Journal of Biotechnology 295 (2019) 41–48

Contents lists available at ScienceDirect

Journal of Biotechnology

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

Design of a multiepitopic *Zaire ebolavirus* protein and its expression in plant cells

Ricardo Nieto-Gómez[a,b], Carlos Angulo[c], Elizabeth Monreal-Escalante[a,b], Dania O. Govea-Alonso[a,b], Annie S. De Groot[d], Sergio Rosales-Mendoza[a,b,*]

[a] *Laboratorio de Biofarmacéuticos recombinantes, Facultad de Ciencias Químicas, Universidad Autónoma de San Luis Potosí, Av. Dr. Manuel Nava 6, SLP, 78210, Mexico*
[b] *Sección de Biotecnología, Centro de Investigación en Ciencias de la Salud y Biomedicina, Universidad Autónoma de San Luis Potosí, Av. Sierra Leona 550, Lomas 2ª Sección, San Luis Potosí, 78210, Mexico*
[c] *Grupo de Inmunología & Vacunología, Centro de Investigaciones Biológicas del Noroeste, SC., Instituto Politécnico Nacional 195, Playa Palo de Santa Rita Sur, La Paz, B. C.S., C.P. 23096, Mexico*
[d] *EpiVax Inc, Providence, RI, USA*

ARTICLE INFO

Keywords:
Low-cost vaccine
Plant-based vaccine
GP antigen
Epitope-based vaccine
EpiMatrix
VaxCAD

ABSTRACT

The recent Ebola virus disease (EVD) outbreaks make the development of efficacious and low cost vaccines against *Ebola virus* (EBOV) an urgent goal. Multiepitopic vaccines allow a rational design rendering vaccines able to induce proper immune responses in terms of polarization and potency. In addition, the pathogen variants can be easily covered by including epitopes conserved among relevant isolates. Other important aspects to consider in vaccination are the costs associated to production, distribution, and administration of the vaccine. Plants provide an advantageous platform for this purpose, since they yield biomass at very low costs and some species can be used to formulate purification-free oral vaccines. In the present study, a multiepitopic protein called Zerola, which carries epitopes from the EBOV glycoprotein (GP), was designed based on immunoinformatic approaches and current experimental evidence on B cell protective GP epitopes. Moreover, expression studies performed in nuclear-transformed tobacco lines confirmed the capacity of the plant cell to synthetize the Zerola antigenic protein. The generation of this plant-based candidate vaccine is a step forward in the development of highly efficient and low cost EBOV vaccines.

1. Introduction

EBOV is part of the *Filoviridae* family whose members are commonly referred to as "filoviruses", which are known to cause hemorrhagic fever diseases in humans and non-human primates (NHPs) (Jun et al., 2015). After an incubation period of 4–10 days, the afflicted suffer from non-specific fever symptoms (nausea, weakness, fatigue, muscular pain, *etc.*) which lead to hemorrhagic manifestations such as vomit, blood vessel ruptures, and gastrointestinal bleeding (Lado et al., 2015). Patients die in shock days after showing clinical symptoms. EBOV is a negative sense single stranded RNA (ssRNA) filovirus comprised of genes that encode seven core structural proteins: (i) four helical nucleocapsid proteins (NP, VP35, VP30, L), (ii) two membrane associated proteins (VP24, VP40), (iii) and the transmembrane glycoprotein (GP) that mediates the viral entry process into host cells (Beer et al., 1999). The Zaire genus comprises the following species: 1) *Sudan ebolavirus* (SEBOV), (2) *Zaire ebolavirus* (ZEBOV), (3) *Côte d'Ivoire ebolavirus* (also known as *Tai Forest ebolavirus*) (TEBOV), (4) *Reston ebolavirus* (REBOV), and (5) *Bundibugyo ebolavirus* (BEBOV); with *Reston ebolavirus* being the only one that is not pathogenic for humans (Marsh et al., 2011; Kuhn et al., 2010). The 2014–2016 EBOV outbreak caused 28,616 reported infections and 11,310 deaths within the African countries of Guinea, Liberia and Sierra Leone (http://www.who.int/csr/disease/ebola/en/), and became the most lethal outbreak since the discovery of EBOV in 1976.

Today, the most effective way to prevent EBOV outbreaks from spreading consists in the isolation of patients with symptoms of Ebola virus disease (EVD) within biocontainment facilities; nevertheless, this strategy is difficult to implement (Kekulé, 2015). Much remains unknown regarding the biological origin of the virus; however, most studies seem to agree that the most common animal sources of human infection include a variety of fruit bats (Gatherer, 2014; Mandl et al., 2015) and possibly, non-human primates (NHPs) (Gale et al., 2016). Vaccination is the most attractive approach to prevent EVD (Wang

* Corresponding author at: Facultad de Ciencias Químicas, Universidad Autónoma de San Luis Potosí, Av. Dr. Manuel Nava 6, SLP, 78210, Mexico.
E-mail address: rosales.s@fcq.uaslp.mx (S. Rosales-Mendoza).

https://doi.org/10.1016/j.jbiotec.2019.02.003
Received 16 December 2018; Received in revised form 12 February 2019; Accepted 13 February 2019
Available online 28 February 2019
0168-1656/

et al., 2017). The first vaccine of high efficacy was reported by Henao-Restrepo, who performed in 2015 a ring vaccination cluster trial in Guinea to evaluate a candidate vaccine based on a recombinant vesicular stomatitis virus expressing the EBOV-GP protein (rVSV-ZEBOV) (Henao-Restrepo et al., 2017). In fact, this vaccine probably played a key role in the control of the 2014 outbreak. However, this type of vaccine is costly and parenteral administration limits its massive application in developing countries since trained personnel is required. Another limitation of such vaccine is related to the adverse effects observed in the preclinical evaluation, which deserves further safety evaluations in the clinical evaluation phases (Geisbert, 2017). The oral route is ideal for vaccine administration since formulations do not require trained personnel and sterile devices for administration; and cold chain-free formulations in the form of capsules could be developed, dramatically reducing vaccination costs. Another advantage is the immune response induced in mucosal compartments upon oral immunization; in fact, the VSVDeltaG/ZEBOVGP oral vaccine induced strong immune responses against GP in NHPs, which is critical for blocking EBOV transmission (Qiu et al., 2009). In this respect, GP is an attractive target for subunit vaccine development due to its key role on the viral entry into the host cells; therefore, humoral responses against GP provide protection against EBOV infection (Lee and Saphire, 2009).

Plants constitute a well-established platform for the production of biopharmaceuticals offering low cost, easy scalability, the synthesis of complex proteins, and the possibility of producing oral formulations that do not require purification or cold chain (Govea-Alonso et al., 2014). These features offer the possibility of counting with economical EBOV vaccines that can be safely administered. Thus far, various plant-made vaccines have reached clinical trials, in which antigens expressed in different plant species were orally administered as raw plant material (Tacket, 2009; Takeyama et al., 2015). In the case of EBOV, plant-made vaccines have been narrowly explored, and only an immunotherapy relying on passive immunity consisting in the plant-made antibody cocktail “ZMapp” has proven to be efficacious against EBOV (Qiu et al., 2014; Zhang et al., 2014).

The selection of appropriate antigens and epitopes are critical steps in the development of effective plant-made vaccines. On this regard, immunobioinformatics allows the development of optimized vaccines in a rational basis, saving time and resources (Moise et al., 2015). For instance, multi-epitopic vaccines can be designed to contain a set of the most antigenic and immunoprotective conserved epitopes to trigger broad and long lasting protective B- and/or T- cell responses against EBOV (Li et al., 2005; Yasmin and Nabi, 2016). One approach for epitope selection relies on targeting T and B cell linear epitopes that bind to a wide set of human leukocyte antigen (HLA) I and II molecules respectively (Bounds et al., 2017; Dikhit et al., 2015; Dash et al., 2017). Within this context, the aim of the present study was to design a multiepitopic chimeric protein carrying EBOV-GP epitopes, elucidated through the use of reported protective EBOV epitopes, conventional bioinformatics tools available online and specialized immunoinformatics tools. Evidence on the expression of the designed multiepitopic protein in plant cells is also presented.

2. Materials and methods

2.1. In silico identification of immunogenic EBOV-GP epitopes

2.1.1. Bibliographical and experimental curated epitopes search

A thorough online search for bibliographical sources that contained EBOV-GP epitope data was performed and resulted in the following findings: (i) Becquart et al. (2014) conducted a study in which B-cell epitopes were identified based on reactivity of the sera of gabonese EBOV infected patients against GP and other proteins of the virus [Nucleoprotein (NP), and matrix Viral Proteins (VP40 and VP35)] across three different outbreaks (1996–2001 period). (ii) The Immune Epitope Database (IEDB, www.iedb.com) integrates the results of multiple experimental analyses on relevant B-cell EBOV-related epitope data (Vita et al., 2014; Ponomarenko et al., 2014). Linear EBOV-GP epitopes were selected from these two sources to aid with the design of the multiepitopic EBOV vaccine (Fig. 1).

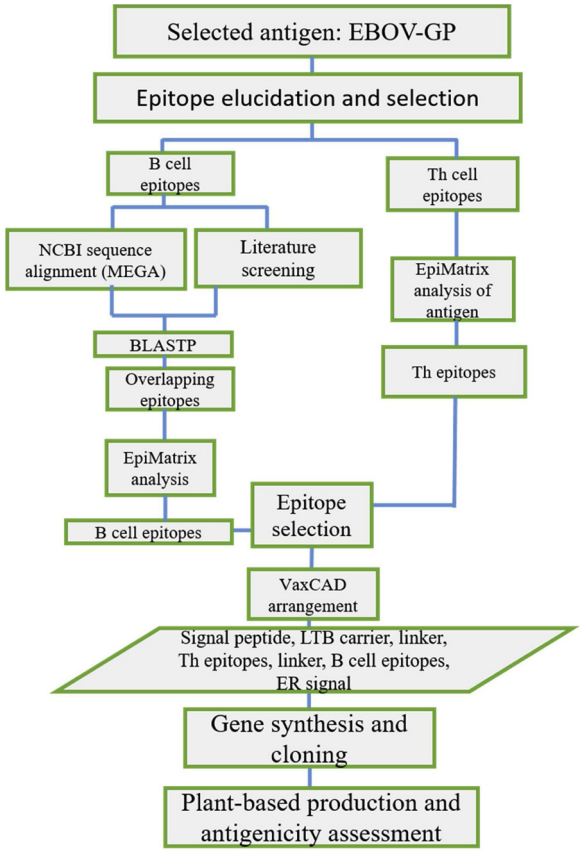

Fig. 1. Flowchart on the steps followed for the design of the Zerola multiepitopic vaccine.

2.1.2. Conventional in silico analysis for the prediction of GP-EBOV candidate epitopes

Protein sequences of Zaire EBOV-GP were downloaded from the National Center for Biotechnology information database (NCBI, https://www.ncbi.nlm.nih.gov/genome/viruses/variation/ebola/), corresponding to infected human and NHPs from Africa for the 1976–2014 period and comprised multiple EVD outbreaks. These sequences were aligned using the MEGA v 6.0 software (Tamura et al., 2013) to identify consensus regions, which were subsequently compared with the B-cell epitopes identified by Becquart et al. (2014) and the IEDB database (Ponomarenko et al., 2014) through the Basic Local Alignment Search Tool for Proteins (BLASTP) (Altschul et al., 1990). Overlapping sequences were selected as putative protective B-cell epitopes.

2.1.3. Epitope prediction and rational epitope arrangement by the Epimatrix and VaxCAD tools

The GP epitopes selected from IEDB and bibliographical sources, as well as the overlapping peptide regions obtained through the conventional *in silico* analysis (MEGA and BLASTP), were subjected to EpiVax’s tools, which included the computational matrix for epitope prediction named Epimatrix as well as VaxCAD. EpiMatrix predicts the binding of the epitopes to the human MHC proteins, which is critical for the activation of T- and B-cells. EpiMatrix maps were obtained by evaluating the binding potential of each of the amino acids contained within the epitopes to the HLA haplotypes, and assigning a score to each of these amino acids that serves as an indicator of the possibility of binding to a

determined HLA haplotype. The software assesses the binding potential to the most common HLA molecules found in certain "supertypes" (DR3, DR11, DR8, DR13, DR1, DR7, DR51, DR15, and DR4). This allows a coverage of more than 90% of the world population without the need of trying each haplotype individually (Moise et al., 2013; De Groot et al., 1997). On the other hand, the VaxCAD algorithm was applied to find an optimal separate arrangement of both Th-cell and B-cell epitopes, which facilitates antigen processing and presentation, and avoid the generation of undesired junctional epitopes (Liu et al., 2013).

2.2. Gene design and plant transformation

A synthetic gene coding for the designed multiepitopic antigen, called Zerola, was obtained by reverse genetics. The Zerola antigen comprises the following elements: a signal peptide for endoplasmic reticulum (ER) translocation (from the *Glycine max* vegetative storage protein), the mature sequence of the B subunit of the heat labile enterotoxin (LTB), a glycine-proline linker, the set of epitopes derived from EpiMatrix analysis, a second glycine-proline linker, the set of EBOV epitopes selected through conventional immunoinformatics and the SEKDEL ER retention sequence. The arrangement of the EBOV epitopes was defined based on VaxCAD analysis. The *Zerola* gene was codon-optimized according to codon usage in plants and obtained from GenScript Co (NJ, USA); and it was subcloned into the pBI121 binary vector through the *Sma*I and *Sac*I restriction sites using conventional molecular cloning procedures (Jefferson, 1987). The obtained construct, named pBI-Zerola, was confirmed by restriction profiles and conventional sequencing; and subsequently transferred to *Agrobacterium tumefaciens* (strain GV3101) using a previously described protocol (Cangelosi et al., 1991).

2.3. Plant transformation and development

Transgenic tobacco (*Nicotiana tabacum* cv. Petit Havana SR1) plants were generated by *A. tumefaciens*-mediated transformation according to Horsch et al. (1985) using the agrobacteria carrying the pBI-Zerola vector. Tobacco cultures were maintained in a controlled environment chamber (25°C) under a 16 h light/8 h dark photoperiod and transferred to fresh medium biweekly. Explants were cultivated on selection medium until shoot development and regeneration was completed in hormone-free medium. Rooted plantlets were subsequently acclimatized in soil and grown in a greenhouse. To assess the transgene presence in the putative transformed plants, total DNA was isolated from leaves of both putative transformants and wild-type plants according to Dellaporta et al. (1983). To perform PCR, 25 µL PCR reaction mixtures were prepared containing 100 ng of DNA, 1X PCR buffer, 1.5 mM magnesium chloride, 2.5 U Taq DNA polymerase (Vivantis), 1 mM dNTPs, and 1 µM of each forward and reverse primers designed to amplify the entire *Zerola* gene (forward, 5′ GACTACACAGGCTTTGTC AACTA 3′; reverse, 5′ GGCTCTCTAAGTGGATGTGAAG 3′); yielding a 1382 bp amplicon. Temperature cycling conditions were 94°C for 5 min (initial denaturation), 35 cycles at 94°C for 30 s (denaturation), 55°C for 30 s (annealing), 72°C for 90 s min (elongation), and a final extension at 72°C for 5 min. This procedure was performed in a MultiGeneTM Mini Personal Thermal Cycler (Labnet). PCR products were analyzed by electrophoresis in 1% agarose gels.

2.4. Protein analysis

Total soluble protein extracts were obtained by milling fresh leaf tissues or seeds from transformed and WT plants (˜250 mg) in the presence of 500 µL of cold protein extraction buffer composed of 750 mM Tris − HCl, pH = 8.0, 15% sucrose, 100 mM β–mercaptoethanol, and 1 mM PMSF.

For in-gel western blot analysis, protein samples from leaf tissues were mixed with one volume of 2× Laemmli reducing protein loading buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue and 0.125 M Tris HCl, pH 6.8) and denatured by boiling for 10 min at 95°C, followed by clarification (centrifugation at 12,000 rpm for 10 min). SDS-PAGE analysis was performed using 4–12% acrylamide gels under denaturing conditions (0.1% SDS). Following electrophoresis, the gel was incubated in a solution containing 50% isopropanol and 5% acetic acid for 15 min. The gel was subsequently washed in ultrapure water for 15 min with an additional incubation in the blocking buffer (PBS plus 0.1% Tween® 20, and 5% BSA) at room temperature for 2 h. The gel was labeled overnight at 4°C with the primary antibody (anti-GP1 serum, 1:500 dilution, plus 5% BSA). A subsequent labeling was performed with a goat CF™488-conjugated anti-mouse IgG antibody (1:2000 dilution; Biotium, Inc.) for 2 h at room temperature. After washing the gel three times for 10 min in PBS plus 0.1% Tween 20, it was scanned on an Odyssey CLx equipment using the Image Studio™ Software.

For dot blot analysis, 5 µl of leaf or seeds extracts were applied onto nitrocellulose membrane and let to dry. Blocking was performed with 5% non-fat milk for 1 h and labeling was then performed overnight using anti-GP1 serum (1:500 dilution, plus 5% BSA). Signal was developed as reported by Ríos-Huerta et al. (2017).

Zerola content in the soluble protein fraction was determined by a GM1 ganglioside-dependent ELISA assay using extracts obtained as described above, but lacking β–mercaptoethanol (Chikwamba et al., 2002). Plates were coated overnight at 4°C with Type III GM1 ganglioside (1.5 µg per well, Sigma). After washing with PBST, plates were blocked for 2 h at room temperature with 5% fat-free milk dissolved in PBST. Protein extracts were added, and plates were incubated overnight at 4°C. Primary labeling was performed overnight at 4°C with a mouse anti-LTB anti-serum (1:400 dilution). Horseradish peroxidase-conjugated secondary anti-mouse antibody (1:2000 dilution, Sigma) was applied for 2 h at room temperature. Reaction was developed by adding an ABTS-based substrate (Sigma) and optical density values were measured at OD 405 nm using a Multiskan FC microplate reader (Thermo Electron Corporation, Waltham, MA, USA). A standard curve prepared with pure LTB standards (0.2–2 ng/µl) was included in the assay.

3. Results

3.1. Epitopes identified by bibliographical search, experimental curated epitope database and conventional in silico analysis

The following two peptides were identified by Becquart et al. (2014) as specific immunodominant GP epitopes through reactivity assays using the sera of symptomatic Gabonese survivors: NISGQSPARTSSDPG and DISEATQVEQHHRRT. From the IEDB online compendium, the following three linear functional EBOV-GP B-cell epitopes were selected: AGNNNTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTR, VEQHHRRTD, and VYKLDISEA. Epitopes retrieved from these resources were further aligned to consensus regions from the GP antigen of the EBOV outbreaks.

On January 27 of 2015, a total of 143 protein sequences were downloaded from the NCBI ebolavirus database which are from infected African human and non-human primates (NHP) during the 1976–2014 period. Retrieved sequences were subjected to a multiple sequence alignment using MEGA software package (v 6.0). The CLUSTALW algorithm that uses Gonnet matrix, along with a 100 bootstrap value and the rest of the parameters left at default settings, were used to perform the alignment. As a result, 25 consensus regions of GP were obtained and compared through BLASTP tool with the B-cell epitopes extracted from the bibliographical and IEDB sources. It was found that certain segments from seven of the consensus regions had a 100% coincidence value with at least one of the epitopes from the mentioned sources (Table 1). Therefore, these segments were used for further analysis.

Table 1
Pairings of epitopes obtained from bibliographical sources with consensus regions obtained from MEGA analysis. Table shows epitopes from Becquart et al. (2014) and IEDB (Ponomarenko et al., 2014) that had a BLASTP coincidence value of 100% with the consensus regions obtained from the analysis of EBOV-GP protein sequences.

Consensus region	Aminoacid position within EBOV-GP	Coinciding IEDB (Ponomaenko et al. 2014) epitopes)	Coinciding Becquart et al. (2014) epitopes
GQSPARTSSDP	320-330	–	NISGQSPARTSSDPG
NTPVYKLD	390-397	VYKLDISEA	–
SEATQ	399-403	–	DISEATQVEQHHRRT
QHHRR	406-410	VEQHHRRTD	DISEATQVEQHHRRT
TAGNNNTHHQDTGEES	458-473	AGNNNTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTR	–
SGKLGLI	476-482	AGNNNTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTR	–
NTIAGVAGLITGGRR	484-498	AGNNNTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTR	–

*(-) without coincidence between GP consensus region and epitopes.

In order to select epitope regions within the GP antigen in which the consensus regions overlapped with the epitopes extracted from the Becquart's study and IEDB database (Ponomarenko et al., 2014), a representative sequence of EBOV-GP (GenBank: AIE11809) was used. The following overlapping regions were selected as B-cell GP epitope candidates for the design of the multiepitopic vaccine: NISGQSPARTSSDPE, NTPVYKLDISEATQVGQHHRRAD, and TAGNNNTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTR. A criterion for the selection of antigenic regions was for these to be at least 15 amino acids in length (15, 23, and 43 aa, respectively).

3.2. Epitopes identified by EpiMatrix analysis

The GP epitopes selected from bibliographical and IEDB sources as well as the overlapping peptide regions obtained through our *in silico* analysis were subjected to an EpiMatrix analysis to improve epitope selection (Fig. 2, highlighted in green). The results retrieved by EpiMatrix revealed a number of positive "hits" (recognizable nonamers by HLA) for class I and II HLA, indicating the possible generation of a T cell dependent response (Table 2).

In parallel, an additional set of epitopes with promising "hits" assessed by the EpiMatrix system (database and prediction tool) was selected to be included in the multiepitopic vaccine (Table 3). The order

Fig. 2. Sequence of the multiepitopic Zerola antigen. The Zerola protein comprises (i) a set of GP-EBOV Th-cell epitopes retrieved by EpiMatrix (in red) indicated as follows: epitope Th-1 (underlined), epitope Th-2 (no underlined), epitope Th-3 (underlined), epitope Th-4 (no underlined), epitope Th-5 (underlined), epitope Th-6 (no underlined), epitope Th-7 (underlined), epitope Th-8 (no underlined), epitope Th-9 (underlined), epitope Th-10 (no underlined), epitope Th-11 (underlined), epitope Th-12 (no underlined), epitope Th-13 (underlined). (ii) a set of B cell epitope regions retrieved by conventional immunoinformatics analysis (in green) indicated as follows: epitope B1 (underlined), epitope B2 (no underlined), epitope B3 (underlined). Regions in blue correspond to the glycine-proline linkers included to aid in epitope display. The dark green section corresponds to the plant endoplasmic reticulum retention signal. The yellow section corresponds to a signal peptide for translocation into endoplasmic reticulum and the grey section corresponds to the LTB sequence. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

in which the selected epitopes were placed in the context of the chimeric protein was determined by the VaxCAD algorithm, which allowed the rearrangement of the peptides in a manner that facilitates protein folding, epitope processing and presentation as well as avoiding the generation of junctional epitopes (Table 4, Fig. 3).

3.3. Approach for the production of the Zerola synthetic gene in plants

A chimeric protein, called Zerola, was designed based on the set of epitopes identified by a conventional immunoinformatic approach and those identified by EpiMatrix along with LTB as a mucosal immunogenic carrier. An expression approach targeting the accumulation of Zerola into the ER of the plant cell was implemented by including a signal peptide and an ER retention signal. The synthetic *Zerola* gene was successfully cloned into the pBI121 binary vector according to restriction profile analysis and sequencing, yielding the pBI-Zerola vector, which allows a constitutive expression in the plant cell driven by the 35SCaMV promoter. The vector comprises the *nptII* marker gene that confers kanamycin resistance under the control of NOS promoter. The pBI-Zerola construct was transferred to *A. tumefaciens* as revealed by PCR analysis of a kanamycin-resistant clone following electroporation.

3.4. Tobacco plants express the Zerola multiepitopic antigen

Following the infection of tobacco leaf segments with the recombinant *Agrobacterium* strain carrying the pBI-Zerola vector, ten putative transformed lines were regenerated in selective media and successfully transferred to soil and grown in a greenhouse. No obvious phenotypic alterations were observed in the putative transgenic plants when compared with the WT line (Fig. 4a). A PCR analysis of the candidate plant lines was conducted to assess the presence of the transgene, observing the presence of the expected amplicon of ≈1382 bp in length in seven of the tested lines (Fig. 4b), which is consistent with the expected size and the amplicon yielded by the positive control consisting of the pBI-Zerola vector as template. No signal was observed for the DNA sample from the WT line used as negative control. The presence of the plant-made Zerola protein was proved through an in-gel Western blot analysis using an anti-GP1 hyperimmune serum, observing a ≈47 kDa immunoreactive band in all tested lines, matching the theoretical molecular weight for Zerola, although a band of higher molecular weight (≈70 kDa) was also observed for the transgenic plants (Fig. 5a). The pure GP1 protein used as positive control in this assay yielded a 60 kDa band as expected, whereas no bands were observed in the negative control (protein extract from a WT plant). We further performed a dot blot analysis labeling with an anti-GP1 serum to assess the antigenicity of Zerola under native conditions, observing a positive reactivity in leaf and seeds extracts from the transformed line P10 whereas null signal was observed for an extract from the WT line. A GM1-ELISA was run to determine the levels of Zerola in leaf tissues, in which significant signal was recorded for transgenic lines and according to the LTB standard curve, accumulation levels of Zerola are in the

Table 2
Results of EpiMatrix analysis of the B-cell epitopes selected based on previous studies looking for GP regions with reactivity against sera from infected patients. EpiMatrix matches sequences of nine amino acids (9-mers) to HLA motifs to find "hits". In order to compare potential epitopes across multiple HLA alleles, EpiMatrix raw scores are converted to a normalized 'Z' scale. Peptides scoring above 1.64 on the EpiMatrix 'Z' scale (typically the top 5% of any given sample) are likely to be MHC ligands.

Peptide	Source	Length (aa)	EpiMatrix Class II "Hits"	EpiMatrix Class I "Hits"
NISGQSPARTSSDPG	Becquart et al. (2014)	15	2	1
NISGQSPARTSSDPE	In silico analysis in the present study	15	0	0
DISEATQVEQHHRRT	Becquart et al. (2014)	23	8	0
NTPVYKLDISEATQVGQHHRRAD	In silico analysis in the present study	15	2	1
VYKLDISEA	IEDB (Ponomarenko et al., 2014)	9	0	0
VEQHHRRTD	IEDB (Ponomarenko et al., 2014)	9	0	0
AGNNNTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTR	Ponomarenko et al. (2014)	42	9	5
TAGNNNTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTR	In silico analysis in the present study	43	9	5

range of 1-2 μg g^{-1} fresh weight leaf tissues (Fig. 5B).

4. Discussion

The 2014 EBOV outbreak, the most lethal to date, led to a substantial increase in the number of vaccine candidates against the virus. Eight of these vaccines reached clinical trials during de 2014–2016 period, and one of them, based in the GP-expressing rVSV-ZEBOV vector, achieved 100% efficacy against the disease (Snape, 2017). However, the expansion of the vaccination alternatives is desirable in view of the limitations identified in the current vaccine candidates. For instance, some of the concerns of the current candidates comprise the lack of data on long-term immunity, the risks on viral vector integration and parenteral administration, as well as poor vaccination acceptance (especially from younger patients) (Mingozzi and High, 2013). New vaccine candidates might also serve as boosting agents in combination with the current candidates.

In the present study a synthetic protein named Zerola was designed, comprising critical conserved B-cell EBOV-GP epitopes from EBOV strains dating back to 1976 (and likely to remain conserved), and a set of Th-cell epitopes identified from an analysis of the full length GP through the EpiMatrix system, having as criterion the lack of human proteome cross-conservation. IEDB served as a first approach to identify curated B-cell epitopes with proved *in vivo* protection against ZEVOB, including epitopes VYKLDISEA and VEQHHRRTD (Ponomarenko et al., 2014), which are the only ones that do not overlapp with MHC epitopes predicted by EpiMatrix. The inclusion of the epitopes obtained through EpiMatrix is intended to increase the vaccine efficacy by increasing T cell populations, supporting the induction and maintenance of robust humoral responses against EBOV-GP. EpiMatrix and VaxCAD are immunoinformatics toolsets that allow the identification and prediction of Th-cell epitopes relevant for vaccine design and their rearrangement for the minimization of epitope junction immunogenicity respectively. The use of these tools has been validated in multiple published studies for Th-cell epitope vaccines in which immunogenicity at the clinical level was demostrated (Koren et al., 2007; Liu et al., 2013; Elfaki et al., 2012); peer reviews have also documented the accuracy of the EpiMatrix algorithm from *in vitro* to *in vivo* models (De Groot and Martin, 2009; Koren et al., 2007).

Herein, the Zerola candidate antigen was expressed in tobacco plants. A binary vector, based on a constitutive promoter was used to transfer the *Zerola* gene into plant cells; so these could serve as a low cost biofactory of the rationally designed antigen while avoiding costly upstream processing required under conventional systems, such as microbial systems and those based on mammalian cell culture. Plant-

Table 4
Optimal order for Th epitope candidates obtained through EpiMatrix for the construction of multiepitopic vaccines.

Order	GP Protein Position	Epitope
1	52-74	VCRDKLSSTNQLRSVGLNLEGNG
2	579 - 595	LRTFSILNRKAIDFLLQ
3	287 – 308	EWAFWETKKNLTRKIRSEELSF
4	336 – 352	TEDHKIMASENSSAMVQ
5	85-99	RWGFRSGVPPKVVNY
6	156-174	EGAFFLYDRLASTVIYRGT
7	555 – 569	ICGLRQLANETTQAL
8	215 – 229	STTIRYQATGFGTNE
9	190-204	KKDFFSSHPLREPVN
10	566 – 583	TQALQLFLRATTELRTFS
11	245 – 265	ESRFTPQFLLQLNETIYTSGK
12	32-49	SIPLGVIHNSTLQVSDVD
13	238 – 252	NLTYVQLESRFTPQF

Table 3
Th cell Epitope candidates obtained through an EpiMatrix analysis of the full GP sequence. The EpiMatrix cluster immunogenicity score represents the deviation in putative epitope content from baseline expectation based on a random peptide standard. Th-cell epitope clusters scoring above +10 are considered to be potentially immunogenic. Peptides with elevated human proteome cross-conservation (JanusMatrix Human Homology Scores > 2) were removed.

GP Protein Position	AA Sequence	EpiMatrix Cluster Score	EpiMatrix Hits	JanusMatrix Human Homology Score
32-49	SIPLGVIHNSTLQVSDVD	18.95	12	0.54
52-74	VCRDKLSSTNQLRSVGLNLEGNG	11.76	10	1.08
85-99	RWGFRSGVPPKVVNY	15.38	8	0
156-174	EGAFFLYDRLASTVIYRGT	17.43	11	1.33
190-204	KKDFFSSHPLREPVN	11.22	6	0
215 – 229	STTIRYQATGFGTNE	10.05	6	0.17
238 – 252	NLTYVQLESRFTPQF	13.8	6	0.71
245 – 265	ESRFTPQFLLQLNETIYTSGK	28.17	16	2
287 – 308	EWAFWETKKNLTRKIRSEELSF	14.44	11	0.5
336 – 352	TEDHKIMASENSSAMVQ	15.84	10	0.42
555 – 569	ICGLRQLANETTQAL	10.5	6	0.14
566 – 583	TQALQLFLRATTELRTFS	22.55	12	1.47
579 - 595	LRTFSILNRKAIDFLLQ	20.25	11	1.8

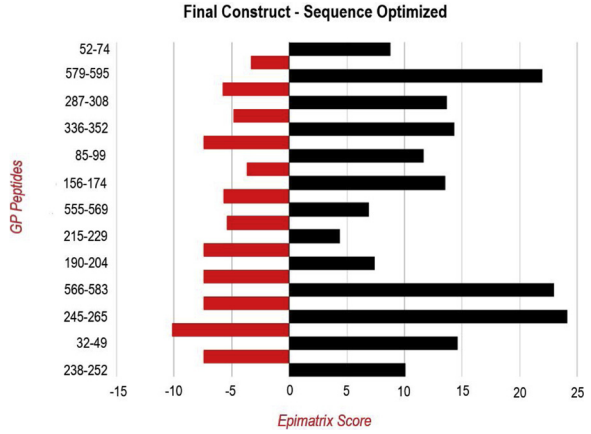

Fig. 3. Epitope order analysis performed by EpiMatrix. The VaxCAD algorithm was used to iteratively rearrange the candidate peptides in order to achieve minimal non-specific "junctional" immunogenicity created at the intersections between epitope sequences (represented by consecutive peptides going on different directions/colors). Peptides were incorporated into the Zerola protein in the order shown here.

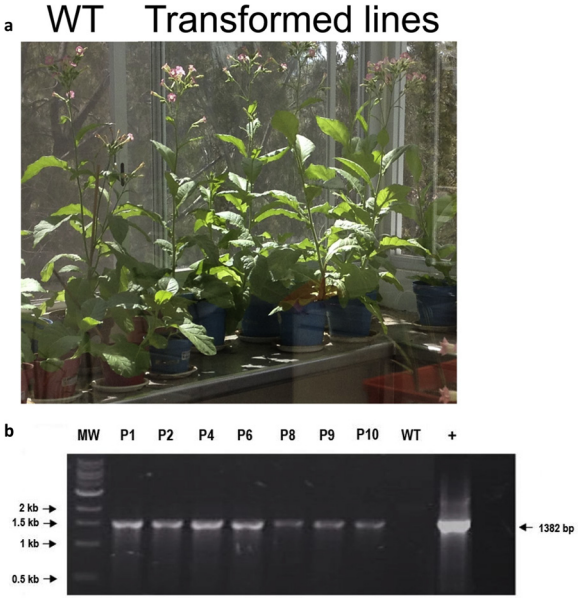

Fig. 4. a Aspect of mature tobacco plants of the WT and transformed lines. Transgenic plants expressing the Zerola antigen showed no phenotypic alterations when compared with the untransformed plant. **b** PCR analysis of putative transgenic tobacco lines carrying the *Zerola* gene. Total DNA preparations were analyzed by end point PCR using specific primers for the *Zerola* gene. Lanes: lane 1, 1 kb molecular weight marker; Lanes 2–8, DNA from the chosen P1-P10 lines, respectively; lane 9, DNA from WT plant; lane 10, positive control (pBI-Zerola, 20 ng).

made production also avoids the risks involved in the production of attenuated vaccines and those based on viral vectors. Expression of Zerola in tobacco leaves reached levels of up to 2.2 μg g^{-1} fresh weight leaf tissues and the in-gel western blot confirmed the presence of the integral Zerola protein according to its theoretical molecular weight (47 kDa). The ≈70 kDa protein that also showed immunoreactivity could correspond to Zerola associated with proteins of the host. The positive results observed in both western blot and dot blot analyses performed with anti-GP1 serum confirmed the antigenicity of the plant-made Zerola antigen and suggest that it properly displays the GP epitopes. The ELISA assay conducted to quantify the recombinant protein was based on GM1 binding, a property of pentameric LTB, thus the

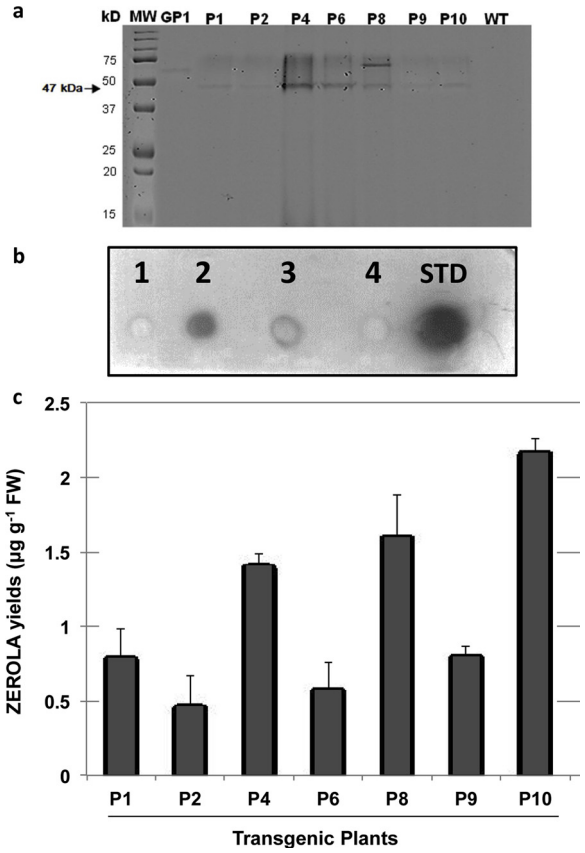

Fig. 5. Integrity and levels of the tobacco-made Zerola protein. **a** In-gel Western blot analysis performed with a mouse anti-GP1 serum to assess the integrity of the Zerola antigen in total protein plant extracts. Lanes: MW, molecular weight marker; GP1, pure *E. coli*-made GP1 protein (250 ng); lanes P1-P10, protein extracts from transgenic tobacco lines; lane WT, protein extract from a wild-type tobacco line. **b** Dot-blot analysis to detect the Zerola native antigen using anti-GP1 serum. Total protein extracts from tobacco leaves or seeds were applied onto nitrocellulose membrane and probed with anti-GP1 serum to assess the antigenicity of Zerola protein. Dots: 1, WT leaf protein extract; 2, P10 leaf protein extract; 3, WT seed protein extract; 4, P10 seed protein extract; STD, pure GP1 protein purified from *E. coli* (500 ng). **c** Levels of the Zerola protein in leaf tobacco tissues estimated by GM1-ELISA, using an anti-LTB serum.

positive reactivity observed in the assay revealed that Zerola assembles in pentamers, which is critical for the functionality of the antigen since the pentameric form mediates an efficient transport across the intestinal barrier allowing the induction of immune responses.

Several studies reporting the expression of LTB-based antigens rely on the use of a strong constitutive promoter fused to an ER retention signal, which results in high accumulation levels of the recombinant protein within the plant biomass (Haq et al., 1995; Mason et al., 1998). However, the toxicity of LTB is also a factor worth monitoring in plants since some groups have reported altered phenotypes in potato and tobacco plants expressing LTB alone or LTB-based chimeric proteins (Mason et al., 1998; Romero-Maldonado et al., 2016). Herein, yields in the range of previous reports were found for the Zerola antigen with no phenotypic alterations, which allow us to project an adequate productivity under greenhouse conditions. Therefore, Zerola is proposed as a plant-based Ebola oral vaccine prototype able to induce wide and strong humoral responses against key components of the virus related to viral entry. Immunization assays need to be performed in order to compare the potential of the Zerola plant-made vaccine following oral immunization schemes. An oral plant-based vaccine was previously reported, which is based on a LTB-based antigen carrying two GP1 epitopes (VEQHHRRTD and VYKLDISEA) and accumulated in

transgenic tobacco tissues at levels up to 14.7 μg g^{-1} fresh leaves (Ríos-Huerta et al., 2017); the antigen induced humoral responses against the target GP epitopes, indicating that the LTB succeeded as a mucosal carrier that favors the induction of adaptive immune responses upon oral administration (Lazorova´ et al., 1993). Remarkably, the monoclonal antibody cocktail "ZMapp" developed in tobacco plants is efficacious in NHPs and human beings (Qiu et al., 2014; Zhang et al., 2014). Therefore, our study strengthens the idea that plants can be effectively used to produce immuno-prophylactic and therapeutic treatments against EVD. The Zerola antigen constitutes a new attractive vaccine candidate that could lead to the induction of optimal humoral responses against EBOV, since the design includes wide sets of B- and Th- cell epitopes that offer the potential to increase the potency and duration of the elicited response.

In conclusion, the plant-made Zerola antigen offers a new tool for the development of rationally-designed vaccines intended to induce highly protective and long-lasting humoral responses, providing a production platform of easy scalability, low production cost and the possibility of oral administration that favors patients comfort, compliance, and safe vaccine delivery.

Conflicts of interest

There are no conflicts of interest between the authors.

Acknowledgements

Current investigations from the group are supported by CONACYT/México (grantINFR-2016-271182 and CB-256063 to SRM and PDCPN2014-01/248033 to CA).

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Becquart, P., Mahlakõiv, T., Nkoghe, D., Leroy, E.M., 2014. Identification of continuous human B-cell epitopes in the VP35, VP40, nucleoprotein and glycoprotein of Ebola virus. PLoS One 9, e96360. https://doi.org/10.1371/journal.pone.0096360.

Beer, B., Kurth, R., Bukreyev, A., 1999. Characteristics of filoviridae: marburg and ebola viruses. Naturwissenschaften 86, 8–17. https://doi.org/10.1007/s001140050562.

Bounds, C.E., Terry, F.E., Moise, L., Hannaman, D., Martin, W.D., De Groot, A.S., et al., 2017. An immunoinformatics-derived DNA vaccine encoding human Class II T cell epitopes of Ebola virus, Sudan virus, and Venezuelan equine encephalitis virus is immunogenic in HLA transgenic mice. Hum. Vaccin. Immunother. 2, 1–13. https://doi.org/10.1080/21645515.2017.1329788.

Cangelosi, G.A., Best, E.A., Martinetti, G., Nester, E.W., 1991. Genetic analysis of Agrobacterium. Methods Enzymol 204, 384–397.

Chikwamba, R., Cunnick, J., Hathaway, D., McMurray, J., Mason, H., Wang, K., 2002. A functional antigen in a practical crop: LT-B producing maize protects mice against Escherichia coli heat labile enterotoxin (LT) and cholera toxin (CT). Transgenic Res. 11, 479–493.

Dash, R., Das, R., Junaid, M., Akash, M.F., Islam, A., Hosen, S.Z., 2017. In silico-based vaccine design against Ebola virus glycoprotein. Adv. Appl. Bioinform. Chem. 10, 11–28. https://doi.org/10.2147/AABC.S115859.

De Groot, A.S., Martin, W., 2009. Reducing risk, improving outcomes: bioengineering less immunogenic protein therapeutics. Clin. Immunol. 131, 189–201. https://doi.org/10.1016/j.clim.2009.01.009.

De Groot, A.S., Jesdale, B.M., Szu, E., Schafer, J.R., Chicz, R.M., Deocampo, G., 1997. An interactive Web site providing major histocompatibility ligand predictions: application to HIV research. AIDS Res. Hum. Retroviruses 13, 529–531. https://doi.org/10.1089/aid.1997.13.529.

Dellaporta, S.L., Wood, J., Hicks, J.B., 1983. A plant DNA minipreparation: Version II. Plant Mol. Biol. Rep. 1, 19–21.

Dikhit, M.R., Kumar, S., Vijaymahantesh, Sahoo B.R., Mansuri, R., Amit, A., et al., 2015. Computational elucidation of potential antigenic CTL epitopes in Ebola virus. Infect. Genet. Evol. 36, 369–375. https://doi.org/10.1016/j.meegid.2015.10.012.

Elfaki, M.E.E., Khalil, E.A.G., De Groot, A.S., Musa, A.M., Gutierrez, A., Younis, B.M., et al., 2012. Immunogenicity and immune modulatory effects of in silico predicted *L. donovani* candidate peptide vaccines. Hum. Vaccin. Immunother. 8, 1769–1774. https://doi.org/10.4161/hv.21881.

Gale, P., Simons, R.R., Horigan, V., Snary, E.L., Fooks, A.R., Drew, T.W., 2016. The challenge of using experimental infectivity data in risk assessment for Ebola virus: why ecology may be important. J. Appl. Microbiol. 120, 17–28. https://doi.org/10.1111/jam.12973.

Gatherer, D., 2014. The 2014 Ebola virus disease outbreak in West Africa. J. Gen. Virol. 95, 1619–1624. https://doi.org/10.1099/vir.0.067199-0.

Geisbert, T.W., 2017. First Ebola virus vaccine to protect human beings? Lancet 389, 479–480. https://doi.org/10.1016/S0140-6736(16)32618-6.

Govea-Alonso, D.O., Cardineau, G.Q., Rosales-Mendoza, S., 2014. Principles of plant-based vaccines. In: Rosales-Mendoza, S. (Ed.), Genetically Engineered Plants as a Source of Vaccines Against Wide Spread Diseases -An Integrated View. Springer Science+Business Media, New York (NY), pp. 1–14 2014.

Haq, T.A., Mason, H.S., Clements, J.D., Arntzen, C.J., 1995. Oral immunization with a recombinant bacterial antigen produced in transgenic plants. Science 268, 714–716.

Henao-Restrepo, A.M., Camacho, A., Longini, I.M., Watson, C.H., Edmunds, W.J., Egger, M., et al., 2017. Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!). Lancet 389, 505–518. https://doi.org/10.1016/S0140-6736(16)32621-6.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G., Fraley, R.T., 1985. A simple and general method for transferring genes into plants. Science 227, 1229–1231. https://doi.org/10.1126/science.227.4691.1229.

Jefferson, R.A., 1987. Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol. Biol. Rep. 5, 387–405.

Jun, S.R., Leuze, M.R., Nookaew, I., Uberbacher, E.C., Land, M., Zhang, Q., et al., 2015. *Ebolavirus* comparative genomics. FEMS Microbiol. Rev. 39, 764–778. https://doi.org/10.1093/femsre/fuv031.

Kekulé, A.S., 2015. Learning from Ebola virus: how to prevent future epidemics. Viruses 7, 3789–3797. https://doi.org/10.3390/v7072797.

Koren, E., De Groot, A.S., Jawa, V., Beck, K.D., Boone, T., Rivera, T., et al., 2007. Clinical validation of the "in silico" prediction of immunogenicity of a human recombinant therapeutic protein. Clin. Immunol. 24, 26–32. https://doi.org/10.1016/j.clim.2007.03.544.

Kuhn, J.H., Becker, S., Ebihara, H., Geisbert, T.W., Johnson, K.M., Kawaoka, Y., et al., 2010. Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations. Arch. Virol. 155, 2083–2103. https://doi.org/10.1007/s00705-010-0814-x.

Lado, M., Walker, N.F., Baker, P., Haroon, S., Brown, C.S., Youkee, D., et al., 2015. Clinical features of patients isolated for suspected Ebola virus disease at Connaught Hospital, Freetown, Sierra Leone: a retrospective cohort study. Lancet Infect. Dis. 15, 1024–1033. https://doi.org/10.1016/S1473-3099(15)00137-1.

Lazorova´, L., Sjo¨lander, A., Russell-Jones, G.J., Linder, J., Artursson, P., 1993. Intestinal tissue distribution and epithelial transport of the oral immunogen LTB, the B subunit of E. coli heat-labile enterotoxin. J. Drug Target. 1, 331–340. https://doi.org/10.3109/10611869308996092.

Lee, J.E., Saphire, E.O., 2009. *Ebolavirus* glycoprotein structure and mechanism of entry. Future Virol. 4, 621–635. https://doi.org/10.2217/fvl.09.56.

Li, G.F., Wang, Y., Zhang, Z.S., Wang, X.J., Ji, M.J., Zhu, X., et al., 2005. Identification of immunodominant Th1-type T cell epitopes from Schistosoma japonicum 28 kDa glutathione-S-transferase, a vaccine candidate. Acta Biochim. Biophys. Sin. (Shanghai) 37, 751–758.

Liu, R., Desrosiers, J., Martin, B., Moise, L., De Groot, A.S., 2013. Immunogenicity assessment of in silico-selected T cell epitopes for a burkholderia biodefense vaccine (P4282). J. Immunol. 190 (1 Supplement) 54.10.

Mandl, J.N., Ahmed, R., Barreiro, L.B., Daszak, P., Epstein, J.H., Virgin, H.W., et al., 2015. Reservoir host immune responses to emerging zoonotic viruses. Cell 160, 20–35. https://doi.org/10.1016/j.cell.2014.12.003.

Marsh, G.A., Haining, J., Robinson, R., Foord, A., Yamada, M., Barr, J.A., 2011. Ebola Reston virus Infection of pigs: clinical significance and transmission potential. J. Infect. Dis. 204, S804–809. https://doi.org/10.1093/infdis/jir300.

Mason, H.S., Haq, T.A., Clements, J.D., Arntzen, C.J., 1998. Edible vaccine protects mice against Escherichia coli heat-labile enterotoxin (LT): potatoes expressing a synthetic LT-B gene. Vaccine 16, 1336–1343.

Mingozzi, F., High, K.A., 2013. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. Blood 122, 23–36. https://doi.org/10.1182/blood-2013-01-306647.

Moise, L., Tassone, R., Latimer, H., Terry, F., Levitz, L., Haran, J.P., et al., 2013. Immunization with cross-conserved H1N1 influenza CD4+ T-cell epitopes lowers viral burden in HLA DR3 transgenic mice. Hum. Vaccin. Immunother. 9, 2060–2068. https://doi.org/10.4161/hv.26511.

Moise, L., Gutierrez, A., Kibria, F., Martin, R., Tassone, R., Liu, R., et al., 2015. iVAX: An integrated toolkit for the selection and optimization of antigens and the design of epitope-driven vaccines. Hum. Vaccin. Immunother. 11, 2312–2321. https://doi.org/10.1080/21645515.2015.1061159.

Ponomarenko, J., Vaughan, K., Sette, A., 2014. Ebola Virus Related Immune Epitope Data Curated in the IEDB. http://help.iedb.org/hc/en-us/articles/114094150431-Ebola-virus-related-immune-epitope-data-curated-in-the-IEDB-2014-.

Qiu, X., Fernando, L., Alimonti, J.B., Melito, P.L., Feldmann, F., Dick, D., et al., 2009. Mucosal immunization of cynomolgus macaques with the VSVDeltaG/ ZEBOVGP vaccine stimulates strong Ebola GP-specific immune responses. PLoS One 4, e5547. https://doi.org/10.1371/journal.pone.0005547.

Qiu, X., Wong, G., Audet, J., Bello, A., Fernando, L., Alimonti, J.B., 2014. Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp. Nature 514, 47–53. https://doi.org/10.1038/nature13777.

Ríos-Huerta, R., Monreal-Escalante, E., Govea-Alonso, D.O., Angulo, C., Rosales-Mendoza, S., 2017. Expression of an immunogenic LTB-based chimeric protein targeting *Zaire ebolavirus* epitopes from GP1 in plant cells. Plant Cell Rep. 36, 355–365. https://doi.org/10.1007/s00299-016-2088-6.

Romero-Maldonado, A., Monreal-Escalante, E., Rosales-Mendoza, S., 2016. Expression in plants of two new antigens with implications in Alzheimer's disease immunotherapy.

Plant Cell Tiss Organ Cult 126, 361–370. https://doi.org/10.1007/s11240-016-0990-9.
Snape, M.D., 2017. Persistence of immune responses induced by Ebola virus vaccines. Lancet Glob. Health 5, e238–e239. https://doi.org/10.1016/S2214-109X(17)30039-6.
Tacket, C.O., 2009. Plant-based oral vaccines: results of human trials. Curr. Top. Microbiol. Immunol. 332, 103–117. https://doi.org/10.1007/978-3-540-70868-1_6.
Takeyama, N., Kiyono, H., Yuki, Y., 2015. Plant-based vaccines for animals and humans: recent advances in technology and clinical trials. Ther. Adv. Vaccines 3, 139–154.
Tamura, K., Stetcher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729. https://doi.org/10.1093/molbev/mst197.
Vita, R., Overton, J.A., Greenbaum, J.A., Ponomarenko, J., Clark, J.D., Cantrell, J.R., 2014. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 43 (Database issue), D405–12. https://doi.org/10.1093/nar/gku938.
Wang, Y., Li, J., Hu, Y., Liang, Q., Wei, M., Zhu, F., 2017. Ebola vaccines in clinical trial: the promising candidates. Hum. Vaccin. Immunother. 13, 153–168. https://doi.org/10.1080/21645515.2016.1225637.
Yasmin, T., Nabi, A.H., 2016. B and T cell epitope-based peptides predicted from evolutionarily conserved and whole protein sequences of Ebola virus as vaccine targets. Scand. J. Immunol. 83, 321–337. https://doi.org/10.1111/sji.12425.
Zhang, Y., Li, D., Jin, X., Huang, Z., 2014. Fighting Ebola with ZMapp: spotlight on plant-made antibody. Sci. China Life Sci. 57, 987–988. https://doi.org/10.1007/s11427-014-4746-7.