0132] In some embodiments, the promoter used in an expression vector to drive transcription of the encoded RNA is a derivative of the P2_8UASxpr-TEF (M Larroude et al., Microbial Biotechnology (2019) 12(6), 1249–1259) promoter and comprises the DNA sequence:

GCTTCGATCGCGTATCGATGATACGCGTCCATGGTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCC
CCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGAGGGTTGTTCCCGCCCACCTCGATCCGGC
ATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTA
CATTATCGAGAGGGTTGTTCCCGCCCACCTCGATCCGACATGTTGAGGTGTCTCACAAGTGCCGTGCAG
TCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGAGGGTTGTTCCCGCCCACCT
CGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGT
AGTGTACGTACATTATCGAGAGGGTTGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGT
GCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGAGGGTTGTTCC

CGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCT
CTTTGTGTGTAGTGTACGTACATTATCGAGAGGGTTGTTCCCGCCCACCTCGATCCGGCATGCACTGATCA
CGGGCAAAAGTGCTTCGATAGAGAGGGGGTTGGCGGCGCATTTGTGTCCCAAAAAACAGCCCCAATTG
CCCCAATTGACCCCAAATTGACCCAGTAGCGGACCCAACCCCGGCGAGAGCCCCCTTCACCCCACA
TATCAAACCTCCCCCGGTTCCCACACTTGCCGTTAAGGGCGTAGGGTACTGCAGTCTGGAATCTACGCT
TGTTCAGACTTTGTACTAGTTTCTTTGTCTGGCCATCCGGGTAACCCATGCCGGACGCAAAATAGACTACT
GAAAATTTTTTTGCTTTGTGGTTGGGACTTTAGCCAAGGGTATAAAAGACCACCGTCCCCGAATTACCTTTC
CTCTTCTTTTCTCTCTCTCCTTGTCAACTCACACCCGAAG (“UAS8-2”; SEQ ID NO:23),

or a sequence having at least 90% sequence identity to SEQ ID NO:23. In some embodiments, the derivative of the P2_8UASxpr-TEF promoter comprises the DNA sequence of SEQ ID NO:23.

[0133] In some embodiments, the promoter used in an expression vector to drive transcription of the encoded RNA is a P1_4UASxpr-TEF (M Larroude et al., *supra*) promoter comprising the DNA sequence:

ACGGCGATACGCGTATCGATACGCGTGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCC
ACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGAGGGTTGTTCCCGCCCACCTCGATCCGGCAT
GCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACA
TTATCGAGAGGGTTGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTC
CCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGAGGGTTGTTCCCGCCCACCTCG
ATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGT
GTACGTACATTATCGAGAGGGTTGTTCCCGCCCACCTCGATCCGGCATGCACTGATCACGGGCAAAAGT
GCGTTCGATAGAGAGGGGGTTGGCGGCGCATTTGTGTCCCAAAAAACAGCCCCAATTGCCCCAATTGA
CCCCAAATTGACCCAGTAGCGGACCCAACCCCGGCGAGAGCCCCCTTCACCCCACATATCAAACCTC
CCCCGGTTCCCACACTTGCCGTTAAGGGCGTAGGGTACTGCAGTCTGGAATCTACGCTTGTTCAGACTTT
GTACTAGTTTCTTTGTCTGGCCATCCGGGTAACCCATGCCGGACGCAAAATAGACTACTGAAAATTTTTTTG
CTTTGTGGTTGGGACTTTAGCCAAGGGTATAAAAGACCACCGTCCCCGAATTACCTTTCCTCTTCTTTTCTC
TCTCTCCTTGTCAACTCACACCCGAAGGATCCAATG (“UAS4-1”; SEQ ID NO:24),

or a sequence having at least 90% sequence identity to SEQ ID NO:24. In some embodiments, the derivative of the P2_8UASxpr-TEF promoter comprises the DNA sequence of SEQ ID NO:24.

[0134] In some embodiments, the promoter used in an expression vector to drive transcription of the encoded RNA is a pTEF promoter comprising the DNA sequence:

ACCGGGTTGGCGGCGTATTTGTGTCCCAAAAAACAGCCCCAATTGCCCCAATTGACCCCAAATTGACC
CAGTAGCGGGCCCAACCCCGGCGAGAGCCCCCTTCACCCCACATATCAAACCTCCCCCGGTTCCCA
CACTTGCCGTTAAGGGCGTAGGGTACTGCAGTCTGGAATCTACGCTTGTTCAGACTTTGTACTAGTTTCTTT
GTCTGGCCATCCGGGTAACCCATGCCGGACGCAAAATAGACTACTGAAAATTTTTTTGCTTTGTGGTTGG
GACTTTAGCCAAGGGTATAAAAGACCACCGTCCCCGAATTACCTTTCCTCTTCTTTTCTCTCTCTCCTTGTC
AACTCACACCCGAAATCGTTAAGCATTTCCTTCTGAGTATAAGAATCATTCAA (SEQ ID NO:36),

or a sequence having at least 90% sequence identity to SEQ ID NO:36. In some embodiments, the pTEF promoter comprises the DNA sequence of SEQ ID NO:36.

[0135] In some embodiments, the *Yarrowia sp.* cell harbors the DNA sequence capable of being transcribed into a heterologous RNA molecule on an expression vector. Such expression vectors can include other polynucleotides in addition to those encoding promoter and those that are transcribed to the heterologous RNA molecule. For example, the expression vectors may include one or more of an origin of replication, an antibiotic resistance gene, an auxotrophic marker, a gene encoding one or one or more proteins that enhance transcription and/or stability of the encoded heterologous RNA molecule, Additional 5'-UTR sequences that are added onto the 5'-end of the heterologous RNA, and/or additional 3'-UTR sequences that are added onto the 3'-end of the heterologous RNA. Antibiotic resistance genes include, but are not limited to, an ampicillin resistance gene, a tetracycline resistance gene and a kanamycin resistance gene. Auxotrophic markers include, but are not limited to, HIS3, LEU2, TRP1, and URA3. Proteins that enhance transcription or RNA stability include, but are not limited to, RNA polymerases; DNA polymerases; RNA cap transferase enzymes, such as a RNA guanine-N7-methyltrasferase, a cap 0 methyltransferase (RNMT), and cap methyltransferases 1 and 2 (CMTr1/2); endoribonuclease inhibitors; enzymes that increase the content or abundance of modified nucleotides such as methyl-transferase enzymes and pseudouridine synthase; an RNAse inhibitor; and/or a ribozyme. Ribonucleotides that may be added to the 5' end of the 5'-UTR by the expression vector include, but are not limited to, nucleotides that bind transcription factors, nucleotides that aid in translation of a heterologous coding RNA payload, nucleotides that aid in the stability of a heterologous RNA payload, and nucleotides that are the target of RNA capping enzymes. Ribonucleotides that may be added to the 3' end of the 3'-UTR sequences by the expression vector include, but are not limited to, nucleotides that encode a polyadenylation signal, nucleotides that encode a poly(A) tail, nucleotides that aid in translation of a heterologous coding RNA payload, nucleotides that aid in the stability of a heterologous RNA payload, nucleotides that aid in the transport of resulting RNA molecule and nucleotides that are terminators. Terminators are RNA sequences that usually occur at the end of a transcript and cause transcription to stop. Terminators can also impact the half-life of the RNA transcript. The choice of expression vector can depend, in part, on the *Yarrowia sp.* cell that ultimately harbors it, as well as the cell type in which it is propagated (if different from the *Yarrowia sp.* cell). Non-limiting examples of expression vectors useful herein include ESC-HIS, ESC-LEU, ESC-TRP, ESC-URA, JMP62-URA, JMP62-HIS, CfB8644, CfB8788, RK-flag-USP19, Tet07-CSII-CR5 and vectors that can be assembled from the Yarrowia lipolytica Golden Gate tool kit (Addgene Kit #1000000167).

[0136] The choice of terminator can depend in part on the promoter used to cause transcription of the heterologous RNA molecule, as well as other factors known in the art. In particular, useful terminators at the 3'-end of the heterologous RNA molecule include the CYC1 terminator:

AUCCGCUCUAACCGAAAAGGAAGGAGUUAGACAACCUGAAGUCUAGGUCCCUAU
UUAUUUUUUUAUAGUUAUGUUAGUAUUAAGAACGUUAUUUAUAUUUCAAAUUU
UUCUUUUUUUUCUGUACAGACGCGUGUACGCAUGUAACAUUAUACUGAAAACCU
UGCUUGAGAAGGUUUUGGGACGCUCGAAG (SEQ ID NO:25),

or a sequence having at least 90% sequence identity to SEQ ID NO:25. In some embodiments, the terminator comprises the DNA sequence of SEQ ID NO:25.

the T1_Tef (M Larroude et al., *supra*) terminator:

UCUAGCUGCUUGUACCUAUGCAACCCCAGUUUGUUAAAAAUUAGUAGUCAAAAA
CUUCUGAGUUAAAAAAAAAAAAAAAAGCUU ("TTEF"; SEQ ID NO:26),

or a sequence having at least 90% sequence identity to SEQ ID NO:26. In some embodiments, the terminator comprises the DNA sequence of SEQ ID NO:26, or a sequence having at least 90% sequence identity to SEQ ID NO:26. In some embodiments, the terminator comprises the DNA sequence of SEQ ID NO:26.

the T2-3_Lip2 (M Larroude et al., *supra*) terminator:

GGAUGUGUCUGUGGUAUCUAAGCUAUUUAUCACUCUUUACAACUUCUACCUCAA
CUAUCUACUUUAAUAAAUGAAUAUCGUUUAUUCUCUAUGAUUACUGUAUAUGCG
UUCCUCUAAGACAGAGU ("TLip2-2"; SEQ ID NO:27),

or a sequence having at least 90% sequence identity to SEQ ID NO:27. In some embodiments, the terminator comprises the DNA sequence of SEQ ID NO:27, or a sequence having at least 90% sequence identity to SEQ ID NO:27. In some embodiments, the terminator comprises the DNA sequence of SEQ ID NO:27.

and the T1-3_Lip2 (M Larroude et al., *supra*) terminator:

GUGUCUGUGGUAUCUAAGCUAUUUAUCACUCUUUACAACUUCUACCUCAACUAU
CUACUUUAAUAAAUGAAUAUCGUUUAUUCUCUAUGAUUACUGUAUAUGCGUUCC
UCUAAGACA (SEQ ID NO:37),

or a sequence having at least 90% sequence identity to SEQ ID NO:37. In some embodiments, the terminator comprises the DNA sequence of SEQ ID NO:37, or a sequence having at least 90% sequence identity to SEQ ID NO:37. In some embodiments, the terminator comprises the DNA sequence of SEQ ID NO:37.

[0137] TMV Capsid Protein

[0138] In order to generate TMV VLPs, the heterologous RNA molecule described herein (e.g., the RNA that comprises in a 5' to 3' order, a 5'-UTR, a heterologous RNA payload, and a 3'-UTR, wherein the 3'-UTR comprises one or more tobacco mosaic virus (TMV) packaging signals) must be co-expressed in a cell with TMV capsid protein. The TMV capsid protein may be a naturally occurring TMV capsid protein or a variant thereof that is capable of forming a VLP and encapsidating the RNA sequence. Variants of TMV capsid proteins are known in the art. In some embodiments, the TMV capsid protein co-expressed in a cell has the amino acid sequence:

MSYSITTPSQFVFLSSAWADPIELINLCTNALGNQFQTQQARTVVQRQFSEVWKPSPQVT
VRFPDSDFKVYRYNAVLDPLVTALLGAFDTRNRIIEVENQANPTTAETLDATRRVDDAT
VAIRSAINNLIVELIRGTGSYNRSSFESSSGLVWTSGPAT (SEQ ID NO:28).

In some embodiments, the TMV capsid protein is encoded by the DNA sequence:

ATGTCTTACAGTATCACTACTCCATCTCAGTTCGTGTTCTTGTCATCAGCGTGGGCCG
ACCCAATAGAGTTAATTAATTTATGTACTAATGCCTTAGGAAATCAGTTTCAAACAC
AACAAGCTCGAACTGTCGTTCAAAGACAATTCAGTGAGGTGTGGAAACCTTCACCA
CAAGTAACTGTTAGGTTCCCTGACAGTGACTTTAAGGTGTACAGGTACAATGCGGT
ATTAGACCCGCTAGTCACAGCACTGTTAGGTGCATTCGACACTAGAAATAGAATAA
TAGAAGTTGAAAATCAGGCGAACCCCACGACTGCCGAAACGTTAGATGCTACTCGT
AGAGTAGACGACGCAACGGTGGCCATAAGGAGCGCGATAAATAATTTAATAGTAG
AATTGATCAGAGGAACCGGATCTTATAATCGGAGCTCTTTCGAGAGCTCTTCTGGTT
TGGTTTGGACCTCTGGTCCTGCAACTTGA (SEQ ID NO:29),

or a DNA sequence having at least 50% sequence identity to and encoding the same amino acid sequence as SEQ ID NO:29.

[0139] In some embodiments the TMV capsid protein amino acid sequence may comprise up to 10 amino acid insertions, deletions or substitutions as long as the resulting polypeptide is capable of forming intact TMV VLPs that encapsidate the heterologous RNA molecule. In some embodiments, the amino acid substitutions are conservative amino acid substitutions, which are well known in the art.

[0140] The TMV capsid protein may be expressed in a *Yarrowia sp.* cell through the use of an expression vector encoding that protein. In some embodiments, the expression vector encoding

the TMV capsid protein is separate from the expression vector encoding the heterologous RNA molecule. In some embodiments, the expression vector encoding the TMV capsid protein is the same as the expression vector encoding the heterologous RNA molecule (e.g., both the heterologous RNA molecule and TMV capsid protein are encoded on a single expression vector).

[0141] Expression of the TMV capsid protein from an expression vector requires that the DNA sequence encoding that protein be operatively linked to a promoter. Such promoter can be the same promoter that drives transcription of the heterologous RNA molecule (e.g., a second copy of that promoter or a single copy of a promoter that is capable of transcribing both the TMV capsid protein encoding DNA and the heterologous RNA molecule, e.g., a bidirectional promoter), or a different promoter. The promoter driving TMV capsid protein expression can be a strong, weak, constitutive, inducible, cell type specific promoter, or a tissue specific promoter. The promoter can be a constitutive promoter or an inducible promoter. The selection of the promoter can be optimized for the *Yarrowia sp.* cell in which the capsid protein is expressed. Specific choices of promoters may be selected from the list of constitutive and inducible promoters set forth above. In some embodiments, the promoter used to drive expression of TMV capsid protein is a GAP promoter. In some embodiments, the promoter used in an expression vector to drive expression of the TMV capsid protein is a GAP promoter comprising the DNA sequence of SEQ ID NO:21. In some embodiments, the promoter used in an expression vector to drive expression of the TMV capsid is a UAS4-1 promoter comprising the DNA sequence of SEQ ID NO:24.

[0142] In some embodiments, the RNA transcript encoding the TMV capsid protein comprises a terminator at its 3'-end. In some embodiments, the terminator is a TTEF terminator or SEQ ID NO:26.

[0143] Host Cells

[0144] As set forth above, the disclosure provides a *Yarrowia sp.* cell harboring a DNA sequence capable of being transcribed into the heterologous RNA molecule disclosed herein and a DMA sequence encoding the TMV capsid protein. As detailed above, in some embodiments, the *Yarrowia sp.* cell will harbor a single expression vector that encodes both the heterologous RNA molecule and the TMV capsid protein. In other embodiments, the host cell will harbor two separate expression vectors – one that encodes the heterologous RNA molecule and another that encodes the TMV capsid protein.

[0145] In some embodiments, the *Yarrowia sp.* cell can express, either naturally or recombinantly, enzymes and proteins capable of promoting transcription of, modifying and/or stabilizing the RNA sequence. Such enzymes and proteins include, but not limited to, RNA polymerases, DNA polymerases, RNA cap transferase enzymes, RNA ligase, RNase inhibitors, and/or enzymes that increase the content of modified nucleotides. Some non-limiting examples of specific enzymes or proteins that may be expressed include 2'-O methyl-transferase enzyme or enzyme complex, which facilitates the addition of a methyl group at the 2'-O position of the first nucleotide adjacent to the cap structure at the 5′ end of the RNA; pseudouridine synthase that facilitates the site-specific isomerization of uridines into pseudouridine or N1-methyl-pseudouridines; a RNA guanine-N7-methyltrasferase; a cap 0 methyltransferase (RNMT); a cap methyltransferase 1 or 2 (CMTr1 or 2); a ribozyme; Tobacco mosaic virus RNA-dependent polymerase; a Brome mosaic virus RNA polymerase; a Turnip crinkle virus RNA polymerase; a Potato virus X RNA polymerase; a cowpea mild mottle virus RNA polymerase; an MS2 bacteriophage RNA polymerase; a T7 RNA polymerase; a T3 RNA polymerase; and an SP6 RNA polymerase. The *Yarrowia sp.* cell may also express, either naturally or recombinantly, an auxotrophic marker. Non-limiting examples of auxotrophic markers include HIS3, LEU2, TRP, MET15, and/or URA3. As set forth above, any of these enzymes or proteins may also be encoded on an expression vector that also encodes the heterologous RNA molecule, an expression vector that also encodes the TMV capsid protein, or on an expression vector that also encodes both the heterologous RNA molecule and the TMV capsid protein.

[0146] In some embodiments, the *Yarrowia sp.* cell is a naturally occurring cell. In other embodiments, the *Yarrowia sp.* cell has been genetically modified (e.g., to express a heterologous enzyme or a auxotrophic marker). In some embodiments, the *Yarrowia sp.* cell is selected from a *Y. lipolytica* cell, *Y. divulgata* cell, and a *Y. keelungensis* cell. In a more specific embodiment, the *Yarrowia sp.* cell is a *Y. lipolytica* cell. In an even more specific embodiment, the *Yarrowia sp.* cell is a *Y. lipolytica* strain PO1h or PO1d cell.

[0147] Methods of Producing and Isolating VLPs Containing the Heterologous RNA Molecule

[0148] In some aspects, the disclosure provides a method of producing a heterologous RNA molecule in a *Yarrowia sp.* cell, wherein the cell comprises:

a. a first promoter operatively linked to a DNA sequence capable of being transcribed into a heterologous mRNA molecule that comprises, in a 5' to 3' order, a 5'-UTR, a heterologous RNA payload, and a 3'-UTR, wherein the 3'-UTR comprises one tobacco mosaic virus (TMV) packaging signal; and

b. a second promoter operatively linked to a DNA sequence encoding a TMV viral capsid protein; and wherein the method comprises growing that cell under conditions in which the heterologous mRNA molecule is bound to and encapsidated by the TMV viral capsid protein in a virus-like particle, thereby producing the heterologous mRNA molecule.

[0149] The expressed heterologous RNA molecule will bind to the TMV capsid protein through the one or more TMV packaging signals present in the RNA molecule, resulting in the self-assembly and production of virus-like particles encapsidating the RNA.

[0150] In some aspects, the disclosure provides VLPs comprising a heterologous RNA molecule disclosed herein isolated from the host cell.

[0151] The VLPs can be isolated from the *Yarrowia sp.* cell by cell lysis followed by centrifugation and treatment of the soluble material with PEG to precipitate the VLPs. The precipitated VLPs may be further purified by density gradient centrifugation. Even further purification of the VLPs may be achieved through the use of ultrafiltration, membrane filter centrifugation and/or size exclusion chromatography.

[0152] In some embodiments, the isolated VLPs are treated with one or more enzymes that digest nucleic acids or polynucleotides. Nucleic acids or polynucleotides produced by the host cell, whether naturally occurring or the result of transcription from one or more expression vectors harbored by the host cell, is not desirable and therefore it is advantageous to remove it from the isolated VLPs. Such nucleic acids or polynucleotides may stick to or are associated with the exterior of the VLP. Thus, treatment of the VLP with an enzyme that digests these species provides a method of removing such contaminants while the VLP protects the encapsidated RNA comprising the heterologous RNA payload from attack by such enzymes. In some embodiments, the isolated VLPs are treated with one or more RNases or nucleases that digest double-stranded RNA. In some embodiments the isolated VLPs are treated with benzonase. In some embodiments, the isolated VLPs are treated with a cocktail of RNases.

[0153] Method of Isolating Heterologous RNA Molecule from VLPs.

[0154] In some aspects, the disclosure provides a heterologous RNA molecule isolated from the VLPs encapsidating such RNA.

[0155] Once the VLPs have been isolated and optionally treated with an enzyme that digests nucleic acids or polynucleotides external to the TMV VLP, the encapsidated RNA therein can be extracted. Such extraction may be achieved using commercially available RNA-extraction columns or other RNA extraction methodologies known in the art. The extracted RNA may then be used directly or further formulated to administer to a subject in need thereof.

[0156] In some aspects, the encapsidated RNA is not extracted from the VLPs and the VLPs themselves are used directly or are further formulated as an agent to deliver the RNA therein to a subject in need thereof.

[0157] All patents, patent applications, and other scientific or technical writings referred to anywhere herein are incorporated by reference herein in their entirety. The embodiments illustratively described herein suitably can be practiced in the absence of any element or elements, limitation or limitations that are specifically or not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising," "consisting essentially of," and "consisting of" can be replaced with either of the other two terms, while retaining their ordinary meanings. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof.

[0158] It should be understood that although the present methods and compositions have been specifically disclosed by embodiments and optional features, modifications and variations of the concepts herein disclosed can be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of the compositions and methods as defined by the description and the appended claims.

[0159] Any single term, single element, single phrase, group of terms, group of phrases, or group of elements described herein can each be specifically excluded from the claims.

[0160] Whenever a range is given in the specification, for example, a temperature range, a time range, a composition, or concentration range, all intermediate ranges and subranges, as well as all individual values included in the ranges given are intended to be included in the disclosure. It will be understood that any subranges or individual values in a range or subrange that are included in the description herein can be excluded from the aspects herein. It will be understood that any elements or steps that are included in the description herein can be excluded from the claimed compositions or methods.

[0161] In addition, where features or aspects of the compositions and methods are described in terms of Markush groups or other grouping of alternatives, those skilled in the art will recognize that the compositions and methods are also thereby described in terms of any individual member or subgroup of members of the Markush group or other group. Similarly, any two or more embodiments of a particular aspect set forth herein may be combined even if such combination is not specifically disclosed.

EXAMPLES

[0162] Example 1. Production of VLPs in *S. cerevisiae* and *Y. lipolytica*.

Table 3, below, shows the content of various plasmids used to produce a heterologous RNA molecule encapsidated by TMV VLPs. Maps of these plasmids are shown in FIGs. 1-12 and 23-32.

[0163] Table 3. Plasmids used in the production of VLPs.

Plasmid Name	**Promoter/ Terminator**	**5'-UTR**	**heterologous RNA payload***	**3'-UTR**
S. cerevisiae				
SC-GFP UTR-1 (FIG. 1)	TEF1 (SEQ ID NO:22)/ CYC1 (SEQ ID NO:25)	SEQ ID NO:1	GFP (SEQ ID NO:30)	UTR-1 (SEQ ID NO:13)
SC-Luc OAS (FIG. 2)			Firefly luciferase (SEQ ID NO:31)	UTR-OAS (SEQ ID NO:15)
SC-Luc UTR-4 (FIG. 3)			Firefly luciferase (SEQ ID NO:31)	UTR-4 (SEQ ID NO:4)
SC-EPO UTR-4 (FIG. 4)			Erythropoietin (SEQ ID NO:32)	UTR-4 (SEQ ID NO:4)
SC-TMV-CP (FIG. 5)	GAP (SEQ ID NO:21)/ CYC1 (SEQ ID NO:25)	plasmid encoded the TMV capsid protein (SEQ ID NO:29)		
Plasmid 78 (FIG. 32)	TEF1 (SEQ ID NO:22) / CYC1 (SEQ ID NO:25)	SEQ ID NO:1	Firefly luciferase (SEQ ID NO:31)	UTR_B_AES (SEQ ID NO:43)
Plasmid 81 (figure not shown)				UTR_B_mtRNR1 (SEQ ID NO:42)
Plasmid 84 (figure not shown)				UTR_B_MRS2 (SEQ ID NO:44)
Plasmid 341 (figure not shown)			GFP (SEQ ID NO:30)	UTR_B_mtRNR1 (SEQ ID NO:42)
Plasmid 342 (figure not shown)				UTR_B_AES (SEQ ID NO:43)
Plasmid 343 (figure not shown)				UTR_B_MRS2 (SEQ ID NO:44)
Plasmid 344 (figure not shown)			Erythropoietin (SEQ ID NO:32)	UTR_B_mtRNR1 (SEQ ID NO:42)
Plasmid 345 (figure not shown)				UTR_B_AES (SEQ ID NO:43)
Plasmid 346 (figure not shown)				UTR_B_MRS2 (SEQ ID NO:44)

Plasmid Name	Promoter/ Terminator	5'-UTR	heterologous RNA payload*	3'-UTR
Y. lipolytica				
YL-TMV-GFP UTR-1** (FIG. 6)	Modified UAS8-2 (SEQ ID NO:23)/ TLip2-2 (SEQ ID NO:27)	SEQ ID NO:1	GFP (SEQ ID NO:30)	UTR-1 (SEQ ID NO:13)
YL-TMV-Luc OAS** (FIG. 7)			Firefly luciferase (SEQ ID NO:31)	UTR-OAS (SEQ ID NO:15)
YL-TMV-Luc UTR-4** (FIG. 8)			Firefly luciferase (SEQ ID NO:31)	UTR-4 (SEQ ID NO:4)
YL-TMV-EPO UTR-4** (FIG. 9)			Erythropoietin (SEQ ID NO:32)	UTR-4 (SEQ ID NO:4)
YL-Luc OAS (FIG. 10)	pTEF (SEQ ID NO:36)/T1-3 Lipt2 (SEQ ID NO:37)	SEQ ID NO:1	Firefly luciferase (SEQ ID NO:31)	UTR-OAS (SEQ ID NO:15)
YL-Luc UTR-4 (FIG. 11)				UTR-4 (SEQ ID NO:4)
YL-TMV-CP (FIG. 12)		plasmid encoded the TMV capsid protein (SEQ ID NO:29)		
Plasmid 230 (FIG. 23)**	Modified UAS8-2 (SEQ ID NO:23)/ TLip2-2 (SEQ ID NO:27)	SEQ ID NO:1	Firefly luciferase (SEQ ID NO:31)	UTR_B_mtRNR1 (SEQ ID NO:42)
Plasmid 333 (FIG. 24)**				UTR_B_AES (SEQ ID NO:43)
Plasmid 334 (FIG. 25)**				UTR_B_MRS2 (SEQ ID NO:44)
Plasmid 335 (FIG. 26)**			GFP (SEQ ID NO:30)	UTR_B_mtRNR1 (SEQ ID NO:42)
Plasmid 336 (FIG. 27)**				UTR_B_AES (SEQ ID NO:43)
Plasmid 337 (FIG. 28)**				UTR_B_MRS2 (SEQ ID NO:44)
Plasmid 338 (FIG. 29)**			Erythropoietin (SEQ ID NO:32)	UTR_B_mtRNR1 (SEQ ID NO:42)
Plasmid 339 (FIG. 30)**				UTR_B_AES (SEQ ID NO:43)
Plasmid 340 (FIG. 31)**				UTR_B_MRS2 (SEQ ID NO:44)

*SEQ ID NOs for heterologous payloads are for the DNA sequence encoding the payload.
**Plasmid also contains the TMV capsid protein coding sequence operatively linked to the UAS4-1 promoter of SEQ ID NO:24 and a TTEF terminator of SEQ ID NO:26.

[0164] *S. cerevisiae* yeast production cells (BY 4741) were cultured in YPD (1% [w/v] yeast extract, 2% [w/v] peptone, 2% [w/v] dextrose) media before dual transformation with both a TMV-CP plasmid and a plasmid selected from one of SC-GFP UTR-1, SC-Luc OAS, SC-Luc UTR-4, or SC-EPO UTR-4 using the Yeast Transformation Kit (Sigma-Aldrich Cat. YEAST1-1KT) with salmon sperm ssDNA (Abcam Cat. ab229278) and following the manufacturer's instructions. Transformations with Plasmid 78, Plasmid 81, Plasmid 84, Plasmid 341, Plasmid

342, Plasmid 343, Plasmid 344, Plasmid 345, or Plasmid 346 were done in BY 4741 cells that had previously been transformed with the TMV-CP plasmid (BY 4741_TMV(GAP) cells). Prior to transformation with any of Plasmid 78, Plasmid 81, Plasmid 84, Plasmid 341, Plasmid 342, Plasmid 343, Plasmid 344, Plasmid 345, or Plasmid 346, the BY 4741_TMV(GAP) cells were cultured in YNB-HIS medium (which selected for the presence of the TMV-CP plasmid) for 24 hours at 30° C. All transformation reactions were incubated at 30°C for 30 mins followed by heat-shock treatment before cells were plated onto solid auxotrophic drop-out media (0.17% [w/v] yeast nitrogen base without supplemented amino acids and nitrogen bases, 0.5% [w/v] ammonium sulfate, 0.14% [w/v] of Yeast Synthetic Drop-out Medium Supplements without histidine, leucine, tryptophan and uracil, 2% [w/v] dextrose, 50 mg/L tryptophan, 50 mg/L uracil). After 4 days at 30°C, positive selection of plasmid-containing cells was performed. Glycerol stocks of the transformed yeasts were cultured by inoculating cells into yeast nitrogen base with the same yeast auxotrophic drop-out medium. Yeasts were cultured for a total of 5 days at 30°C with constant shaking. Cells were harvested by centrifugation and supernatants removed, cells were then resuspended, and the centrifugation repeated. Cells from transformations with SC-GFP UTR-1, SC-Luc OAS, SC-Luc UTR-4, or SC-EPO UTR-4 next resuspended in TEN buffer (50 mM Tris pH 7,4; 500 mM NaCl; 2mM EDTA; 1 cOmplete™ Mini protease inhibitor cocktail tablet/50 mL (Roche Cat.04693124001). Those cells were then lysed using mechanical lysis by glass bead beating (TissueLyserII, 425-600 µm, Sigma-Aldrich, Cat. G8772). Cell lysates containing the VLPs were then collected and centrifuged to remove beads. Cells from transformations with Plasmid 78, Plasmid 81, Plasmid 341, Plasmid 342, Plasmid 344, or Plasmid 345, or were resuspended in PBS containing 1 cOmplete™ Mini protease inhibitor cocktail tablet/50 mL and then lysed by shearing lysis using a Microfluidizer® (LM20, Microfluidics). Cells from transformations with Plasmid 84, Plasmid 343, and Plasmid 346 were grown and stored a glycerol stocks for later analysis.

[0165] *Y. lipolytica* yeast production cells (PO1h or PO1d) were cultured in YPD (1% [w/v] yeast extract, 2% [w/v] peptone, 2% [w/v] D(+) glucose) media until efficient biomass was achieved. Cells were then prepared for transformation by transferral into 0.1M lithium acetate buffer pH 6. Plasmids were linearized by restriction digestion using rCutSmart buffer (NEB, Cat. B6004S) with NotI-HF (NEB, Cat. R3189L) prior to transformation. Reactions were incubated at 37 °C for 2 hours, followed by 20 min at 65°C, then stored on ice until transformation. Transformation was achieved by the addition of one of plasmids YL-TMV-GFP UTR-1, YL-TMV-Luc OAS, YL-TMV-Luc UTR-4, YL-TMV-EPO UTR-4, Plasmid 230,

Plasmid 333, Plasmid 334, Plasmid 335, Plasmid 336, Plasmid 337, Plasmid338, Plasmid 339, or Plasmid 340, or by co-transformation with YL-TMV-CP and one of YL-Luc UTR-4 or YL-Luc OAS, and the transformation mix. The transformation mix was: salmon sperm ssDNA (0.3 mg/mL Abcam Cat. ab229278), 40% PEG3350, 0.05M DTT and 0.1M lithium acetate pH 6, followed by incubation at 39°C for 60 min. The transformed cells were then isolated and plated onto solid auxotrophic drop-out media (0.17% [w/v] yeast nitrogen base without supplemented amino acids and nitrogen bases, 0.5% [w/v] ammonium sulfate, 2% [w/v] D(+)-glucose, 0.19% [w/v] Yeast Synthetic Drop-out Medium Supplements without uracil) plates and colonies were selected after 4 days. Glycerol stocks of *Y. lipolytica* yeasts transformed with YL-TMV-GFP UTR-1, YL-TMV-Luc OAS, YL-TMV-Luc UTR-4, YL-TMV-EPO UTR-4, Plasmid 230, Plasmid 333, Plasmid 335, Plasmid 336, Plasmid338, or Plasmid 339 were cultured by inoculating cells into yeast nitrogen base with yeast synthetic drop-out medium supplemented with glucose. Yeasts were cultured for a total of 3 days at 30°C with constant shaking. Cells were harvested by centrifugation and supernatants were removed. The yeast cells were then resuspended, and the centrifugation was repeated. Cells were next resuspended in PBS buffer pH 7.2 with 1 cOmplete™ Mini protease inhibitor cocktail tablet/50 mL (Roche Cat. 04693124001). The cell suspension was then treated with lyticase (Sigma Cat: L2524) at 750 U/mL and was incubated at 37°C for 1 hour before shearing lysis using a Microfluidizer® (LM20, Microfluidics).

[0166] Example 2. Purification of VLPs.

[0167] VLPs were purified from lysed yeast cells using either density centrifugation or chromatography. For experiments using any of Plasmid 78, Plasmid 81, Plasmid 341, Plasmid 342, Plasmid 344, or Plasmid 345 in *S. cerevisiae*, or Plasmid 230, Plasmid 333, Plasmid 335, Plasmid 336, Plasmid 338, or Plasmid 339 in *Y. lipolytica*, chromatography was used.

[0168] *Density Centrifugation*

[0169] Clarified soluble material containing VLPs from yeast cell lysates was transferred to centrifuge tubes and PEG master solution (20% (w/v) PEG 6000, 1M NaCl) was added to each sample resulting in a final concentration of 4% (w/v) PEG 6000, 200mM NaCl. The samples were gently mixed before incubation overnight at 4°C using gentle rotation. Samples were then centrifuged at 13,000 x g for 45 minutes at 20°C to pellet the visible precipitated material. Samples were examined by denaturing SDS-PAGE to confirm the presence of TMV capsid protein both in the lysate and the PEG pellet. Density-based VLP purification was next performed by resuspending PEG-precipitated pellets into PBS before adding CsCl solution (35%

w/v) and centrifugation at 180,000 x g. VLPs were visible in the CsCl ultracentrifugation tube as a blue band. The solution was fractionated by serial removal of 500 µL aliquots and the presence of TMV capsid was analyzed by denaturing SDS-PAGE of each aliquot. Capsid-containing fractions were pooled and transferred into PBS (150 mM) pH 7.2 using membrane filter centrifugation. Total capsid protein concentrations were measured by BCA assay (Pierce BCA kit, Thermo Scientific Cat. 23225).

[0170] *Chromatography*

[0171] TMV-RNA particles were also purified from cell lysate by size-exclusion chromatography using a CIMultus™ OH (2 µm) column with 1.5 M potassium phosphate buffer pH 7 as the equilibration buffer and 50 mM potassium phosphate buffer pH 7 as the elution buffer. Fractions containing TMV VLPs were collected and pooled before being confirmed for purity by SDS-PAGE analysis. Pooled fractions were transferred into PBS (150 mM) pH 7.2 using membrane filter centrifugation before concentration and total capsid protein concentrations were measured by BCA assay (Pierce BCA kit, Thermo Scientific Cat. 23225).

[0172] Example 3. Extraction of Heterologous RNA Molecules from VLPs

[0173] VLPs in PBS were treated with a nuclease enzyme or a mixture of enzymes (Benzonase or specific cocktails of RNases) designed to digest nucleic acid in the presence of $MgCl_2$ for 1 hour at 37°C. Digestion reactions were terminated by the addition of EDTA. VLPs were then exchanged into PBS by membrane centrifugal filtration before the encapsidated RNA was extracted using commercial RNA-extraction columns based on chaotropic salt silica membrane separation principles (RNeasy Mini kit, Qiagen, cat. 74104, Monarch RNA cleanup, NEB, cat. T2023). The integrity of isolated RNA was assessed by capillary electrophoresis using either Agilent 2200 TapeStation or 5300 Fragment Analyzer System. RNA concentrations were determined using absorption at 260 nm by NanoDrop. The yield per liter of culture of both TMV protein (FIG. 13A) and RNA (FIG. 13B) was measured. For both TMV protein and RNA isolated from the TMV VLPs, yields were significantly higher when VLPs were produced in *Y. lipolytica* compared to *S. cerevisiae* both for GFP-encoding RNA containing a single TMV OAS (compare SC-GFP UTR-1 to YL-TMV-GFP UTR-1 in FIG. 13, FIG. 13A and FIG. 13B) and for luciferase-encoding RNA containing two TMV OAS (compare SC-Luc UTR-4 to YL-TMV-Luc UTR-4 in FIG. 13, FIG. 13A and FIG. 13B). The integrity of purified RNA produced from cells transformed with SC-GFP UTR-1 (FIG. 14A) and YL-TMV-GFP UTR-1 (FIG. 14B) was confirmed by capillary electrophoresis.

[0174] For experiments utilizing any of Plasmid 78, Plasmid 81, Plasmid 341, Plasmid 342, Plasmid 344, or Plasmid 345 in *S. cerevisiae* and the comparative Plasmid 230, Plasmid 333, Plasmid 335, Plasmid 336, Plasmid 338, or Plasmid 339 in *Y. lipolytica*, total TMV capsid protein (a measure of VLP yield), total RNA isolated from the TMV VLPs, and total target RNA (the heterologous RNA encoded by plasmid, i.e., GFP, luciferase, or erythropoietin) were measured. The results are also shown in Table 4, below. Table 4 shows that except in the case of Plasmid 78/ Plasmid 333 (firefly luciferase with a 3'-UTR derived from AES), a higher yield of total TMV capsid protein was recovered from the transformed *Y. lipolytica* cells as compared to the corresponding *S. cerevisiae* cells. Table 4 also shows that in all cases, the amount of total mRNA isolated from the *Y. lipolytica* TMV VLPs was vastly higher than that isolated from corresponding *S. cerevisiae* TMV VLPs and that the yield of total mRNA from the *S. cerevisiae* TMV VLPs was consistently very low (less than 3 µg/L). The same was true for target mRNA yield, with the yield of target mRNA from *S. cerevisiae* TMV VLPs never exceeding about 0.4 µg/L, while the yield of target mRNA from *Y. lipolytica* TMV VLPs ranged from about 2.7 µg/L to about 126 µg/L. For each pair of corresponding plasmids, the target mRNA yield from *Y. lipolytica* TMV VLPs was between about 36 and 490-fold greater than that from the corresponding *S. cerevisiae* TMV VLPs. In all cases except for luciferase as the heterologous mRNA, the target mRNA integrity (target mRNA yield/total mRNA yield) was at least 4-fold higher for mRNA isolated from *Y. lipolytica* TMV VLPs than the corresponding *S. cerevisiae* TMV VLPs.

[0175] Table 4. TMV CP, Total RNA and Target RNA Yields from Comparative *S. cerevisiae* and *Y. lipolytica* TMV VLPs.

Plasmid #	Host Cell	Target mRNA	3'-UTR	CP (mg/L)	Total mRNA (ug/L)	Target mRNA (ug/L)	Target mRNA integrity (%)	Δ target mRNA (YL over SC; fold)
341	SC	GFP	mtRNR1	1.379	2.956	0.408	13.8	
335	YL	GFP	mtRNR1	11.379	31.862	19.448	61.0	48
342	SC	GFP	AES	1.221	0.824	0.100	12.1	
336	YL	GFP	AES	12.742	131.250	67.750	51.6	678
344	SC	hEPO	mtRNR1	3.950	2.033	0.388	19.1	
338	YL	hEPO	mtRNR1	8.233	129.958	114.708	88.3	296
345	SC	hEPO	AES	3.492	1.374	0.258	18.8	
339	YL	hEPO	AES	7.488	155.417	126.500	81.4	490
81	SC	Luc	mtRNR1	8.738	0.423	0.075	17.7	
230	YL	Luc	mtRNR1	11.207	10.776	2.724	25.3	36
78	SC	Luc	AES	9.333	0.375	0.088	23.3	
333	YL	Luc	AES	7.724	118.155	28.586	24.2	327

[0176] Example 4. In Vitro Transcription Production of RNA

[0177] Double-stranded DNA templates that produce a luciferase-encoding RNA sequence during IVT, were generated from plasmids containing, in a 5'-to-3' order, a DNA sequence capable of being transcribed into the 5'-UTR of SEQ ID NO:1, a luciferase coding sequence (SEQ ID NO:31), and a DNA sequence capable of being transcribed into UTR-4 SEQ ID NO:4 or UTR-OAS (SEQ ID NO:15). Primers were designed to target the sequence only between the beginning of the 5'-UTR to the end of the 3'-UTR of the luciferase RNA encoding template within the plasmid backbone. The primers also added 5' to the 5'-UTR a sequence that corresponded to the T7 promoter (SEQ ID NO:38) and an additional 3' adenosine (A) so that, together with 5'-most G of the 5'-UTR, that AG created a template for co-transcriptional capping with the CleanCap® reagent. The CleanCap® Reagent AG was used during the IVT reaction to generate cap-1 capped mRNA. At the 3' end of the 3'-UTR, the primers added an 80 nucleotide-long polyadenosine tail. The positive-sense strand of the resulting DNA templates had the nucleotide sequences set forth in SEQ ID NO:33 and SEQ ID NO:34. PCR reactions were then used to amplify the template using High-Fidelity Master Mix (NEB, Cat: M0541S), following the manufacturer's protocol. These DNA templates were purified (GeneJET® PCR Purification Kit, Thermo Fisher Scientific) and transcribed into RNA using the HiScribe® T7 mRNA Kit (NEB, Cat E2080S) in the presence of CleanCap® reagent AG at 4mM. All other reagents were used in accordance with the HiScribe kit instructions. The reaction was incubated at 37°C for 150 min before termination with DNase I and a further 30 min incubation. The RNA was then extracted from the reaction mixture using Oligo d(T)25 beads (NEB, Cat: S1550S) following the manufacturer's instructions. The eluted RNA was then further purified using Monarch RNA Cleanup Kit (NEB, Cat: T2040L) before analysis by capillary electrophoresis. The successful production and the correct length of the RNA were confirmed by capillary electrophoresis before transfection.

[0178] For certain experiments (as indicated) we used commercially available IVT-produced RNA that encoded the same expression product. That commercial IVT-produced RNA did not comprise the same 5'- or 3'-UTRs as the RNA produced from the plasmids that we used for VLP production or our in-house made IVT-generated RNA, nor did it comprise any TMV packaging sequences.

[0179] Example 5. Transfection of Cells with RNA

[0180] THP-1 Dual™ cells NF-κB-SEAP IRF-Luc Reporter Monocytes (InvivoGen Cat. thpd-nfis) were modified to express Lucia luciferase enzyme when the interferon regulatory factor

pathway is stimulated as a result of TLR-3, MDA5, RIG-1 or cCAS receptor activation. THP-1 Dual™ cells were cultured using standard mammalian cell culture techniques in RPMI 1640 + 10% FBS (heat inactivated) with 1% Pen/Strep and the selective antibiotic Normocin. At cell passage P4, the THP-1 cells, which are monocyte-like, were differentiated into macrophage-like cells using phorbol 12-myristate 13-acetate (PMA) (Thermo Fisher Cat. J63916) as follows. THP-1 cells were seeded into flat-bottomed well plates before treatment with PMA at a final concentration of 50 ng/mL in growth medium for 3 hours at 37°C. The treatment medium was removed and replaced with growth medium before the cells were incubated for 3 days. The resulting adherent cells were gently washed into fresh media before being transfected with RNA samples.

[0181] Adherent HEK cells were cultured in DMEM + 10% FBS with 1% Pen/Strep and were passaged every 2-4 days, using trypsin-EDTA to detach the cells from the plate. Cells were then seeded in 96-well plates using OptiMEM serum-free media and were left to adhere for at least 2 hours before transfection with RNA samples.

[0182] HepG2 were treated similarly however the growth media was MEM rather than DMEM.

[0183] RNA samples (RNA isolated from VLPs; RNA produced by IVT; and controls) were prepared for transfection with Lipofectamine (Invitrogen Cat.LMRNA008) reagent according to the manufacturer's instructions and immediately added to cells in an amount that was calculated to contain 125 ng of reporter gene-specific (i.e., RNA that encoded GFP, firefly luciferase or erythropoietin (EPO)) RNA per well. Poly (dA:dT) (InvivoGen Cat. tlrl-patn) was used as an immunogenicity-positive control transfected with lipofectamine at 125 ng per well.

[0184] Example 6. Expression of Protein Encoded by RNA Payload in Transformed Cells

[0185] After 24 hours of incubation at 37°C after transfection, the cells were assessed for transfection efficiency by the expression of GFP, firefly luciferase or erythropoietin (EPO) reporter genes. GFP expression was read directly from the plate, after washing the cells into PBS, using a plate reader at 485 nm excitation and 520 emission wavelengths. Luciferase expression was analyzed by measuring bioluminescence resulting from the addition of the substrate luciferin to transfected cells expressing luciferase. The protocol from the kit Promega Luciferase Assay System (Cat: E1501) was followed. In short, cells were lysed in the 96 well plate with the provided buffer and 50 µL of the lysate was transferred to a white, opaque plate. The substrate D-luciferin was added to the lysate and bioluminescence was measured immediately using a plate reader. EPO protein expression was measured at 24 hours post-transfection by taking 50 µL of the supernatants from wells containing the transfected THP-1

derived macrophages and diluting 1:100 with ELISA kit diluent reagent. The human EPO quantification ELISA was performed following the manufacturer's instructions (R&D Systems Cat: DY286-05).

[0186] The functionality of the RNA produced in *S. cerevisiae* and *Y. lipolytica* in TMV VLPs, was analyzed *in vitro* and compared sequence-matched IVT-produced RNA, as well as commercially purchased IVT-produced RNA encoding the corresponding reporter protein in the above-described THP1-Dual™ cell assays. THP1-Dual™ derived macrophages are sensitive to proinflammatory contaminants during transfection and will reduce translation of reporter proteins if they are immune stimulated.

[0187] As shown in FIG. 15, the amount of GFP produced from VLPs isolated from *Y. lipolytica* transformed with YL-TMV-GFP UTR-1 was significantly higher than that produced from VLPs isolated from *S. cerevisiae* co-transformed with SC-GFP UTR-1 and TMV-CP. Also, the amount of GFP produced from those *Y. lipolytica* VLPs was significantly greater than GFP produced from both the corresponding IVT-generated RNA and a commercially obtained IVT-generated RNA that encoded the same GFP (Trilink CleanCap® EGFP Cat: L-7601).

[0188] FIG. 16A, shows that firefly luciferase expression in THP-1 cells transformed from RNA produced from VLPs isolated from *Y. lipolytica* transformed with YL-TMV-Luc UTR-4 or co-transformed with YL-Luc UTR-4 and YL-TMV-CP had higher than expression from RNA produced from VLPs isolated from *S. cerevisiae* co-transformed with SC-Luc UTR-4 and TMV-CP. It was also higher than luciferase expressed from RNA generated by IVT from the corresponding template of SEQ ID NO:33 and similar to a commercially available IVT-generated RNA encoding that same luciferase (Trilink CleanCap® FLuc mRNA Cat: L-7602). FIG. 16B, shows firefly luciferase expression in these THP-1 cells transformed from RNA produced from VLPs isolated from *Y. lipolytica* co-transformed with YL-TMV-CP and either YL-Luc OAS or YL-Luc UTR-4 compared to corresponding IVT-generated RNA from the template of SEQ ID NO:34 or SEQ ID NO:33, respectively. VLP-derived RNA from *Y. lipolytica* transformed with YL-Luc UTR-4 produced significantly more firefly luciferase than the corresponding IVT-derived RNA (from the template of SEQ ID NO:33). VLP-derived RNA from *Y. lipolytica* transformed with YL-Luc UTR-OAS produced slightly less firefly luciferase than the corresponding IVT-derived RNA (from the template of SEQ ID NO:34), but was substantially less immunogenic (see below).

[0189] FIG. 17 shows the expression of erythropoietin in THP-1 cells transformed from RNA produced from VLPs isolated from *Y. lipolytica* transformed with YL-TMV-EPO UTR-4 is

significantly greater than from RNA produced from VLPs isolated from *S. cerevisiae* co-transformed with SC-EPO UTR-4 and TMV-CP.

[0190] FIG. 33 shows that THP-1 cells transformed with VLP-derived RNA from *Y. lipolytica* transformed with Plasmid 335 ("YL-eGFP-mtRNR1") produced a statistically significant greater amount of GFP than THP-1 cells transformed with VLP-derived RNA from *S. cerevisiae* harboring the TMV-capsid encoding SC-TMV-CP and corresponding GFP-encoding plasmids Plasmid 341 ("SC_eGFP_mtRNR1"). The amount of GFP produced from THP-1 cells transformed with VLP-derived RNA obtained from *Y. lipolytica* transformed with Plasmid 336 ("YL_eGFP_AES") was substantially greater than the amount of GFP produced from THP-1 cells transformed with VLP-derived RNA obtained from *S. cerevisiae* transformed with Plasmid 342 ("SC_eGFP_AES"). However, no measure of statistical significance could be calculated because the low amount of VLP-derived RNA obtained from *S. cerevisiae* transformed with Plasmid 342 was only enough to run a single sample.

[0191] FIG. 34 shows the same statistically significant superior hEPO expression in THP-1 cells using VLP-derived RNA from *Y. lipolytica* as compared to corresponding VLP-derived RNA from *S. cerevisiae*. The mean amount of hEPO expressed in THP-1 cells transformed with VLP-derived RNA from *Y. lipolytica* transformed with Plasmid 338 ("YL-hEPO-mtRNR1") was 39-fold greater than the hEPO expressed from VLP-derived RNA from *S. cerevisiae* transformed with the corresponding plasmid (Plasmid 344; "SC_hEPO_mtRNR1") (7830 units vs 197 units of hEPO). The mean amount of hEPO expressed in THP-1 cells transformed with VLP-derived RNA from *Y. lipolytica* transformed with Plasmid 339 ("YL-hEPO-AES") was 26-fold greater than the hEPO expressed from VLP-derived RNA from *S. cerevisiae* transformed with the corresponding plasmid (Plasmid 345; "SC_hEPO_AES") (4120 units vs 157 units of hEPO).

[0192] Example 7. Immunogenicity of Heterologous RNA Molecules in Transformed Cells

[0193] The THP-1 Dual™ derived macrophages are genetically modified to secrete the enzyme Lucia luciferase if the IRF proinflammatory pathway is activated. Thus, the level of Lucia luciferase in the media is directly correlated to immune stimulation resulting from the transfection of those cells with the RNA samples. Twenty-four hours after transfection with RNA, 20 µL of the cell supernatant from the 96-well plates was removed and transferred to a white, opaque plate. The Lucia luciferase enzyme substrate coelenterazine was then added following the instructions from the kit QUANTI-Luc™-4 (Invivogen Cat:rep-qlc4ig1). The resulting bioluminescence was then read immediately by a plate reader.

[0194] FIG. 18 shows that the level of secreted Lucia luciferase from RNA derived from VLPs produced in *Y. lipolytica* transformed with YL-TMV-GFP UTR-1 was significantly lower than that observed from RNA derived from VLPs produced in *S. cerevisiae* co-transformed with SC-GFP UTR-1 and TMV-CP. The level of secreted Lucia luciferase from IVT-generated RNA that corresponded to the RNA produced in VLPs (templated from SEQ ID NO:35) and from commercially obtained IVT-generated RNA that encoded GFP was somewhat lower than that produced by the *Y. lipolytica* VLP-generated RNA.

[0195] FIG. 19 shows that the level of secreted Lucia luciferase from RNA derived from VLPs produced in *Y. lipolytica* transformed with YL-TMV-Luc UTR-4 or co-transformed with YL-Luc UTR-4 and YL-TMV-CP were significantly lower than that observed from RNA derived from VLPs produced in *S. cerevisiae* co-transformed with SC-Luc UTR-4 and TMV-CP. The *Y. lipolytica* YL-TMV-Luc UTR 4 VLP-generated RNA also produced somewhat lower levels of secreted Lucia luciferase as compared to IVT-generated RNA that corresponded to the RNA produced in VLPs (templated from SEQ ID NO:33) and from commercially obtained IVT-generated RNA that encoded firefly luciferase.

[0196] FIG. 20 shows that the level of secreted Lucia luciferase from RNA derived from VLPs produced in *Y. lipolytica* co-transformed with YL-TMV-CP and either YL-Luc OAS or YL-Luc UTR-4 was significantly less than the corresponding IVT-derived RNA.

[0197] FIG. 21 shows that the level of secreted Lucia luciferase from RNA derived from VLPs produced in *Y. lipolytica* transformed with YL-TMV-EPO UTR-4 was significantly less than RNA derived from VLPs produced in *S. cerevisiae* transformed with SC-EPO UTR-4.

[0198] The immunogenicity of RNA produced from TMV VLPs produced in *Y. lipolytica* and the corresponding RNA produced by IVT was further assessed using primary human PBMC derived dendritic cell (DCs). These cells are considered the most sensitive *in vitro* assay for immunogenicity and are routinely used in the safety assessment of therapeutics. DCs upregulate the expression of costimulatory molecules on their cell surface and secrete cytokines when stimulated by proinflammatory molecules. The upregulation of these receptors is measured using fluorescently labeled antibiotics and flow cytometry as described below.

[0199] Commercially sourced human donor PBMCs (iQBiosciences) were thawed and CD14(+) cells were isolated using magnetic beads (Miltenyi Biotec CD14 MircoBeads Cat:130-050-201). CD14(+) cells were then differentiated into immature dendritic cells by culturing in the presence of human IL-4 (Biolegend Cat:574002) and GM-CSF (Biolegend Cat:572905) RPMI, 10% heat-inactivated FBS for 9 days 250,000 cells per well in 250 µL. The maturation capability of the

dendritic cells was confirmed through treatment with a cytokine cocktail (PGE2 1μg/mL, IL1β 10ng/mL, IFNγ 20ng/mL, TNFα 50ng/mL) for 24 hours before flow cytometry analysis for the upregulation of CD80, 86, 54 and 40. RNA samples from VLPs produced in *Y. lipolytica* transformed with YL-TMV-GFP UTR-1 were prepared for transfection with Lipofectamine (Invitrogen cat.LMRNA008) reagent according to the manufacturer's instructions and immediately added to cells at a final concentration of 1 μg of RNA per well. A commercial sample of IVT-generated RNA was included as a control (Trilink CleanCap® EGFP-mRNA Cat:L-7601) and transfected at the same concentration as the *in vivo* made RNA samples. To account for any differences in purification treatment between the commercial and the *in vivo* made samples, *in vivo* RNAs had been additionally purified with Oligo d(T)25 beads (NEB, Cat: S1550S) and silica-based isolation, prior to transfection (Monarch RNA cleanup, NEB, Cat. T2040L). RNA concentrations were then remeasured by absorption at 260 nm. After 24 hours incubation with the RNA samples the dendritic cells were removed and assayed for GFP reporter expression and costimulatory molecule upregulation as result of immune stimulation by the RNA. The supernatants from the wells were removed and diluted twofold before IL-6 cytokine analysis by ELISA following the manufacturer's kit instructions (Biolegend Cat:430504) . Cells were stained for viability with Zombie Violet™ viability kit (Biolegend Cat:423113) before staining with anti CD80-APC and CD86-PE or anti CD40-APC and CD54-PE labeled antibodies (Biolegend Cat: 375404, 374205, 334309, 353105). The cells were then analyzed by flow cytometry, viable cells were gated and analyzed for GFP expression as a percentage of viable, GFP(+) cells. Viable cells were also analyzed for PE and APC fluorescence intensity.

[0200] FIG. 22A, shows that RNA isolated from VLPs produced from two separate batches of *Y. lipolytica* transformed with YL-TMV-GFP UTR-1 and a commercial IVT-produced RNA resulted in an equivalent percentage of GFP positive cells, measured by flow cytometry. FIG. 22B-FIG. 22E, show that the VLP-derived RNA from two separate batches was either non-immunogenic or very slightly immunogenic, as measured by upregulation of costimulatory molecules CD86, CD80, CD54, and CD40 over a negative control. In contrast, commercial IVT GFP RNA caused substantial upregulation of each of these costimulatory molecules. In addition, the VLP-derived RNA did not cause significant IL-6 cytokine secretion above the negative control, while the IVT-derived RNA caused a substantial increase (FIG. 22F).

[0201] Table 5. Sequences

SEQ ID NO:	Description	Sequence
1	native 5'-UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUA UAAGACCCCGGCGCCGCCACC
2	native human α-globin 3'-UTR	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGG GCCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGU ACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG C
3	modified human α-globin 3'-UTR	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGG GCCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGA UGGGGGCACAGAAACUUAUUUCAGACUCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
4	3' UTR sequence with two TMV 186 packaging sequences located within a native α-globin 3'-UTR and short polyA tail (UTR-4)	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGG GCCUCCCCCCCAGCCCCUCCUCCGUUGUUUAUAGAAAU AAUAUAAAAUUAGGUUUGAGAGAGAAGAUUACAAAC GUGAGAGACGGAGGGCCCAUGGAACUUACAGAAGAA GUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCG AUCAGGCUUGCAAAGUUUCGAUCUCGAACCGGAAAA AAGAGUGAUGUCCGCAAAGGGAAAAAUACCUUCCUG CACCCUUGUUUAUAGAAAUAAUAUAAAAUUAGGUUU GAGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCC CAUGGAACUUACAGAAGAAGUCGUUGAUGAGUUCAU GGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAAAGUU UCGAUCUCGAACCGGAAAAAAGAGUGAUGUCCGCAA AGGGAAAAAUAUACCCCCGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGCAAAAAAAAA
5	3' UTR sequence with two TMV 127 packaging sequences located within a native α-globin 3'-UTR having a short polyA tail	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGG GCCUCCCCCCCAGCCCCUCCUCCGGUUUGAGAGAGAAG AUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUU ACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUC CCUAUGUCGAUCAGGCUUGCAAAGUUUCGAUCUCGA ACCGGCCUUCCUGCACCCGUUUGAGAGAGAAGAUUA CAAACGUGAGAGACGGAGGGCCCAUGGAACUUACAG AAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUA UGUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCG GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGC GGCAAAAAAAAA

6	3' UTR sequence with two TMV 127 packaging sequences flanking a modified α-globin 3'-UTR with a polyA tail and ending with a polyA tail	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGAUGGGGGCACAGAAACUUA UUUCAGACUCACCCGUACCCCCGUGGUCUUUGAAUA AAGUCUGAGUGGGCGGCAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAGUUUGAGAGAG AAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGAA CUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAU GUCCCUAUGUCGAUCAGGCUUGCAAAGUUUCGAUCU CGAACCGGAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA
7	3' UTR sequence with two copies of a TMV 127 packaging sequence-modified α-globin 3'-UTR sequence and ending with a polyA tail	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGAUGGGGGCACAGAAACUUA UUUCAGACUCACCCGUACCCCCGUGGUCUUUGAAUA AAGUCUGAGUGGGCGGCGUUUGAGAGAGAAGAUUAC AAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGA AGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAU GUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGG GCCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGA UGGGGGCACAGAAACUUAUUUCAGACUCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGCA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAA
8	3' UTR sequence with two TMV 127 packaging sequences at the 5' and 3' end of a modified α-globin 3'-UTR sequence	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGAUGGGGGCACAGAAACUUA UUUCAGACUCACCCGUACCCCCGUGGUCUUUGAAUA AAGUCUGAGUGGGCGGCGUUUGAGAGAGAAGAUUAC AAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGA AGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAU GUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG

9	3' UTR sequence with two TMV 127 packaging sequences at the 5’ and 3’ end of a modified α-globin 3’-UTR sequence and ending with a polyA tail	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGAUGGGGGCACAGAAACUUA UUUCAGACUCACCCGUACCCCCGUGGUCUUUGAAUA AAGUCUGAGUGGGCGGCGUUUGAGAGAGAAGAUUAC AAACGUGAGAGACGGAGGGCCCAUGGAACUUACAGA AGAAGUCGUUGAUGAGUUCAUGGAAGAUGUCCCUAU GUCGAUCAGGCUUGCAAAGUUUCGAUCUCGAACCGG AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAA
10	3’-UTR with two TMV 186 packaging sequences connected by a linker and 5’ to a modified α-globin 3’-UTR sequence	UUGUUUAUAGAAAUAAUAUAAAAUUAGGUUUGAGAG AGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGG AACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAG AUGUCCCUAUGUCGAUCAGGCUUGCAAAGUUUCGAU CUCGAACCGGAAAAAAGAGUGAUGUCCGCAAAGGGA AAAAUACCUUCCUGCACCCUUGUUUAUAGAAAUAAU AUAAAAUUAGGUUUGAGAGAGAAGAUUACAAACGUG AGAGACGGAGGGCCCAUGGAACUUACAGAAGAAGUC GUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGAUC AGGCUUGCAAAGUUUCGAUCUCGAACCGGAAAAAAG AGUGAUGUCCGCAAAGGGAAAAAUAGCUGGAGCCUC GGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAG CCCCUCCUCCCCUUCCUGCACCCGAUGGGGGCACAGA AACUUAUUUCAGACUCACCCGUACCCCCGUGGUC1UU UGAAUAAAGUCUGAGUGGGCGGC
11	3’-UTR with two TMV 127 packaging sequences connected by a linker and 5’ to a modified α-globin 3’-UTR sequence	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGGCCUUCCUGCACCCGUUUG AGAGAGAAGAUUACAAACGUGAGAGACGGAGGGCCC AUGGAACUUACAGAAGAAGUCGUUGAUGAGUUCAUG GAAGAUGUCCCUAUGUCGAUCAGGCUUGCAAAGUUU CGAUCUCGAACCGGGCUGGAGCCUCGGUGGCCUAGC UUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCC CUUCCUGCACCCGAUGGGGGCACAGAAACUUAUUUC AGACUCACCCGUACCCCCGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGC

12	3’-UTR with two TMV 127 packaging sequences connected by a linker and flanked on both the 5’ and 3’-end by polyA stretches	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAGUUUGAGAGAGAAGAUUACAAACGUGAGAGAC GGAGGGCCCAUGGAACUUACAGAAGAAGUCGUUGAU GAGUUCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUU GCAAAGUUUCGAUCUCGAACCGGCCUUCCUGCACCCG UUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGG GCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGUU CAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAAA GUUUCGAUCUCGAACCGGAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAA
13	3’-UTR with one TMV 127 packaging sequence located within a native α-globin 3’-UTR and short polyA tail (UTR-1)	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGG GCCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGG UUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGG GCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGUU CAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAAA GUUUCGAUCUCGAACCGGUACCCCCGUGGUCUUUGA AUAAAGUCUGAGUGGGCGGCAAAAAAAAA
14	3’-UTR with one TMV 186 packaging sequence located within a native α-globin 3’-UTR and short polyA tail (UTR-3)	GCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGG GCCUCCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGU UGUUUAUAGAAAUAAUAUAAAAUUAGGUUUGAGAGA GAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGGA ACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGA UGUCCCUAUGUCGAUCAGGCUUGCAAAGUUUCGAUC UCGAACCGGAAAAAAGAGUGAUGUCCGCAAAGGGAA AAAUAUACCCCCGUGGUCUUUGAAUAAAGUCUGAGU GGGCGGCAAAAAAAAA
15	3’-UTR with one TMV127 packaging sequence located 5’ to a modified α-globin 3’-UTR (UTR-OAS)	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGAUGGGGGCACAGAAACUUA UUUCAGACUCACCCGUACCCCCGUGGUCUUUGAAUA AAGUCUGAGUGGGCGGC
16	75 nt TMV packaging sequence (TMV 75)	UGAGAGACGGAGGGCCCAUGGAACUUACAGAAGAAG UCGUUGAUGAGUUCAUGGAAGAUGUCCCUAUGUCGA UCA
17	127 nt TMV packaging sequence (TMV 127)	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGG
18	186 nt TMV packaging sequence (TMV 186)	UUGUUUAUAGAAAUAAUAUAAAAUUAGGUUUGAGAG AGAAGAUUACAAACGUGAGAGACGGAGGGCCCAUGG AACUUACAGAAGAAGUCGUUGAUGAGUUCAUGGAAG AUGUCCCUAUGUCGAUCAGGCUUGCAAAGUUUCGAU CUCGAACCGGAAAAAAGAGUGAUGUCCGCAAAGGGA AAAAUA

19	99 nt TMV packaging sequence (TMV 99)	AUUACAAACGUGAGAGACGGAGGGCCCAUGGAACUU ACAGAAGAAGUCGUUGAUGAGUUCAUGGAAGAUGUC CCUAUGUCGAUCAGGCUUGCAAAGUUU
20	234 nt TMV packaging sequence (TMV 234)	GCAAGUUUUAGUUAAUAUUAGAAAUGUGAAGAUGUC AGCGGGUUUCUGUCCGCUUUCUCUGGAGUUUGUGUC GGUGUGUAUUGUUUAUAGAAAUAAUAUAAAAUUAGG UUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAGG GCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGUU CAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAAA GUUUCGAUCUCGAACCGG
21	GAP Promoter	TCATTATCAATACTGCCATTTCAAAGAATACGTAAATA ATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAA AAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCC CAAAATAGGGGGCGGGTTACACAGAATATATAACATC GTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCAC TAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAG AAAAAAAAAGAATCCCAGCACCAAAATATTGTTTTCTT CACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCA ACTACAGAGAACAGGGGCACAAACAGGCAAAAAACG GGCACAACCTCAATGGAGTGATGCAACCTGCCTGGAG TAAATGATGACACAAGGCAATTGACCCACGCATGTAT CTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTC TCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACC AGTTCCCTGAAATTATTCCCCTACTTGACTAATAAGTA TATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATC TATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAGT CTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGA ATAAACACACATAAACAAACAAA
22	TEF1 promoter	CCACACACCATAGCTTCAAAATGTTTCTACTCCTTTTTT ACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTA CCACTTCAAAACACCCAAGCACAGCATACTAAATTTCC CCTCTTTCTTCCTCTAGGGTGTCGTTAATTACCCGTACT AAAGGTTTGGAAAAGAAAAAAGACACCGCCTCGTTTC TTTTTCTTCGTCGAAAAAGGCAATAAAAATTTTTATCA CGTTTCTTTTTCTTGAAAATTTTTTTTTTGATTTTTTTC TCTTTCGATGACCTCCCATTGATATTTAAGTTAATAAA CGGTCATCAATTTCTCAAGTTTCAGTTTCATTTTTCTTG TTCTATTACAACTTTTTTTACTTCTTGCTCATTAGAAAG AAAGCATAGCAATCTAATCTAAG

23	modified UAS8-2 promoter	GCTTCGATCGCGTATCGATGATACGCGTCCATGGTGAG GTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGC TTCTCTTTGTGTGTAGTGTACGTACATTATCGAGAGGG TTGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTC TCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTC TTTGTGTGTAGTGTACGTACATTATCGAGAGGGTTGTT CCCGCCCACCTCGATCCGACATGTTGAGGTGTCTCACA AGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTG TGTAGTGTACGTACATTATCGAGAGGGTTGTTCCCGCC CACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCC GTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGT GTACGTACATTATCGAGAGGGTTGTTCCCGCCCACCTC GATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAG TCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGT ACATTATCGAGAGGGTTGTTCCCGCCCACCTCGATCCG GCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGC CCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTA TCGAGAGGGTTGTTCCCGCCCACCTCGATCCGGCATGC ACTGATCACGGGCAAAAGTGCTTCGATAGAGAGGGGG TTGGCGGCGCATTTGTGTCCCAAAAAACAGCCCCAATT GCCCCAATTGACCCCAAATTGACCCAGTAGCGGACCC AACCCCGGCGAGAGCCCCCTTCACCCCACATATCAAAC CTCCCCCGGTTCCCACACTTGCCGTTAAGGGCGTAGGG TACTGCAGTCTGGAATCTACGCTTGTTCAGACTTTGTA CTAGTTTCTTTGTCTGGCCATCCGGGTAACCCATGCCG GACGCAAAATAGACTACTGAAAATTTTTTTGCTTTGTG GTTGGGACTTTAGCCAAGGGTATAAAAGACCACCGTC CCCGAATTACCTTTCCTCTTCTTTTCTCTCTCTCCTTGTC AACTCACACCCGAAG

24	UAS4-1 promoter	ACGGCGATACGCGTATCGATACGCGTGCATGCTGAGG TGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCT TCTCTTTGTGTGTAGTGTACGTACATTATCGAGAGGGT TGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTCT CACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCT TTGTGTGTAGTGTACGTACATTATCGAGAGGGTTGTTC CCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAA GTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGT GTAGTGTACGTACATTATCGAGAGGGTTGTTCCCGCCC ACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCG TGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTG TACGTACATTATCGAGAGGGTTGTTCCCGCCCACCTCG ATCCGGCATGCACTGATCACGGGCAAAAGTGCGTTCG ATAGAGAGGGGGTTGGCGGCGCATTTGTGTCCCAAAA AACAGCCCCAATTGCCCCAATTGACCCCAAATTGACCC AGTAGCGGACCCAACCCCGGCGAGAGCCCCCTTCACC CCACATATCAAACCTCCCCCGGTTCCCACACTTGCCGT TAAGGGCGTAGGGTACTGCAGTCTGGAATCTACGCTTG TTCAGACTTTGTACTAGTTTCTTTGTCTGGCCATCCGGG TAACCCATGCCGGACGCAAAATAGACTACTGAAAATT TTTTTGCTTTGTGGTTGGGACTTTAGCCAAGGGTATAA AAGACCACCGTCCCCGAATTACCTTTCCTCTTCTTTTCT CTCTCTCCTTGTCAACTCACACCCGAAGGATCCAATG
25	CYC1 terminator	AUCCGCUCUAACCGAAAAGGAAGGAGUUAGACAACC UGAAGUCUAGGUCCCUAUUUAUUUUUUUAUAGUUAU GUUAGUAUUAAGAACGUUAUUUAUAUUUCAAAUUUU UCUUUUUUUUCUGUACAGACGCGUGUACGCAUGUAA CAUUAUACUGAAAACCUUGCUUGAGAAGGUUUUGGG ACGCUCGAAG
26	TTEF terminator	UCUAGCUGCUUGUACCUAUGCAACCCCAGUUUGUUA AAAAUUAGUAGUCAAAAACUUCUGAGUUAAAAAAAAA AAAAAAAAGCUU
27	TLip2-2 terminator	GGAUGUGUCUGUGGUAUCUAAGCUAUUUAUCACUCU UUACAACUUCUACCUCAACUAUCUACUUUAAUAAAU GAAUAUCGUUUAUUCUCUAUGAUUACUGUAUAUGCG UUCCUCUAAGACAGAGU
28	TMV capsid protein	MSYSITTPSQFVFLSSAWADPIELINLCTNALGNQFQTQQ ARTVVQRQFSEVWKPSPQVTVRFPDSDFKVYRYNAVLD PLVTALLGAFDTRNRIIEVENQANPTTAETLDATRRVDDA TVAIRSAINNLIVELIRGTGSYNRSSFESSSGLVWTSGPAT

29	DNA encoding TMV capsid protein	ATGTCTTACAGTATCACTACTCCATCTCAGTTCGTGTTC TTGTCATCAGCGTGGGCCGACCCAATAGAGTTAATTAA TTTATGTACTAATGCCTTAGGAAATCAGTTTCAAACAC AACAAGCTCGAACTGTCGTTCAAAGACAATTCAGTGA GGTGTGGAAACCTTCACCACAAGTAACTGTTAGGTTCC CTGACAGTGACTTTAAGGTGTACAGGTACAATGCGGTA TTAGACCCGCTAGTCACAGCACTGTTAGGTGCATTCGA CACTAGAAATAGAATAATAGAAGTTGAAAATCAGGCG AACCCCACGACTGCCGAAACGTTAGATGCTACTCGTAG AGTAGACGACGCAACGGTGGCCATAAGGAGCGCGATA AATAATTTAATAGTAGAATTGATCAGAGGAACCGGAT CTTATAATCGGAGCTCTTTCGAGAGCTCTTCTGGTTTG GTTTGGACCTCTGGTCCTGCAACTTGA
30	DNA sequence encoding eGFP	ATGGTCAGCAAAGGAGAGGAACTGTTCACTGGAGTCG TGCCTATTCTGGTCGAACTGGATGGAGATGTCAATGGG CATAAATTCAGCGTGTCCGGAGAGGGAGAGGGAGACG CAACCTACGGCAAGCTGACACTGAAGTTCATCTGCACC ACAGGCAAGCTGCCCGTGCCTTGGCCAACCCTGGTGAC CACACTGACATACGGCGTGCAGTGTTTTAGCAGGTATC CAGACCACATGAAGCAGCACGATTTCTTTAAGTCCGCC ATGCCAGAGGGATACGTGCAGGAGCGGACCATCTTCT TTAAGGACGATGGCAATTATAAGACCAGAGCCGAGGT GAAGTTCGAGGGCGACACACTGGTGAACAGGATCGAG CTGAAGGGCATCGACTTTAAGGAGGATGGCAATATCC TGGGCCACAAGCTGGAGTACAACTATAATTCTCACAAC GTGTATATCATGGCCGATAAGCAGAAGAACGGCATCA AAGTGAATTTCAAGATCCGCCACAATATCGAGGACGG CAGCGTGCAGCTGGCCGATCACTACCAGCAGAACACC CCTATCGGAGACGGACCCGTGCTGCTGCCTGATAATCA CTATCTGAGCACACAGTCCGCCCTGTCTAAGGACCCCA ACGAGAAGCGGGATCACATGGTGCTGCTGGAATTTGT CACCGCTGCTGGAATCACACTGGGAATGGATGAGCTG TATAAATAA

31	DNA sequence encoding luciferase	ATGGAGGACGCCAAGAACATCAAGAAGGGCCCCGCCC CCTTCTACCCCCTGGAGGACGGCACCGCCGGCGAGCA GCTGCACAAGGCCATGAAGCGGTACGCCCTGGTGCCC GGCACCATCGCCTTCACCGACGCCCACATCGAGGTGG ACATCACCTACGCCGAGTACTTCGAGATGAGCGTGCG GCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACC AACCACCGGATCGTGGTGTGCAGCGAGAACAGCCTGC AGTTCTTCATGCCCGTGCTGGGCGCCCTGTTCATCGGC GTGGCCGTGGCCCCCGCCAACGACATCTACAACGAGC GGGAGCTGCTGAACAGCATGGGCATCAGCCAGCCCAC CGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAAGATC CTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAAGA TCATCATCATGGACAGCAAGACCGACTACCAGGGCTTC CAGAGCATGTACACCTTCGTGACCAGCCACCTGCCCCC CGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCTTCG ACCGGGACAAGACCATCGCCCTGATCATGAACAGCAG CGGCAGCACCGGCCTGCCCAAGGGCGTGGCCCTGCCC CACCGGACCGCCTGCGTGCGGTTCAGCCACGCCCGAG ACCCCATCTTCGGCAACCAGATCATCCCCGACACCGCC ATCCTGAGCGTGGTGCCCTTCCACCACGGCTTCGGCAT GTTCACCACCCTGGGCTACCTGATCTGCGGCTTCCGGG TGGTGCTGATGTACCGGTTCGAGGAGGAGCTGTTCCTG CGGAGCCTGCAGGACTACAAGATCCAGAGCGCCCTGC TGGTGCCCACCCTGTTCAGCTTCTTCGCCAAGAGCACC CTGATCGACAAGTACGACCTGAGCAACCTGCACGAGA TCGCCAGCGGCGGCGCCCCCCTGAGCAAGGAGGTGGG CGAGGCCGTGGCCAAGCGGTTCCACCTGCCCGGCATCC GGCAGGGCTACGGCCTGACCGAGACCACCAGCGCCAT CCTGATCACCCCCGAGGGCGACGACAAGCCCGGCGCC GTGGGCAAGGTGGTGCCCTTCTTCGAGGCCAAGGTGGT GGACCTGGACACCGGCAAGACCCTGGGCGTGAACCAG CGGGGCGAGCTGTGCGTGCGGGGCCCCATGATCATGA GCGGCTACGTGAACAACCCCGAGGCCACCAACGCCCT GATCGACAAGGACGGCTGGCTGCACAGCGGCGACATC GCCTACTGGGACGAGGACGAGCACTTCTTCATCGTGGA CCGGCTGAAGAGCCTGATCAAGTACAAGGGCTACCAG GTGGCCCCCGCCGAGCTGGAGAGCATCCTGCTGCAGC ACCCCAACATCTTCGACGCCGGCGTGGCCGGCCTGCCC GACGACGACGCCGGCGAGCTGCCCGCCGCCGTGGTGG TGCTGGAGCACGGCAAGACCATGACCGAGAAGGAGAT CGTGGACTACGTGGCCAGCCAGGTGACCACCGCCAAG AAGCTGCGGGGCGGCGTGGTGTTCGTGGACGAGGTGC CCAAGGGCCTGACCGGCAAGCTGGACGCCCGGAAGAT CCGGGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAG ATCGCCGTGTGA

32	DNA sequence encoding erythropoietin	ATGGGGGTGCACGAATGTCCTGCCTGGCTGTGGCTTCT CCTGTCCCTGCTGTCGCTCCCTCTGGGCCTCCCAGTCCT GGGCGCCCCACCACGCCTCATCTGTGACAGCCGAGTCC TGGAGAGGTACCTCTTGGAGGCCAAGGAGGCCGAGAA TATCACGACGGGCTGTGCTGAACACTGCAGCTTGAATG AGAATATCACTGTCCCAGACACCAAAGTTAATTTCTAT GCCTGGAAGAGGATGGAGGTCGGGCAGCAGGCCGTAG AAGTCTGGCAGGGCCTGGCCCTGCTGTCGGAAGCTGTC CTGCGGGGCCAGGCCCTGTTGGTCAACTCTTCCCAGCC GTGGGAGCCCCTGCAGCTGCATGTGGATAAAGCCGTC AGTGGCCTTCGCAGCCTCACCACTCTGCTTCGGGCTCT GGGAGCCCAGAAGGAAGCCATCTCCCCTCCAGATGCG GCCTCAGCTGCTCCACTCCGAACAATCACTGCTGACAC TTTCCGCAAACTCTTCCGAGTCTACTCCAATTTCCTCCG GGGAAAGCTGAAGCTGTACACAGGGGAGGCCTGCAGG ACAGGGGACAGATGA

33	IVT DNA template comprising a T7 promoter with an additional 3'-adenosine, the 5'-UTR of SEQ ID NO:1, the luciferase coding sequence of SEQ ID NO:31, a DNA sequence corresponding to the 3'-UTR of SEQ ID NO:4, and a DNA sequence corresponding to a polyA tail	GACGTAATACGACTCACTATAAGGGAAATAAGAGAGA AAAGAAGAGTAAGAAGAAATATAAGACCCCGGCGCCG CCACCATGGAGGACGCCAAGAACATCAAGAAGGGCCC CGCCCCCTTCTACCCCCTGGAGGACGGCACCGCCGGCG AGCAGCTGCACAAGGCCATGAAGCGGTACGCCCTGGT GCCCGGCACCATCGCCTTCACCGACGCCCACATCGAGG TGGACATCACCTACGCCGAGTACTTCGAGATGAGCGTG CGGCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACA CCAACCACCGGATCGTGGTGTGCAGCGAGAACAGCCT GCAGTTCTTCATGCCCGTGCTGGGCGCCCTGTTCATCG GCGTGGCCGTGGCCCCCGCCAACGACATCTACAACGA GCGGGAGCTGCTGAACAGCATGGGCATCAGCCAGCCC ACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAAGA TCCTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAA GATCATCATCATGGACAGCAAGACCGACTACCAGGGC TTCCAGAGCATGTACACCTTCGTGACCAGCCACCTGCC CCCCGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCT TCGACCGGGACAAGACCATCGCCCTGATCATGAACAG CAGCGGCAGCACCGGCCTGCCCAAGGGCGTGGCCCTG CCCCACCGGACCGCCTGCGTGCGGTTCAGCCACGCCCG AGACCCCATCTTCGGCAACCAGATCATCCCCGACACCG CCATCCTGAGCGTGGTGCCCTTCCACCACGGCTTCGGC ATGTTCACCACCCTGGGCTACCTGATCTGCGGCTTCCG GGTGGTGCTGATGTACCGGTTCGAGGAGGAGCTGTTCC TGCGGAGCCTGCAGGACTACAAGATCCAGAGCGCCCT GCTGGTGCCCACCCTGTTCAGCTTCTTCGCCAAGAGCA CCCTGATCGACAAGTACGACCTGAGCAACCTGCACGA GATCGCCAGCGGCGGCGCCCCCCTGAGCAAGGAGGTG GGCGAGGCCGTGGCCAAGCGGTTCCACCTGCCCGGCA TCCGGCAGGGCTACGGCCTGACCGAGACCACCAGCGC CATCCTGATCACCCCCGAGGGCGACGACAAGCCCGGC GCCGTGGGCAAGGTGGTGCCCTTCTTCGAGGCCAAGGT GGTGGACCTGGACACCGGCAAGACCCTGGGCGTGAAC CAGCGGGGCGAGCTGTGCGTGCGGGGCCCCATGATCA TGAGCGGCTACGTGAACAACCCCGAGGCCACCAACGC CCTGATCGACAAGGACGGCTGGCTGCACAGCGGCGAC ATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT GGACCGGCTGAAGAGCCTGATCAAGTACAAGGGCTAC CAGGTGGCCCCCGCCGAGCTGGAGAGCATCCTGCTGC AGCACCCCAACATCTTCGACGCCGGCGTGGCCGGCCTG CCCGACGACGACGCCGGCGAGCTGCCCGCCGCCGTGG TGGTGCTGGAGCACGGCAAGACCATGACCGAGAAGGA GATCGTGGACTACGTGGCCAGCCAGGTGACCACCGCC AAGAAGCTGCGGGGCGGCGTGGTGTTCGTGGACGAGG TGCCCAAGGGCCTGACCGGCAAGCTGGACGCCCGGAA GATCCGGGAGATCCTGATCAAGGCCAAGAAGGGCGGC AAGATCGCCGTGTGAGCTGGAGCCTCGGTGGCCTAGCT TCTTGCCCCTTGGGCCTCCCCCCAGCCCCTCCTCCGTTG TTTATAGAAATAATATAAAATTAGGTTTGAGAGAGAA GATTACAAACGTGAGAGACGGAGGGCCCATGGAACTT

		ACAGAAGAAGTCGTTGATGAGTTCATGGAAGATGTCC CTATGTCGATCAGGCTTGCAAAGTTTCGATCTCGAACC GGAAAAAAGAGTGATGTCCGCAAAGGGAAAAATACCT TCCTGCACCCTTGTTTATAGAAATAATATAAAATTAGG TTTGAGAGAGAAGATTACAAACGTGAGAGACGGAGGG CCCATGGAACTTACAGAAGAAGTCGTTGATGAGTTCAT GGAAGATGTCCCTATGTCGATCAGGCTTGCAAAGTTTC GATCTCGAACCGGAAAAAAGAGTGATGTCCGCAAAGG GAAAAATATACCCCCGTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAA

34	IVT DNA template comprising a T7 promoter with an additional 3'-adenosine, a DNA sequence corresponding to the 5'-UTR of SEQ ID NO:1, the luciferase coding sequence of SEQ ID NO:31, a DNA sequence corresponding to the 3'-UTR of SEQ ID NO:15, and a DNA sequence corresponding to a polyA tail	GACGTAATACGACTCACTATAAGGGAAATAAGAGAGA AAAGAAGAGTAAGAAGAAATATAAGACCCCGGCGCCG CCACCATGGAGGACGCCAAGAACATCAAGAAGGGCCC CGCCCCCTTCTACCCCCTGGAGGACGGCACCGCCGGCG AGCAGCTGCACAAGGCCATGAAGCGGTACGCCCTGGT GCCCGGCACCATCGCCTTCACCGACGCCCACATCGAGG TGGACATCACCTACGCCGAGTACTTCGAGATGAGCGTG CGGCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACA CCAACCACCGGATCGTGGTGTGCAGCGAGAACAGCCT GCAGTTCTTCATGCCCGTGCTGGGCGCCCTGTTCATCG GCGTGGCCGTGGCCCCCGCCAACGACATCTACAACGA GCGGGAGCTGCTGAACAGCATGGGCATCAGCCAGCCC ACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAAGA TCCTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAA GATCATCATCATGGACAGCAAGACCGACTACCAGGGC TTCCAGAGCATGTACACCTTCGTGACCAGCCACCTGCC CCCCGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCT TCGACCGGGACAAGACCATCGCCCTGATCATGAACAG CAGCGGCAGCACCGGCCTGCCCAAGGGCGTGGCCCTG CCCCACCGGACCGCCTGCGTGCGGTTCAGCCACGCCCG AGACCCCATCTTCGGCAACCAGATCATCCCCGACACCG CCATCCTGAGCGTGGTGCCCTTCCACCACGGCTTCGGC ATGTTCACCACCCTGGGCTACCTGATCTGCGGCTTCCG GGTGGTGCTGATGTACCGGTTCGAGGAGGAGCTGTTCC TGCGGAGCCTGCAGGACTACAAGATCCAGAGCGCCCT GCTGGTGCCCACCCTGTTCAGCTTCTTCGCCAAGAGCA CCCTGATCGACAAGTACGACCTGAGCAACCTGCACGA GATCGCCAGCGGCGGCGCCCCCCTGAGCAAGGAGGTG GGCGAGGCCGTGGCCAAGCGGTTCCACCTGCCCGGCA TCCGGCAGGGCTACGGCCTGACCGAGACCACCAGCGC CATCCTGATCACCCCCGAGGGCGACGACAAGCCCGGC GCCGTGGGCAAGGTGGTGCCCTTCTTCGAGGCCAAGGT GGTGGACCTGGACACCGGCAAGACCCTGGGCGTGAAC CAGCGGGGCGAGCTGTGCGTGCGGGGCCCCATGATCA TGAGCGGCTACGTGAACAACCCCGAGGCCACCAACGC CCTGATCGACAAGGACGGCTGGCTGCACAGCGGCGAC ATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT GGACCGGCTGAAGAGCCTGATCAAGTACAAGGGCTAC CAGGTGGCCCCCGCCGAGCTGGAGAGCATCCTGCTGC AGCACCCCAACATCTTCGACGCCGGCGTGGCCGGCCTG CCCGACGACGACGCCGGCGAGCTGCCCGCCGCCGTGG TGGTGCTGGAGCACGGCAAGACCATGACCGAGAAGGA GATCGTGGACTACGTGGCCAGCCAGGTGACCACCGCC AAGAAGCTGCGGGGCGGCGTGGTGTTCGTGGACGAGG TGCCCAAGGGCCTGACCGGCAAGCTGGACGCCCGGAA GATCCGGGAGATCCTGATCAAGGCCAAGAAGGGCGGC AAGATCGCCGTGTGAGTTTGAGAGAGAAGATTACAAA CGTGAGAGACGGAGGGCCCATGGAACTTACAGAAGAA GTCGTTGATGAGTTCATGGAAGATGTCCCTATGTCGAT CAGGCTTGCAAAGTTTCGATCTCGAACCGGGCTGGAGC

		CTCGGTGGCCTAGCTTCTTGCCCCTTGGGCCTCCCCCC AGCCCCTCCTCCCCTTCCTGCACCCGATGGGGGCACAG AAACTTATTTCAGACTCACCCGTACCCCCGTGGTCTTT GAATAAAGTCTGAGTGGGCGGCAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA
35	IVT DNA template comprising a T7 promoter with an additional 3'-adenosine, a DNA sequence corresponding to the 5'-UTR of SEQ ID NO:1, the GFP coding sequence of SEQ ID NO:30, a DNA sequence corresponding to the 3'-UTR of SEQ ID NO:13, and a DNA sequence corresponding to a polyA tail	GACGTAATACGACTCACTATAAGGGAAATAAGAGAGA AAAGAAGAGTAAGAAGAAATATAAGACCCCGGCGCCG CCACCATGGTCAGCAAAGGAGAGGAACTGTTCACTGG AGTCGTGCCTATTCTGGTCGAACTGGATGGAGATGTCA ATGGGCATAAATTCAGCGTGTCCGGAGAGGGAGAGGG AGACGCAACCTACGGCAAGCTGACACTGAAGTTCATC TGCACCACAGGCAAGCTGCCCGTGCCTTGGCCAACCCT GGTGACCACACTGACATACGGCGTGCAGTGTTTTAGCA GGTATCCAGACCACATGAAGCAGCACGATTTCTTTAAG TCCGCCATGCCAGAGGGATACGTGCAGGAGCGGACCA TCTTCTTTAAGGACGATGGCAATTATAAGACCAGAGCC GAGGTGAAGTTCGAGGGCGACACACTGGTGAACAGGA TCGAGCTGAAGGGCATCGACTTTAAGGAGGATGGCAA TATCCTGGGCCACAAGCTGGAGTACAACTATAATTCTC ACAACGTGTATATCATGGCCGATAAGCAGAAGAACGG CATCAAAGTGAATTTCAAGATCCGCCACAATATCGAG GACGGCAGCGTGCAGCTGGCCGATCACTACCAGCAGA ACACCCCTATCGGAGACGGACCCGTGCTGCTGCCTGAT AATCACTATCTGAGCACACAGTCCGCCCTGTCTAAGGA CCCCAACGAGAAGCGGGATCACATGGTGCTGCTGGAA TTTGTCACCGCTGCTGGAATCACACTGGGAATGGATGA GCTGTATAAATAAGCTGGAGCCTCGGTGGCCTAGCTTC TTGCCCCTTGGGCCTCCCCCCAGCCCCTCCTCCCCTTCC TGCACCCGGTTTGAGAGAGAAGATTACAAACGTGAGA GACGGAGGGCCCATGGAACTTACAGAAGAAGTCGTTG ATGAGTTCATGGAAGATGTCCCTATGTCGATCAGGCTT GCAAAGTTTCGATCTCGAACCGGTACCCCCGTGGTCTT TGAATAAAGTCTGAGTGGGCGGCAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA
36	pTEF promoter	ACCGGGTTGGCGGCGTATTTGTGTCCCAAAAAACAGCC CCAATTGCCCCAATTGACCCCAAATTGACCCAGTAGCG GGCCCAACCCCGGCGAGAGCCCCCTTCACCCCACATAT CAAACCTCCCCCGGTTCCCACACTTGCCGTTAAGGGCG TAGGGTACTGCAGTCTGGAATCTACGCTTGTTCAGACT TTGTACTAGTTTCTTTGTCTGGCCATCCGGGTAACCCAT GCCGGACGCAAAATAGACTACTGAAAATTTTTTTGCTT TGTGGTTGGGACTTTAGCCAAGGGTATAAAAGACCAC CGTCCCCGAATTACCTTTCCTCTTCTTTTCTCTCTCCT TGTCAACTCACACCCGAAATCGTTAAGCATTTCCTTCT GAGTATAAGAATCATTCAA

37	T1-3 Lipt2 terminator	GUGUCUGUGGUAUCUAAGCUAUUUAUCACUCUUUAC AACUUCUACCUCAACUAUCUACUUUAAUAAAUGAAU AUCGUUUAUUCUCUAUGAUUACUGUAUAUGCGUUCC UCUAAGACA
38	T7 promoter	GACGTAATACGACTCACTATA
39	3'-UTR from the Mitochondrial Encoded 12S RNA gene (mtRNR1)	CAAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCU AGCCACACCCCCACGGGAAACAGCAGUGAUUAACCUU UAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAAC CCCAGGGUUGGUCAAUUUCGUGCCAGCCACACC
40	3'-UTR from the amino-terminal enhancer of split gene (AES)	CUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUC CCGUCCUGGGUACCCCGAGUCUCCCCCGACCUCGGGU CCCAGGUAUGCUCCCACCUCCACCUGCCCCACUCACC ACCUCUGCUAGUUCCAGACACCUCC
41	3'-UTR from the magnesium transporter MRS2 gene (MRS2)	CUGCUGCCUGCUUCUUGCUCCAGCACCAUGGAAUGCC UGCGCAGUUUACCCUGCCUCCUGCCCCGCGCGAUGAG ACUUCCCCGGCGGACGCUGUGUGCCCUGGCCUUGGAC GUGACCUCUGUGGGUCCUCCCGUUGCUGCCU
42	3'-UTR with one TMV127 packaging sequence located 5' to a mtRNR1 3'-UTR (UTR-B_mtRNR1)	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGGCAAGCACGCAGCAAUGC AGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGA AACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGU UUAACUAAGCUAUACUAACCCCAGGGUUGGUCAAUU UCGUGCCAGCCACACC
43	3'-UTR with one TMV127 packaging sequence located 5' to a AES 3'-UTR (UTR-B_AES)	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGGCUGGUACUGCAUGCACG CAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCG AGUCUCCCCCGACCUCGGGUCCCAGGUAUGCUCCCAC CUCCACCUGCCCCACUCACCACCUCUGCUAGUUCCAG ACACCUCC
44	3'-UTR with one TMV127 packaging sequence located 5' to a MRS2 3'-UTR (UTR-B_MRS2)	GUUUGAGAGAGAAGAUUACAAACGUGAGAGACGGAG GGCCCAUGGAACUUACAGAAGAAGUCGUUGAUGAGU UCAUGGAAGAUGUCCCUAUGUCGAUCAGGCUUGCAA AGUUUCGAUCUCGAACCGGCUGCUGCCUGCUUCUUG CUCCAGCACCAUGGAAUGCCUGCGCAGUUUACCCUGC CUCCUGCCCCGCGCGAUGAGACUUCCCCGGCGGACGC UGUGUGCCCUGGCCUUGGACGUGACCUCUGUGGGUC CUCCCGUUGCUGCCU

[0202] Although the invention has been described with reference to the presently preferred embodiment, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

What is Claimed Is:

1. A *Yarrowia sp.* cell comprising:
   a. a first promoter operatively linked to a DNA sequence capable of being transcribed into a heterologous mRNA molecule that comprises, in a 5' to 3' order, a 5'-UTR, a heterologous RNA payload, and a 3'-UTR, wherein the 3'-UTR comprises one or more tobacco mosaic virus (TMV) packaging signal; and
   b. a second promoter operatively linked to a DNA sequence encoding a TMV viral capsid protein.

2. The *Yarrowia sp.* cell of claim 1, wherein the cell harbors a first vector comprising the DNA sequence capable of being transcribed into the heterologous RNA molecule; and a second vector comprising the DNA sequence encoding a TMV viral capsid protein.

3. The *Yarrowia sp.* cell of claim 1, wherein the cell harbors a vector comprising both the DNA sequence capable of being transcribed into the heterologous RNA molecule; and the DNA sequence encoding a TMV viral capsid protein.

4. The *Yarrowia sp.* cell of any one of claims 1-3, wherein the heterologous RNA molecule contains only one TMV packaging signal.

5. The *Yarrowia sp.* cell of any one of claims 1-3, wherein the heterologous RNA molecule contains only two or more TMV packaging signals.

6. The *Yarrowia sp.* cell of any one of claims 1-5, wherein the heterologous RNA payload encodes a mammalian polypeptide or protein.

7. The *Yarrowia sp.* cell of any one of claims 1-5, wherein the heterologous RNA payload does not encode a polypeptide.

8. The *Yarrowia sp.* cell of any one of claims 1-7, wherein the cell is a *Yarrowia lipolytica* cell.

9. The *Yarrowia sp.* cell of any one of claims 1-8, wherein the 5'-UTR comprises SEQ ID NO:1.

10. The *Yarrowia sp.* cell of any one of claims 1-3 or 5-9, wherein the 3'-UTR comprises a nucleotide sequence selected from any one of SEQ ID NOs:4-12, or an RNA sequence having at least 90% sequence identity to any one of SEQ ID NOs:4-12 and comprising at least two copies of the minimal TMV packaging sequence of SEQ ID NO:16.

11. The *Yarrowia sp.* cell of claim 10, wherein the 3'-UTR comprises any one of SEQ ID NOs: 4-12.

12. The *Yarrowia sp.* cell of claim 11, wherein the 3'-UTR comprises SEQ ID NO:4.

13. The *Yarrowia sp.* cell of any one of claims 1-4 or 6-9, wherein the 3'-UTR comprises a nucleotide sequence selected from any one of SEQ ID NOs:13-15 or 42-44, or an RNA sequence having at least 90% sequence identity to any one of SEQ ID NOs:13-15 or 42-44 and comprising one copy of the minimal TMV packaging sequence of SEQ ID NO:16.

14. The *Yarrowia sp.* cell of claim 13, wherein the 3'-UTR comprises any one of SEQ ID NOs:13-15.

15. The *Yarrowia sp.* cell of claim 14, wherein the 3'-UTR comprises SEQ ID NO:13 or SEQ ID NO:15.

16. The *Yarrowia sp.* cell of claim 13, wherein the 3'-UTR comprises any one of SEQ ID NOs:42-44.

17. The *Yarrowia sp.* cell of any one of claims 1-16, wherein each TMV packaging signal present comprises a nucleotide sequence independently selected from any one of SEQ ID NOs:16-20, or a nucleotide sequence comprising SEQ ID NO:16 and having at least 90% sequence identity to any one of SEQ ID NOs:16-20.

18. The *Yarrowia sp.* cell of claim 17, wherein each TMV packaging signal present is independently selected from any one of SEQ ID NOs:16-20.

19. The *Yarrowia sp.* cell of claim 17 or 18, wherein each TMV packaging signal present is the same.

20. The *Yarrowia sp.* cell of claim 19, wherein each TMV packaging signal present is SEQ ID NO:17 or SEQ ID NO:18.

21. The *Yarrowia sp.* cell of any one of claims 1-20, wherein each of the first promoter and second promoter comprises a nucleotide sequence independently selected from SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:36, or a sequence having at least 90% sequence identity to any one of SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:36.

22. The *Yarrowia sp.* cell of claim 21, wherein each of the first promoter and second promoter comprises a nucleotide sequence independently selected from SEQ ID NO:23, SEQ ID NO:24 or SEQ ID NO:36.

23. The *Yarrowia sp.* cell of any one of claims 1-22, wherein each of the DNA sequence capable of being transcribed into a heterologous mRNA molecule and the DNA sequence encoding a TMV viral capsid protein further encodes upon transcription a terminator, wherein each terminator has a nucleotide sequence independently selected from SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:37, or a sequence having at least 90% sequence identity to any one of SEQ ID NO:25, SEQ I DNO:26, SEQ ID NO:27, or SEQ ID NO:37.

24. The *Yarrowia sp.* cell of claim 23, wherein the nucleotide sequence of each terminator is independently selected from SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, or SEQ ID NO:37.

25. A method of producing a heterologous mRNA molecule comprising the step of growing the *Yarrowia sp.* cell of any one of claims 1-24 under conditions in which the heterologous RNA molecule and the TMV capsid protein are expressed and, through the presence of a TMV packaging signal, the heterologous RNA molecule is bound to and encapsidated by the TMV

viral capsid protein in a virus-like particle, thereby producing the heterologous mRNA molecule.

26. The method of claim 25, further comprising the step of isolating the virus-like particle from the *Yarrowia sp.* cell.

27. The method of claim 26, further comprising treating the isolated virus-like particles with one or more RNases or nucleases that digest double-stranded RNA.

28. The method of claim 26 or 27, further comprising purifying the encapsidated RNA from the isolated virus-like particles.

29. An isolated virus-like particle produced by the method of any one of claims 25-27.

30. An isolated RNA molecule produced by the method of claim 28.