UNIVERSIDADE DE LISBOA

FACULDADE DE FARMÁCIA

POLYMERASE BASIC PROTEIN 1 (PB1) AS A MOLECULAR DETERMINANT OF FITNESS AND ADAPTATION IN INFLUENZA A VIRUS

Marta Tiago Gíria

Orientadores: Professora Doutora Helena Rebelo de Andrade

Professor Doutor José Moniz Pereira

Tese especialmente elaborada para a obtenção do grau de Doutor em Farmácia, especialidade Microbiologia.

2017

UNIVERSIDADE DE LISBOA

FACULDADE DE FARMÁCIA

POLYMERASE BASIC PROTEIN 1 (PB1) AS A MOLECULAR DETERMINANT OF FITNESS AND ADAPTATION IN INFLUENZA A VIRUS

Marta Tiago Gíria

Orientadores: Professora Doutora Helena Rebelo de Andrade

Professor Doutor José Moniz Pereira

Tese especialmente elaborada para a obtenção do grau de Doutor em Farmácia, especialidade Microbiologia.

Júri:

Presidente: Professora Doutora Matilde da Luz Santos Duque da Fonseca e Castro

Vogais: Professor Doutor Paulo Jorge Pereira Cruz Paixão

Professora Doutora Perpétua da Conceição Rodrigues Gomes Cavaco Silva

Professor Doutor Miguel Agostinho Sousa Pinto Torres Fevereiro

Professor Doutor José Miguel Azevedo Pereira

Professora Doutora Helena Paula Lopes Henriques Rebelo de Andrade

Marta Gíria teve apoio financeiro da Fundação para a Ciência e Tecnologia através de uma bolsa de Doutoramento (SFHR/BD/65211/2009).

2017

Todas as afirmações efetuadas no presente documento são da exclusiva responsabilidade do seu autor, não cabendo qualquer responsabilidade à Faculdade de Farmácia da Universidade de Lisboa pelos conteúdos nela apresentados.

Acknowledgments

I would like to thank everyone that has contributed to making this project possible.

I begin by thanking Prof. Doutora Helena Rebelo de Andrade for being my mentor for the past 16 years, with the most respect and admiration for her scientific knowledge, her perseverance and her impressive ability to resolve and coordinate in extreme and emergency situations. Thank you for your supervision, for your constant guidance and for continuously pushing me forward, but most of all thank you for providing me such different work experiences which have contributed so much to my professional training and to my personal development.

I would like to thank Prof. Doutor José Moniz Pereira for permitting the development of this research project at the Departamento de Microbiologia e Imunologia da Faculdade de Farmácia da Universidade de Lisboa and for co-supervising the work developed.

I would also like to thank Prof. Doutor José Miguel Azevedo Pereira for integrating me in his lab but above all for his constant availability and for being so attentive and considerate.

To Doutora Madalena Almeida Santos from the Hospital de Curry Cabral, Lisboa, I would like to thank the essential contribution to this work by providing the clinical specimens that were a basis of our experimental approach.

Additionally, I would like to thank Dr. Ruben Donis from the Influenza Division of the National Center for Immunization & Respiratory Diseases of the Centers for Disease Control and Prevention, Atlanta, for providing my internship in reverse genetics that has proven determinant for the accomplishment of this research, and Dr. Li-Mei Chen and Dr. Guaniri Mateu-Petit for the attentive supervision.

Also, I acknowledge the Fundação para a Ciência e Tecnologia for the financial support that has permitted the development of this research.

To my colleagues and friends Luis Santos, Vanessa Correia and João Louro, I thank you for your contribution to the experimental aspects of this work. To you and to Anabela Coelho, Paulo Gonçalves and Sónia Pedro, I thank for always accompanying me.

To my colleagues and friends Rita Calado, Marta Calado, Inês Bártolo, João Perdigão, Pedro Borrego, Cheila Rocha, Joana Duarte, Andreia Martins, Francisco Martin, Carla Silva, Ana Rita Dinis, Inês Figueiredo, Inês Moranguinho, Jaciara Guimarães and Claudia Palladino, I particularly thank the extremely relevant and irrelevant conversations that have contributed so much to my support in these past years. João Perdigão, thank you for being so patiently helpful.

To my family…thank you so much. To my parents, for the high standards set and for being so caring. To my sisters Joana, Rita e Sara for being such a strong brotherhood and a reference.

Pedro, thank you for having changed my life and for always being with me, my love. Frederica, Vicente and baby, you are my life.

Preface

The research described in this thesis was performed from Abril 2010 to October 2016, under the supervision of Prof. Doutora Helena Rebelo de Andrade and the co-supervision of Prof. Doutor José Moniz Pereira.

All experimental work was performed at the Host-pathogen Interaction Unit, Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa and the major findings were communicated in the form of research papers published in peer review international journals, oral communications and poster communications. The research on influenza and epidemiology of influenza infections has contributed to strengthen our knowledge and background on emerging infections and public health, which ultimately resulted in the additional authorship of a scientific book chapter.

Papers published in peer review international journals:

Gíria M, Santos L, Louro J, Rebelo de Andrade H. Reverse genetics vaccine seeds for influenza: Proof of concept in the source of PB1 as a determinant factor in viral growth and antigen yield. Virology. 2016; 486:21-27. doi: 10.1016/j.virol.2016.05.015. Epub 2016 May 27.

Gíria M, Rebelo de Andrade H. Genetic evolution of PB1 in the zoonotic transmission of influenza A(H1) virus. Infect Genet Evol. 2014; 27: 234-243. doi: 10.1016/j.meegid.2014.07.024. Epub 2014 Aug 1.

Gíria MT, Rebelo de Andrade H, Santos LA, Correia VM, Pedro SV, Santos MA. Genomic signatures and antiviral drug susceptibility profile of A(H1N1)pdm09. J Clin Virol. 2012 Feb;53(2):140-4. doi: 10.1016/j.jcv.2011.11.002. Epub 2011 Dec 15.

Manuscript in preparation:

Gíria M, Louro J, Rebelo de Andrade H. Compatibility between influenza PB1 and HA: phenotypic evaluation of putative residues for genomic interaction and their relevance in the overall viral fitness. Manuscript in preparation.

The present thesis is based on the previous papers published in peer review international journals and manuscript in preparation.

Oral communications in international conferences:

Gíria M, Rebelo de Andrade H. Reverse genetic vaccine seeds for Influenza, presented at the Influenza Vaccines for the World - Fifth International Conference, 2015, Albufeira, Portugal.

Oral communications in national conferences:

Gíria M, Rebelo de Andrade H. Vírus Influenza do tipo A : Emergência na população humana, presented at the VIII Congresso Científico AEFFUP Doenças Infecciosas Emergentes, 2013, Porto, Portugal.

Gíria M, Rebelo de Andrade H. Compatibilidade funcional do complexo de replicação como determinante de virulência em vírus Influenza A. Presented at Seminários CPM-URIA, Faculdade de Farmácia de Universidade de Lisboa, 2012, Lisboa, Portugal.

Book chapter:

Gíria M, Rebelo de Andrade H, Valadas E. 2014. Doenças infecciosas emergentes. In: Barroso H, Melo Cristino J, Taveira N. Microbiologia Médica. Lidel – edições técnicas, 1.ª Edição, (ISBN:978-972-757-576-3). capítulo 51, pag.541-551.

Poster presentation:

Gíria M, Santos L, Louro J, Rebelo-de-Andrade H. Functional compatibility between PB1 and antigenic proteins as a determinant of viral fitness and adaptation in the A(H1N1)pdm09, presented at OPTIONS IX for the control of influenza, 2014, Chicago, USA.

Other communications:

Papers published in peer review international journals:

Correia V, Santos LA, Gíria M, Almeida-Santos MM, Rebelo-de-Andrade H. Influenza A(H1N1)pdm09 resistance and cross-decreased susceptibility to oseltamivir and zanamivir antiviral drugs. J Med Virol. 2015; 87: 45-56. doi: 10.1002/jmv.23986. Epub 2014 Jul 21.

Santos, LA, Correia V, Gíria M, Pedro S, Santos MM, Silvestre MJ, Rebelo-de-Andrade H. Genetic and Antiviral Drug Susceptibility Profiles of Pandemic A(H1N1)v Influenza Virus Circulating in Portugal. Influenza and Other Respiratory Viruses. 2011; 5 (Suppl. 1), 294?300. doi: 10.1111/j.1750-2659.2011.00221.x

Oral communications in international conferences:

L Santos, V Correia, M Gíria, S Pedro, M Santos, M Silvestre, H Rebelo-de-Andrade. Genetic and Antiviral Drug Susceptibility Profiles of Pandemic A(H1N1)v Influenza Virus Circulating in Portugal, presented at the Options for the Control of Influenza VII, 2010, Hong Kong SAR, China.

Resumo

A Organização Mundial da Saúde e o *National Institute of Allergy and Infectious Diseases* reportaram deficiências no crescimento de vírus vacinais do subtipo A(H1N1)pdm09, produzidos por genética reversa, que comprometeram a imunização eficiente e atempada durante a pandemia de 2009 e que acentuaram a necessidade de melhorar o processo de produção de vacinas. Esta situação tinha também sido anteriormente detetada em investigação com vacinas pré-pandémicas para o subtipo A(H5N1) e atribuída a uma possível interação menos eficiente entre as proteínas virais. A dinâmica da evolução genética dos vírus influenza A sugere que poderá existir um padrão de cosegregação entre a subunidade Básica da Polimerase 1 (PB1) e as proteínas antigénicas Hemaglutinina (HA) e Neuraminidase (NA). Nos episódios de rearranjo genómico que resultaram na emergência dos vírus pandémicos de 1957 e 1968, os vírus sazonais adquiriram segmentos PB1 juntamente com proteínas antigénicas com origem em vírus aviários. Em 1947 foi identificado um padrão semelhante, em que um episódio de rearranjo entre vírus sazonais envolvendo o segmento PB1 e proteínas antigénicas resultou num aumento da dispersão geográfica face ao que carateriza as epidemias sazonais. Nestas situações, a presença do segmento PB1 homólogo às proteínas antigénicas parece ter favorecido o *fitness* viral. Em estudos retrospetivos acerca da composição genómica de vírus vacinais de elevado rendimento, produzidos por rearranjo clássico, foi frequentemente detetada a incorporação do segmento PB1 juntamente com os das proteínas antigénicas, sugerindo também que a sua interação poderá ter impacto no *fitness* viral.

Neste contexto, propusemos avaliar se o segmento PB1 é um determinante molecular do *fitness* viral, se especificamente a sua compatibilidade com as proteínas antigénicas condiciona o *fitness*

e a adaptação dos vírus influenza A e, também, se este conceito será passível de ser explorado para melhorar a produção de vírus vacinais. O vírus A(H1N1)pdm09 foi utilizado como modelo nesta investigação por ser um produto de rearranjo com uma combinação única de segmentos com origem aviária, suína e humana. Adicionalmente, o vírus vacinal pandémico apresentou deficiências no crescimento e, uma vez que o A(H1N1)pdm09 se mantém em circulação com um perfil sazonal, permanece a necessidade de produzir vacinas anualmente para fazer face aos frequentes *drifts* antigénicos.

Os objetivos desta investigação foram definidos como 1) avaliar a evolução genética do segmento PB1 na transmissão zoonótica de vírus suínos e inferir a sua possível contribuição para o *fitness* viral e a adaptação e 2) determinar se a compatibilidade funcional ou estrutural entre o segmento PB1 e as proteínas antigénicas é um determinante do *fitness* viral, em protótipo de vírus vacinal A(H1N1)pdm09 utilizado como modelo, produzido por genética reversa.

Para a concretização do objetivo 1, foi selecionada uma amostra de sequências nucleotídicas de segmentos PB1 de vírus suínos que infetaram o homem e efetuada uma analise filogenética e uma avaliação de mutações. Os resultados mais relevantes incluíram a confirmação de que é possível identificar a história evolutiva do segmento PB1 no que diz respeito à distinção entre linhagens e hospedeiros. Ainda, foram detetados possíveis marcadores de adaptação a hospedeiros mamíferos no segmento PB1, como os 336I, 361R, 468K e 584Q, e a novas composições genómicas, o que ocorre provavelmente na sequência de episódios de rearranjo, como os 638D e 618D. Os resíduos 298I, 386K e 517V na PB1 foram identificados como possivelmente associados a uma maior compatibilidade entre PB1 e HA do subtipo H1, em hospedeiros mamíferos. O possível papel destes marcadores na adaptação viral foi inferido com base na epidemiologia molecular dos vírus, na localização genómica dos marcadores e nas propriedades dos aminoácidos envolvidos. O impacto fenotípico da aquisição das mutações L298I, R386K e I517V pelo A(H1N1)pdm09 durante a sua história evolutiva, foi também avaliado *in vitro* através da construção de vírus recombinantes por genética reversa. Os resultados obtidos permitiram-nos sugerir que a aquisição das mutações poderá ter resultado em alterações na conformação da proteína PB1, mas também no aumento do número de nucleótidos complementares aos do segmento HA e que poderão estar envolvidos na interação entre os dois segmentos ao nível do RNA. Por outro lado, verificámos que a aquisição destas mutações teve

um impacto negativo na cinética de crescimento viral *in vitro*, o que sugere que a interação entre os segmentos genómicos ao nível do RNA pode ser um fator determinante para a cosegregação mas que os mecanismos que lhe estão associados provavelmente não serão dependentes de uma vantagem replicativa. Esta situação seria concordante com um modelo de empacotamento seletivo das partículas virais proposto por outros autores.

Para a concretização do objetivo 2) determinar se a compatibilidade funcional ou estrutural entre o segmento PB1 e as proteínas antigénicas é um determinante do *fitness* viral em protótipo de vírus vacinal A(H1N1)pdm09 produzido por genética reversa, foi selecionada uma estirpe imunogénica A(H1N1)pdm09 protótipo. Com base nesta estirpe foram construídos protótipos de vírus vacinais com a composição genómica do vírus vacinal, embora contendo o segmento PB1 homólogo ou heterólogo às proteínas antigénicas, e avaliados parâmetros da sua cinética de crescimento e do rendimento em antigénio. Para a seleção de uma estripe imunogénica protótipo foi avaliado o perfil genético de uma amostra de estirpes de A(H1N1)pdm09, que circularam em Portugal durante o período pandémico, e selecionada uma estirpe semelhante ao consenso. Os protótipos de vírus vacinais foram construídos por genética reversa num esqueleto de A/PuertoRico/08/34, com a composição genómica 6:2 do protótipo vacinal clássico (PR8:HA,NA A(H1N1)pdm09) e a composição 5:3 na qual o segmento PB1 incorporado é o da estirpe imunogénica, juntamente com as proteínas antigénicas HA e NA (PR8:HA,NA,PB1 A(H1N1)pdm09). Esta abordagem permitiu identificar que a presença do segmento PB1 homólogo às proteínas antigénicas resultou num aumento significativo na cinética de crescimento viral, capacidade hemaglutinante e atividade da Neuraminidase. Com base nestes resultados, consideramos que poderá ser possível obter uma melhoria significativa do crescimento viral e rendimento em antigénio em protótipos vacinais PR8:A(H1N1)pdm09 produzidos por genética reversa, em comparação com o protótipo vacinal clássico, através da introdução do segmento PB1 da estirpe imunogénica.

Consideramos ainda que, adicionalmente ao papel da subunidade PB1 da polimerase na replicação viral, o segmento genómico PB1 poderá ser um determinante molecular do *fitness* e um fator determinante na epidemiologia molecular dos vírus influenza, através de interações que estabelece com outros segmentos genómicos ao nível do RNA e da sua aparente capacidade de acumular alterações genéticas adaptativas que são o maior fator impulsionador do *fitness*

viral. É necessária mais investigação para clarificar os mecanismos de empacotamento viral, o papel das interações ao nível do RNA em estabelecer padrões de cosegregação e as especificidades destas interações de acordo com o subtipo viral. No entanto, a compatibilidade funcional ou estrutural entre proteínas ou segmentos genómicos parece poder ser explorada para aumentar o *fitness* viral e melhorar a produção de vírus vacinais de outros subtipos. Ainda, uma vez que se reconhece que a composição genómica dos vírus influenza pode ter um impacto significativo no *fitness* viral, e consequentemente constituir um determinante de virulência, consideramos que incluir a sua análise na avaliação de risco de novas estirpes seria extremamente relevante no contexto sazonal e de ameaça pandémica.

Abstract

The World Health Organization and the National Institute of Allergy and Infectious Diseases reported growth deficits of influenza A(H1N1)pdm09 reverse genetic pandemic vaccine virus seeds. These have compromised the effective and timely distribution of vaccines in the 2009 pandemics and accentuated the need to improve the process of vaccine production. In pre-pandemic A(H5N1) research, seed viruses produced by reverse genetics have also been reported to present growth deficits. These deficits have been attributed to a putative sub-optimal protein interaction. The dynamics of the genetic evolution of influenza A viruses appears to suggest a gene segregation pattern between the Polymerase Basic protein 1 (PB1) and antigenic proteins Hemagglutinin (HA) and Neuraminidase (NA). In the reassortment events that lead to the emergence of the 1957 e 1968 pandemic viruses, the contemporary seasonal viruses acquired PB1 genomic segment together with antigenic glycoproteins originating from avian viruses. A similar pattern was identified in 1947, where a reassortment event between seasonal viruses, involving PB1 and antigenic proteins, has altered the epidemiology of the infection to a near-pandemic geographic dispersion. In both situations, viral fitness appears to have benefitted from acquiring a PB1 genomic segment homologous to antigenic proteins. Also, in retrospective studies on the genomic composition of high yield seasonal vaccine seeds produced by classical reassortment, PB1 is frequently co-incorporated with antigenic proteins HA and NA, further suggesting that the interaction between these proteins could have an impact in viral fitness. In this context, we proposed to address the question of PB1 genomic segment being a molecular determinant of fitness and adaptation in influenza A virus and, particularly, of the functional compatibility between PB1 and antigenic proteins being a driver of the overall viral fitness and

putatively exploitable to improve seed virus production. The A(H1N1)pdm09 virus was used a model for this research because it is a product of viral reassortment with an unprecedented genomic composition of segments originating from avian, swine and human seasonal viruses. Additionally, the 2009 pandemic vaccine virus presented severe growth deficits and, since the A(H1N1)pdm09 persists in circulation with a seasonal epidemiologic profile, the demand for high yield A(H1N1)pdm09 vaccine seeds will be continuous and the need to adequate the immunogenic strain to the circulating viruses will be recurrent because of antigenic *drifts*. The objectives of this research were defined as 1) to evaluate the genetic evolution of PB1 in the zoonotic transmission of swine influenza virus and infer its putative contribution towards viral fitness and adaptation, and 2) to determine if the functional or structural compatibility between PB1 and antigenic proteins is a molecular determinant of the overall virus fitness in the reverse genetics A(H1N1)pdm09 vaccine seed model.

The approach followed to accomplish objective 1 was to select a study sample of PB1 nucleotide sequences from swine virus that have infected the human host, to analyze phylogeny and mutation trends and to search for putative markers for viral adaptation on the basis of viral molecular epidemiology, genomic location of the polymorphisms and amino-acid properties. Our major findings were that the evolutionary history of PB1 is traceable in terms of lineage and host origin. Specific genomic markers in PB1 appear to putatively relate to the viral adaptation to mammalian hosts, 336I, 361R, 468K and 584Q, and to the viral adaptation to new genomic backgrounds possibly in the sequence of reassortment events, such as 638D and 618D. Residues 298I, 386K and 517V have been found to putatively relate to an enhanced compatibility between PB1 and HA of the H1 subtype, in the mammalian host. A subsequent *in vitro* investigation of the phenotypic impact of mutations L298I, R386K and I517V acquired by the A(H1N1)pdm09 during its evolutionary history, was performed by generating an A(H1N1)pdm09 recombinant virus and an A(H1N1)pdm09 reassortant in which the specific mutations have been reverted, by reverse genetics. This approach has resulted in two major findings. Acquiring these mutations has been found to putatively promote conformational changes in PB1 and enhance the span of complementary nucleotides possibly involved in PB1 interaction with HA at the RNA level and, on the other hand, has proven detrimental to viral growth kinetics *in vitro*. These findings have lead us to suggest that the interaction between genomic segments at the RNA level could be a

determinant of co-segregation, concordant with a selective packaging model proposed by other authors, but that the mechanisms that drive this process are probably not dependent on a replicative advantage.

Our approach to accomplishing objective 2) to determine if the functional or structural compatibility between PB1 and antigenic proteins is a molecular determinant of the overall virus fitness in the reverse genetic A(H1N1)pdm09 vaccine seed model, was to determine the genetic profile of A(H1N1)pdm09 strains circulating in Portugal during the pandemic period and select a prototype immunogenic strain, to generate reassortant viruses with the genomic composition of A(H1N1)pdm09 seed viruses prototypes bearing PB1 homologous and heterologous to antigenic proteins, and to evaluate viral growth and antigen yield *in vitro*. A sample of specimens collected from the pandemic period in Portugal were evaluated for genetic and phenotypic features and a strain similar to the consensus was selected as a prototype strain. Vaccine seed prototypes of the selected A(H1N1)pdm09 strain in an A/PuertoRico/08/34 backbone were generated by reverse genetics to present the genomic compositions of the 6:2 classical vaccine seed (PR8:HA,NA A(H1N1)pdm09) and a 5:3 seed prototype in which the PB1 segment from the immunogenic strain is co-incorporated with the antigenic proteins (PR8:HA,NA,PB1 A(H1N1)pdm09). Our major findings were that the presence of PB1 homologous to antigenic protein significantly increased viral replication, hemagglutination capacity and Neuraminidase activity. We have establishing proof of concept that, in the PR8:A(H1N1)pdm09 seed virus model, viral growth and antigen yield can be significantly improved by the inclusion of PB1 from the immunogenic strain when compared to the classical seed virus prototype.

We consider that, additionally to the role of PB1 protein in viral replication, PB1 genomic segment may be a molecular determinant of the overall virus fitness and a determinant factor in the molecular epidemiology of the viruses by establishing interactions with other segments at the RNA level and by, apparently, being able to genetically change and adapt to improve these interactions. Further research is necessary to clarify the mechanisms of viral genome packaging, the role of interactions at the RNA level in establishing the co-segregation patterns and the specificities of this interactions at the subtype level. However, it becomes clear that the functional compatibility between PB1 and antigenic proteins is a driver of the overall viral fitness in the A(H1N1)pdm09 and is putatively exploitable to improve seed virus production.

We also consider that exploring the concept of the compatibility between gene segments or proteins being a determinant factor in the overall viral fitness, can result in major improvements in the production of reverse genetics seed viruses of different influenza subtypes. Also, being aware of the fact that the genomic composition of influenza viruses can have a major phenotypic impact, and that consequently is a determinant of virulence even though the mechanisms that drive the selective packaging remain unclear, we consider that its inclusion in the risk assessment of influenza strains would be extremely relevant for seasonal and pandemic preparedness.

Abbreviations

µl	Microliters
µg	Micrograms
A	Alanine
aa	Amino acid
AIV	Avian Influenza Virus
BLAST	Basic Local Alignment Search Tool
b	Nucleotide base
bp	Base pair
C	Cysteine
CDC	Centers for Disease Control and Prevention
cDNA	Complementary Deoxyribonucleic Acid
D	Aspartic Acid
dN	Non-synonymous substitution
DNA	Deoxyribonucleic Acid
dS	Synonymous substitution
E	Glutamic Acid
ECDC	European Centers for Disease Control and Prevention
EMA	European Medicines Agency
FDA	Food and Drug Administration
G	Glycine
GPRBC	Guinea Pig Red Blood Cells

H	Histidine
HA	Hemagglutinin
HRBC	Human Red Blood Cells
I	Isoleucine
ILI	Influenza-Like illness
K	Lysine
L	Leucine
LRT	Lower Respiratory Tract
M	Matrix protein
M	Methionine
MDCK	Madin Darby Canine Kidney Cells
ml	Milliliters
MOI	Multiplicity of Infection
mRNA	Messenger Ribonucleic Acid
MTS	Mitochondrial Targeting Sequence
N	Asparagine
NA	Neuraminidase
NAI	Neuraminidase Inhibitor
NEP	Nuclear Export protein
NIAID	National Institute of Allergy and Infectious Diseases
NLS	Nuclear Localization Signal
NP	Nucleoprotein
NS	Non-Structural protein
nt	Nucleotide
ORF	Open Reading Frame
P	Proline
PA	Polymerase Acidic protein
PB1	Polymerase Basic protein 1
PB2	Polymerase Basic protein 2
PCR	Polymerase Chain Reaction

PR8	A/PuertoRico/08/34
Q	Glutamine
R	Arginine
rfu	Relative Fluorescence Units
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
RT	Reverse Transcription
RT-PCR	Reverse Transcription Polymerase Chain Reaction
rtRT-PCR	Real Time Reverse Transcription Polymerase Chain Reaction
S	Serine
SIV	Swine Influenza Virus
SOIV	Swine Origin Influenza Virus
T	Threonine
TCID	Tissue Culture Infectious Dose
TRIG	Triple Reassortment Internal Genes
TR-SOIV	Triple Reassortant Swine Origin Influenza Virus
URT	Upper Respiratory Tract
V	Valine
vRNA	Viral Ribonucleic Acid
W	Tryptophan
WHO	World Health Organization
wt	Wild-Type
Y	Tyrosine

Table of contents

Table index

Figure index

Figure index

Chapter 1

General introduction

1.1 Biology of influenza A viruses and human disease

1.1.1 Classification

Influenza viruses belong to the *Orthomyxoviridae* family, which includes types A through D distinguished by host range, variability of the antigenic surface glycoproteins Hemagglutinin (HA) and Neuraminidase (NA) and genome organization[1-4]. They are enveloped, pleomorphic viruses with segmented single strand negative RNA (ss(-)RNA)) genomes[1, 2].

Influenza A viruses infect a wide range of mammalian and avian species. They are characterized by the subtype of the antigenic surface glycoproteins and further distinguished into strains. There are 18 subtypes of HA and 11 of NA identified to date[4]. Avian influenza viruses (AIV) are further classified as highly pathogenic (Highly Pathogenic Avian Influenza, HPAI) or low pathogenic (Low Pathogenic Avian Influenza, HPAI), according to its pathogeny and mortality rate in chickens in laboratory setting[4]. Because of the genetic and antigenic diversity, host-range and ability to reassort, influenza A viruses have the potential to cause pandemics[4, 5].

Influenza type B predominantly infects the human host. Type B viruses are classified into lineages and further divided into strains[4]. The genetic diversity is lower when compared to influenza A viruses and because of this and the more restricted host range, influenza B viruses may cause seasonal epidemics but do not cause pandemics[4]. Influenza type C viruses also predominantly infect the human host but are not known to cause epidemics. Type D viruses infects cattle and are not known to infect the human host[4].

1.1.2 General organization of the genome and viral proteins

Influenza A genome is composed of 8 segments that code for up to 15 proteins identified to date, although for some of these proteins the levels of expression in different viral subtypes or different hosts may vary and the exact role has not been clarified[1, 6-8]. Segments range from 890 to 2341 bases (b) long, in a total genome size of 13,5Kb. Each viral RNA (vRNA) segment is composed of a central coding region in antisense orientation. Flanking this region are 3' and 5' untranslated short sequences (up to 58nt long) that include a 12nt long and a 13nt long sequence,

respectively, highly conserved between all vRNAs[6]. The vRNA segments are associated with the polymerase complex and the Nucleoprotein (NP) forming viral ribonucleoproteins vRNPs[6].

The segments PB2, PB1 and PA code for the polymerase complex, composed by the Polymerase Basic protein 2 (PB2), Polymerase Basic protein 1 (PB1) and Polymerase Acidic protein (PA), responsible for replication and transcription. PB2 and PB1 genomic segments are 2341b long each. PB1 segment, additionally to PB1 protein, codes for PB1-F2 and N40 proteins. PB1-F2 is only found in infected cells. It is coded from a different open reading frame (ORF), ORF(+1), and is primarily associated with the induction of cellular apoptosis in a late stage of infection. N40 is a truncated form of PB1 at the N-terminal with no ability to bind to PA. It is thought to have a role in the regulation of transcription, but it remains unclear to what extent[1, 2]. PA genomic segment is 2233b long and, additionally to the PA protein, codes for PA-X, PA-155 and PA-182. PA-X is a truncated form of PA which appears to be involved in the modulation of the host response to infection by repressing cellular expression, although further research is necessary to clarify its role[7]. PA-155 and PA-182 are N-truncated forms of the PA protein whose function is not clarified[8].

The HA and NA segments, 1778 and 1413b long respectively, encode the Hemagglutinin and Neuraminidase antigenic transmembrane glycoproteins[1, 2].

The NP segment is 1565b long and encodes the Nucleoprotein, which has the primary function of encapsulating the virus genome for the purposes of RNA transcription, replication and packaging[1, 2].

The M segment is 1027b long and codes for M1 and M2 proteins. M1 is the Matrix protein, the fundamental structural component of the virus particle. It binds the membrane, interacts with both viral RNA and NP of the RNP complex and interacts with components of the host cells mostly for regulatory functions. M2 is an ionic-channel with the main function of regulating pH across the viral membrane, which is important in both the early and late stage of the replication cycle[1, 2].

The NS segment is 890b long and codes for Non-Structural proteins 1 and 2 (NS1 and NS2) which are present in the cell but are not incorporated in the infectious virus particle. NS1 protein is mostly associated with the activation of anti-apoptotic mechanisms in the early stage of cell infection. NS2, also designated as the Nuclear Export protein (NEP), mediates the export of RNP complexes from the nucleus[1, 2].

1.1.3 Life cycle and replication

Transmission of influenza A viruses occurs via aerosols and droplets containing viruses expelled by infected individuals, either by direct or indirect contact and airborne[9].

The virus HA binds to sialic acid (N-acetyl-neuraminic acid) cellular receptors and enters the cell by receptor-mediated endocytoses in an endosome[9, 10]. The acidic pH of the endosome leads to a conformational change in HA exposing the HA2 subunit, which is a fusion peptide that drives the fusion between the endosomal and viral membranes[9, 10]. The acidic pH of the endosome also leads to the opening of the M2 ionic channel, the acidification of the virion and the consequent release of the vRNPs from M1 protein and entry into the cells cytoplasm[9, 10]. The proteins composing the vRNP, PB2, PB1, PA and NP have Nuclear Localization Signals (NLS) which are recognized by the cellular nuclear import mechanisms[9, 10].

The influenza A virus replication process is multigenic. In the nucleus, the negative single strands of RNA are converted into positive sense RNA, designated as cRNA, which is the template for vRNA replication. The viral RNA-dependent RNA polymerase transcribes vRNA to mRNA[9, 10]. The mRNA is transported to the cytoplasm for translation into viral proteins using the cellular mechanisms[9, 10]. Still in the nucleus, the viral RNA-dependent RNA polymerase replicates vRNA which is assembled into vRNP together with the new PB2, PB1, PA and NP proteins, which enter the nucleus after being translated in the cytoplasm. The vRNPs are transported to the cytoplasm and used in the assembly of new virus particles[9, 10]. RNP export from the nucleus occurs via NS2/NEP. M1 protein association with NEP is essential for the export signaling of RNP from the nucleus to the cytoplasm, for protein synthesis, and to prevent reuptake of RNP to the nucleus. The new virus particles are formed using the cellular membrane, where HA, NA and M2 are incorporated[9, 10].

The same gene segments coding for the proteins involved in the replication process, code for additional proteins that are associated with the control of cellular apoptosis during the viral infection. PB1-F2 and NS1 initiate antagonist mechanisms of cell apoptosis in different stages of viral infection. NS1 is associated with the activation of anti-apoptotic mechanisms on the early stage of cell infection. These mechanisms will inhibit programmed cell death induced by the viral

infection, permitting replication to occur[11, 12]. PB1-F2 has a known pro-apoptotic function in the host cells in a later stage of infection, promoting the release of newly formed infectious virus particles. Also, this protein has been associated with an increase of the polymerase activity by binding to the viral polymerase PB1 subunit[13]. The expression of PB1-F2 is a known determinant of virulence, although the mechanisms behind the increased virulence are not fully clarified.

The cassette of internal genes of influenza A virus, coding for non-membrane non-antigenic proteins, therefore controls the major processes by which viral fitness is determined: replication and induction of apoptosis.

1.1.4 Human disease, prevention and treatment

Influenza A viruses infect the human respiratory epithelium and causes a respiratory illness mostly associated with fever, cough, sore throat and fatigue. Influenza can lead to severe complications and death, predominantly in risk groups including children, elderly, pregnant women and individuals with medical conditions such as weakened immune systems, extreme obesity or chronic diseases[14]. It is highly contagious, with an estimated reproduction number (R) of 1,28 secondary cases generated for typical infectious case in seasonal epidemics (seasonal influenza addressed in section 1.2 Ecology and epidemiology of influenza A viruses of the General Introduction chapter)[15]. Pandemic influenza viruses are associated with the introduction of new antigenic proteins in the human host and the absence of protective immunity can lead to an increase in the number of cases. This has occurred in the past pandemics, where R has ranged from 1,46 to 1,8 secondary cases generated for typical infectious case (pandemic influenza is reviewed in section 1.2 Ecology and epidemiology of influenza A viruses, of the General Introduction chapter)[15].

Vaccination against influenza is recommended by the World Health Organization (WHO) as the most effective measure to prevent infection and severe outcomes in risk groups with increased risk of exposure, such as the health care workers, and in risk groups for developing severe complications, identified before[16]. Influenza vaccines have to be reformulated in a regular basis, given the genetic evolution and the molecular epidemiology of the viruses (the genetic

evolution and epidemiology of influenza A viruses is reviewed in sections *1.3.3 Generation of diversity in Influenza A viruses* and *1.2.2 Historical perspective of influenza A emergence in the human host and current seasonal epidemiology*, of the General Introduction chapter). Particularly, in seasonal vaccination, the selection of strains that compose the vaccine needs to be addressed annually because of viral antigenic *drifts*, to avoid mismatches and to maximize immune protection for every season, for both southern and northern hemispheres[16]. Also, new antigenic glycoproteins are recurrently introduced in the human host by zoonotic transmission and novel vaccines may have to be developed for the prevention or mitigation of the infection by these emergent immunogenic strains[16].

Currently, Inactivated Influenza Vaccines (IIV), such as whole virus, split and subunit vaccines, and Live Attenuated Influenza Vaccines (LAIV) are licensed and available[16, 17]. IIV consist of influenza vaccine viruses propagated in embryonated eggs or cell cultures and subsequently inactivated[18, 19]. These constitute whole virus vaccines, which can also be disrupted and used as subvirion preparations designated as split vaccines, retaining the immunogenic characteristics but being less reactive when compared to whole virus[18]. Split vaccines have been the most extensively used in seasonal influenza[20]. For subunit vaccines, the antigenic proteins HA and NA are isolated and further purified, also retaining immunogenicity and reduced reactivity[19]. All IIV are recommended for intramuscular administration[18]. LAIV consist of vaccine viruses attenuated to cold adaptation and temperature sensitive features, which permits its replication in the nasopharynx but limits it in the lower respiratory tract (LRT)[19, 21]. LAIV are administered intranasally and can also induce mucosal immunity in addition to antibody response[20].

IIV and LAIV are intended to induce the production of antibodies directed towards the HA protein, which contribute primarily to prevent the illness, and the NA protein, contributing to reduce the severity[18, 20]. Because these are highly variable proteins, the effectiveness of the vaccines may be reduced due to mismatches with the current circulating strains, as referred above[20].

Given the restriction in the use of LAIV in risk groups for developing severe complications or their care givers because of virus shedding, these vaccines have not been recommended by the Advisory Committee on Immunization Practices (ACIP) for the forthcoming 2016-2017 season[20, 22].

Most seasonal influenza vaccines include two influenza A strains, of the A(H1N1) and the A(H3N2) subtypes, and a type B strain, which are selected based on prediction of strains likely to circulate in the subsequent influenza season (current seasonal epidemiology of influenza A viruses reviewed in the *1.2.2 Historical perspective of influenza A emergence in the human host and current seasonal epidemiology* section). Quadrivalent vaccines including 2 influenza type B strains of the major lineages circulating have also been licensed in the Unites States of America in 2012 and in Europe in 2013[16, 17, 23]. In Europe, and particularly at the national level, vaccination with IIV is recommended to risk groups of developing serious complications, such as chronic patients, individuals >65 years and pregnant women, and to risk groups for transmitting the infections such as the health professionals and other care givers[24, 25].

Antivirals are available for treatment, but the recommendation for their use differs worldwide, from a more generalized use to reduce the probability of developing illness, as in the United States, to a more restricted use in situations where the vaccine may fail or is not recommended, as occurs in Europe. Specifically, at the national level, the use of antivirals is recommended by the General Directorate of Health to prevent severe complications or to prevent transmission by health professionals to risk groups[26-28].

The two major classes of antivirals are Neuraminidase Inhibitors (NAI) and M2 Inhibitors. Currently, the European Medicines Agency (EMA) approved influenza antiviral drugs are NAI Oseltamivir and Zanamivir, whose mechanisms of action is to block NA function preventing the release and spread of infectious virus particles. The Food and Drug Administration (FDA) has licensed a third NAI, Peramivir[26, 29, 30]. Although some strains with reduced inhibition for Oseltamivir have been detected worldwide, most viruses analyzed from the 2015/2016 influenza season were susceptible to Oseltamivir and all were susceptible to Zanamivir[30, 31]. Because of structural differences, mutations associated with a reduced inhibition are not necessarily the same for these drugs and may not have the same effect in all influenza subtypes[29]. Influenza virus circulating in the 2015/2016 season continued to be susceptible to Neuraminidase Inhibitors and their use is recommended for the forthcoming season of 2016/2017[26, 32].

The mechanisms of action of M2 Inhibitors is to specifically block the ion channel function of the M2 protein, which consequently interferers with the steps of the life cycle that involve this protein.

These have been the first class of antivirals developed but its effectiveness has been limited by the rapid spread of drug resistant mutations. All viruses analyzed from the 2015/2016 influenza season were resistant to the two FDA and EMA approved M2 Inhibitors Amantadine and Rimantadine, and their use is not recommended for the forthcoming season of 2016/2017[26, 30, 32].

1.2 Ecology and epidemiology of influenza A viruses

1.2.1 Reservoir, intermediate hosts and the human-animal interface

Wild aquatic birds are the primary natural reservoir of influenza viruses. All influenza A subtypes circulate in this reservoir, with the exception of H17N10 and H18N11 which have only been identified in bats. Consequently, this species may also be acting as a reservoir[33]. In the avian reservoir, infection by most subtypes is asymptomatic or causes mild disease but HPAI subtypes such as H5 or H7 can cause severe symptoms and death. AIV spread worldwide because of migrations and transmit to domestic poultry where they become closer to human and other animal hosts. AIV are considered species-specific but interspecies transmission occasionally occurs and the avian reservoir is thought to be the source of influenza A virus in other animal species[34, 35]. Additionally to the human host, a broad range of different animal hosts have been infected with influenza A viruses such as swine, poultry, canine, equine, marine mammals and small number of other incidental hosts[33, 36].

The major risk associated with the zoonotic transmission of AIV is the possibility of these viruses being able to adapt and effectively transmit within the human host. This may have been the origin of the 1918 influenza pandemics, described in section *1.2.2 Historical perspective of the emergence of influenza A viruses in the human host and current seasonal epidemiology*, although aspects of the interspecies transmission profile between the avian reservoir and the human and swine host populations remain unclear.

AIV subtypes H5, H6, H7, H9 and H10 have intermittently infected the human host in 1996, 1997, 2003-5, 2007, 2012-4 and 2016[37-45]. Most infections have been associated with the A(H5N1) subtype, over 800 from 2003 to 2016 with an approximate 50% fatality rate, and the A(H7N9)

subtype, which has caused over 570 human infections since 2013, with a fatality rate of approximately 40%[46-48].

The zoonotic transmission of AIV was thought to require an intermediate host, such as the swine, to permit the adaptation of the virus to effectively replicate and transmit within the human host [49, 50]. However, in 1997, during an outbreak of avian A(H5N1) in China, direct transmission from poultry to the human host was recognized for the first time and the intermediate infection of swine may now be considered a facilitating factor other than a requirement for human infection with an AIV[18, 19].

The pathogenicity in poultry is not necessarily reflected in the human host. HPAI A(H5N1) is highly pathogenic in humans but the LPAI A(H7N9), for example, has also been found to cause severe human disease and death, and both LPAI and HPAI A(H7) strains have been associated with conjunctivitis and influenza-like-illness (ILI) [46, 51, 52]. Also, there is insufficient data to clarify if the human cases of infection are proportional to the level of circulation of these viruses in the reservoir or if some subtypes infect the human host more easily[51]. AIV A(H10N7) and A(H10N8), for example, have caused human infections but its circulation in the avian reservoir appears to be very limited[53-55].

Most of the reported cases of zoonotic transmission of AIV occurred in China and the Middle East, where the viruses circulate among poultry, subsequently to close exposure to infected poultry or contaminated environments, and this remains the primary risk factor for human host infection[56]. The situations where human-to-human transmission have been detected were occasional, associated with very close, unprotected and prolonged contact between an infected individual and a caregiver, and eventually self-exhausted[56, 57].

In parallel to the threat of direct human infection with an AIV, and its adaptation to replicate and transmit within the human host, and of the infection of an intermediate mammalian host where this adaptation can occur, AIVs pose an additional threat. Their reassortment with a human, or other mammal adapted influenza virus, may result in the emergence of a new virus with an already established ability to effectively transmit within the human host. This has been the origin of the 1957 and 1968 pandemics, described in section *1.2.2 Historical perspective of the emergence of*

influenza A viruses in the human host and current seasonal epidemiology of the General Introduction chapter.

The swine population, additionally to possibly acting as an intermediate host to human infection with AIV and facilitate the adaption of the virus to a mammalian host, has also been considered a possible missing vessel, where avian and human or swine viruses can reassort. Swine can be naturally infected with avian and human viruses, because of the sialic acid receptors expressed, and co-infection can potentiate reassortment. However, the human and swine epithelia has been proposed to have similar receptor distribution, which raises the possibility of human host also acting a mixing vessel[58, 59].

Swine influenza virus (SIV) subtypes A(H1N1) and A(H3N2) circulate and cause outbreaks in the swine population, mostly reported in the United States. SIV do not generally infect the human host but sporadic infection with swine subtypes A(H1N1), A(H3N2) and with viral reassortants and genetic variants of these subtypes, have occurred[60-62]. Because it is clinically similar to infection with seasonal viruses, the majority of cases are presumed to remain undiagnosed. There were over 50 cases reported from 1970 to 2000 by SIV A(H1N1) and 11 notifications of human infection by a Triple-Reassortant SIV A(H3N2) since 2005. In 2009 the emergence of a SIV A(H1N1) caused a pandemic in the human host, designated as A(H1N1)pdm09, as described in section *1.2.2 Historical perspective of the emergence of influenza A viruses in the human host and current seasonal epidemiology*. After this period, since 2011, approximately 300 cases of human infection with a genetic variant of SIV A(H3N2) which has acquired the M gene from the A(H1N1)pdm09 virus, A(H3N2)v, have been detected[51, 63-66]. In 2015 and 2016, 7 cases of human infection with SIV A(H1N1) and A(H3N2)v were reported[45, 67]. The rise in the number of cases of human infection with SIV is mostly attributed to an enhancement in the surveillance system. Mostly all cases of zoonotic transmission have occurred by close contact with infected swine, in the United States, and, with the exception of the 2009 pandemics, only limited human-to-human transmission has been detected[62, 68].

Infection of the swine host with avian A(H9N12) and A(H4N6) viruses have also been reported, but the viruses do not seem to be enzootic in pigs and no zoonotic transmissions to the human host has been detected[69, 70].

The transmission of influenza virus from the human to the swine host has also been shown to occur[71]. Since 1990, at least 20 introductions of human seasonal viruses in the swine population have been detected in the United States and, between 2009 and 2011, approximately 50 introductions of the A(H1N1)pdm09 pandemic virus from the human host to the swine population were reported globally[72].

Other animal hosts include horses and dogs. Influenza A subtype A(H3N8) circulate in horses causing outbreaks and has been transmitted to the canine host where it now co-circulates with an A(H3N2) subtype[73]. In 2011, an outbreak of infection with A(H3N8) was also detected in seals. A(H3N8) is thought to have been transmitted directly from avian species, because of the genetic homology to the avian virus and because similar mammal adaptive mutations were found in these hosts[74].

The biodiversity of the reservoir and animal hosts is a determinant source of emergent subtypes. On the other hand, the genetic diversity and evolutionary dynamics of influenza viruses are crucial factors that may permit the efficient zoonotic transmission and the sustained transmission in the new host species. Nevertheless, the human activity remains a critical factor driving the emergence of influenza subtypes in the human host because it extends the interface by creating exposure.

The interspecies transmission of influenza virus is a continual threat to public health and understanding the determinants of viral fitness and the mechanisms of adaptation is crucial for the risk assessment of emerging subtypes or variants.

1.2.2 Historical perspective of the emergence of influenza A viruses in the human host and current seasonal epidemiology

In 1918, an AIV A(H1N1) emerged in the human and swine populations. The interspecies transmission profile, from the avian reservoir to the human and swine host populations, remains unclear[64, 75]. The virus caused a pandemic in the human host, designated as the "Spanish Flu"

and is estimated to have resulted in 20–50 million deaths. Since the post-pandemic period, the virus has acquired a seasonal epidemiologic profile and its subsequent genetic evolution appears to have been characterized by mutation and intra-subtype reassortment events[76]. These have resulted in the co-circulation of different clades which have caused epidemic outbreaks of increased severity in 1928-1929, 1932-1933, 1936-1937 and in 1943-1944, in Europe and the United States of America[76].

In 1947, a seasonal A(H1N1) strain from a minor clade, not dominant and undetected at the time, acquired PB1, NA and NP genomic segments from a strain belonging to the dominant A(H1N1) clade by a reassortment event. The emerging virus designated as A-Prime has caused a total vaccine failure and has spread with a near pandemic geographical dispersion although reduced mortality[64, 76]. Again, in 1951, a seasonal A(H1N1) intra-subtype reassortment event has caused a strain to acquire PB2 and HA segments from a second strain, contemporary but phylogenetically more similar to the ones that circulated around 1940. The severity of the disease was enhanced when compared to a common influenza season, but the HA had little antigenic change and virulence was attributed to a putative increase in replication capacity[64, 76]. Intra-subtype reassortment between seasonal viruses is an important factor in the dynamics of influenza A evolution and these events demonstrate a putatively relevant epidemiologic consequence in terms of severity and geographic dispersion.

The circulation of the seasonal HA and NA antigenic proteins of the A(H1N1) subtype has been eliminated from human host in 1957, in an inter-subtype reassortment event where the seasonal virus acquired new HA, NA and PB1 genomic segments from an A(H2N2) avian origin. The emerging reassortant caused a pandemic, designated as the "Asian Flu" and estimated to have caused 1–4 million deaths, after which has acquired a seasonal epidemiologic profile[64, 76].

In 1968, the seasonal A(H2N2) virus acquired HA and PB1 from an avian subtype A(H3) origin, by reassortment, and the emerging A(H3N2) caused the third documented pandemic in the human host designated the "Hong Kong Flu" and also estimated to have caused approximately 1–4 million deaths. The pandemic virus has subsequently acquired a seasonal epidemiologic profile in the human population[64, 76].

Twenty years past the elimination of the A(H1N1) subtype, in 1977, a putative accidental laboratory release causes the emergence of an A(H1N1) strain phylogenetically similar to the seasonal strains that circulated in 1950, but distinct from the ones isolated in 1947 and 1957[77]. Its detection occurred primarily in the former Soviet Union, and then also in Thailand and northeast China affecting mostly children who were immunologic naive[64]. The emerging A(H1N1) subtype assumed a seasonal epidemiologic profile but did not cause the elimination of the previous seasonal A(H3N2) virus[64]. Unprecedented, these viruses co-circulated with a seasonal epidemiologic profile, alternating in dominant circulation until the emergence of the 2009 pandemic virus. The 2009 pandemic virus was a swine origin A(H1N1) subtype, product of a reassortment in which the TR-SIV acquired NA and M genes from a swine virus from the Eurasian lineage[78]. The 2009 pandemic is estimated to have caused over 200 000 deaths in the first 12 months, after which the A(H1N1)pdm09 assumed a seasonal epidemiologic profile that is maintained to date, and eventually eliminated the previous seasonal A(H1N1) subtype but not the seasonal A(H3N2)[79, 80].

Seasonal epidemics occur every year in temperate climates and in tropical regions more irregular outbreaks occur throughout the year[1, 52]. The estimated annual attack rate for annual epidemics and outbreaks ranges from 5%–10% in adults and 20%–30% in children and are estimated to result in about 3 to 5 million cases of severe illness and about 250 000 to 500 000 deaths [81]. Additionally to the direct costs associated with care providing, indirect costs of productivity losses and premature death and disability also contribute to the economic burden[82]. Influenza seasons differ in duration and severity, depending on the prevailing virus and the immunity of the population, and generally peak between December and March in the northern hemisphere[67]. Since the 2009 pandemics, A(H1N1)pdm09 and A(H3N2) co-circulate with influenza B virus alternating dominance and in the 2015/2016 season A(H1N1)pdm09 was the prevailing virus in circulation in Europe[67].

1.2.3 Molecular determinants of virulence and adaptation

Influenza A viruses differ in the presence of genetic features that translate into virulence phenotypes[35, 83]. Some of these features are characteristic of AIV and others are product of viral adaptation to mammalian hosts, acquired in the interspecies transmission.

The HA cleavage site is a known virulence determinant. The HA protein is synthetized as a precursor HA0 protein that is cleaved in HA1 and HA2 subunits, exposing the antigenic head and the transmembrane stalk, respectively, and permitting the virus to be infectious[35, 83]. In LPAI and in human and other mammalian influenza viruses, the cleavage site is monobasic and recognized by trypsin-like proteases present in the human and respiratory tract and in the avian respiratory and intestinal tracts[35, 83]. In HPAI, the cleavage site is multibasic, recognized by proteases with a broad distribution, and systemic infection can occur[35, 83]. The removal of the multibasic cleavage site is an attenuation factor in HPAI A(H5), although, in mammalian hosts, tissue tropism for viral replication is also determined by other factors and the introduction of a multibasic cleavage site may not solely promote systemic replication and enhance virulence[35, 83].

The receptor binding affinity of influenza A viruses is also a known virulence determinant. Human influenza viruses preferably bind α2,6-linked sialic acid receptors which are located in the upper respiratory tract (URT), as described in section *1.1.2 Life cycle and replication*. AIVs preferably bind α2,3-linked sialic acid receptors which are present in the intestinal tract of avian species. For the A(H1), A(H2) and A(H3) subtypes, several mutations have been identified to promote a receptor switch use from α2,3 to α2,6, specifically E190D and D225G for A(H1) and Q226L and G228S for A(H2) and A(H3), which means that the host range and cell tropism are changed from avian to human[35].

In the A(H5N1), the switch from avian to human receptor usage was accomplished *in vitro*, proving that it is possible for the virus to naturally acquire the capacity to infect the human URT[84]. An additional threat posed by non-human influenza virus is that some subtypes have been proven to acquire dual use receptor capacity, meaning that they can use both α2,3 to α2,6-linked sialic acid receptors, such as the 1918 A(H1N1) pandemic virus, the avian A(H5N1) and A(H7N9) and some strains of the 2009 pandemic virus, A(H1N1)pdm09. The infection of the URT of the human host can evolve to the LRT where α2,3-linked sialic acid receptors are found and result in a more

severe infection[85-87]. The removal of glycosylation sites in the HA of A(H1N1)pdm09 and A(H5N1), such as in position 158-160, have also been demonstrated to result in changes in virulence in animal models, because they play an important role in antigenicity and in receptor binding[88, 89].

In parallel to the binding of HA to sialic acid containing cell receptors, the cleavage of the sialic acids from the cell receptors by the NA may also contribute to virulence changes[90]. Specific mutations in the NA of AIV A(H7N7) have been shown to possibly increase the severity of the disease in mammalian hosts, by enhancing NA activity and consequently promoting the release of infectious virus particles and prevent their agglomeration[91, 92]. Deletions in the NA stalk have also been associated with the transmission of AIV in poultry, resulting in an enhanced virulence, although the mechanisms are not fully understood[93, 94].

The PB2 subunit of the viral polymerase protein is a known essential factor in the adaptation of avian virus to mammalian hosts. The translocation of vRNP to the nucleus requires PB2 and NP to bind to an adaptor protein, importin-α. There are several isoforms of importin-α and the interspecies transmission of AIV to the human host requires a switch from the preferential binding of PB2 to importin-α3 to importin-α7, which can be acquired by the accumulation of specific mutations in PB2, K627E and D701N or compensatory mutations G590S and Q591R[83, 95]. This is thought to be related to overcoming the constraint of viral replication at a body temperature of 33ºC in the human URT, for a virus originally adapted to replicate at a temperature of 41ºC in the intestinal tract of avian species[83]. The affinity of an avian PB2 binding to importin-α7 has a significant impact in the efficiency of the viral replication in the human host and therefore is a determinant factor in pathogenicity and virulence[35, 95].

The induction of cellular apoptosis by the PB1-F2 protein is known to increase virulence in influenza viruses and, additionally to the full length coding of the protein, mutation N66S has been found to further increase pathogenicity in mice by reducing the production of interferon which is part of the innate immune response[96, 97]. PB1-F2 protein has also been found to promote

inflammatory response of the lungs and more severe secondary bacterial infections, which is thought to be associated with the 1918 pandemic virulence[98-100].

The antagonizing of the antiviral response by limiting the production of interferon is a major determinant of virulence and is mostly accomplished by NS1 protein, by limiting IFN production, blocking IFN-induced gene expression or suppression of the effector molecule involved in the IFN signaling pathway[11, 101, 102]. NS1 is highly variable and several polymorphisms have been detected, but the core functions appear to be conserved in influenza subtypes[101]. A D92E substitution has been shown to particularly increase virulence in the A(H5N1) by enhancing viral replication in the presence of interferon, but the mechanisms remain unclear and further research is necessary the clarify the specific role of polymorphisms in the virulence phenotype[102, 103].

1.2.4 Risk assessment

The risk assessment for influenza is focused in determining the probability and the possible impact in public health of a specific event, such as a genetic or antigenic change, a reassortment event or the zoonotic transmission of animal non-human viruses, and is intended to adequate the prevention and control measures to minimize the impact. The main factors currently considered are the characteristics of the particular viral threat, the level of exposure and the context where it has occurred or will putatively occur. The viral threat is evaluated on the basis of its biology and epidemiology, the evaluation of the human exposure to the threat contemplates the population groups exposed or in risk of exposure and, particularly, their level of immunity to the threat, and the analyses of the context will mainly reflect the social, economic and political environment that could be relevant to the implementation and effectiveness of the prevention and control measures.

Seasonal influenza risk assessment is specifically intended to provide a description of the epidemiologic profile of the disease from the beginning of the season, to identify the populations affected and infer the impact in health services and to evaluate the vaccine effectiveness and viral susceptibility to antiviral drugs. The basic clinical, epidemiologic and virologic information is generally collected at a national level through the surveillance systems, when existent.

In the zoonotic transmission of influenza A viruses, risk assessment is critical in providing estimates of the pandemic potential of emerging strains and the possible impact in public health. In an historical perspective, in both mild and severe pandemics, a large proportion of the population is affected and the health, economic and social burdens are very high. Pandemics are not predictable and the early detection of the first cases is critical to the implementation of containment plans. Enhanced surveillance and risk assessment of emergent strains are essential to pandemic preparedness[104].

Currently, avian A(H5), A(H7) and A(H9) and swine A(H3) are considered potential pandemic threats. The HPAI A(H5N1) avian virus represents a high risk to human health because it is enzootic in poultry in several regions, genetically diverse because of the widespread circulation and consequently more likely to adapt, has a singular large host-range when compared to other AIV and is highly pathogenic in the human host[52]. The LPAI A(H7N9) also constitutes a high threat to human health because it is widely spread among poultry and because, since it is low pathogenic in the reservoir, it is more rarely diagnosed[52, 62].

The south and south-eastern Asia region are prone to the emergence of avian influenza viruses in the human host mostly because of the population density that characterize some areas of this regions and of the extended human-animal interface particularly in live bird markets. Considering the high risk AIV A(H5N1) and A(H7N9), and because most human infections have occurred following exposure to infected species or their secretions or excretions by direct or close contact, the most critical factor for reducing the incidence of human infection remains to decrease the possibility of human exposure and to limit their circulation in poultry[56, 57, 105].

The risk factors for a pandemic to occur are the emergence of a virus able to cause sustained transmission within the human host and the human insufficient immunity against this emerging virus. Given the globalization of short term travel and trade, a pandemic can be rapidly established with little initial public health response time. However, non-human animal viruses with HA unknown to the human immune system have occasionally infected the human host as described in the section *1.2.1 Reservoir, intermediate hosts and the human-animal interface*, but have not caused pandemics.

A major difference between the zoonotic viruses that have infected the human host and the pandemic viruses is the transmissibility phenotype[52]. All pandemic viruses isolated from the human host were airborne transmissible in the ferret and guinea-pig models, as are seasonal viruses and as opposed to avian viruses[57, 58]. The acquisition of this capacity is determinant to the pandemic potential and therefore a major factor to consider in the risk assessment of zoonotic transmitted strains. The extrapolation of animal transmission models to the human host must consider that some mammalian adaptation requirements may be host specific, but these have been considered adequate models and have been extensively used in influenza research because of their similarity to the human infection and clinical outcome[106-110].

Some phenotypic characteristics that are known determinants of virulence promote the capacity for airborne transmission in the mammalian host and are main factors to consider in the risk assessment of zoonotic transmitted strains. Primarily the viral attachment to the URT, mediated by the HA receptor biding preference and facilitated by the deletion of a glycosylation site and by pH stabilization and the improved replication at the URT temperature conditions, mediated essentially by PB2[35].

In previous pandemic viruses the receptor binding site of HA recognize α2,6-linked sialic acid receptors as do seasonal viruses and contrary to avian viruses which recognize α2,3-linked sialic acid receptors, and this feature has been proven determinant for airborne transmission in the process of adaptation to the human host[111-114]. The enhancement of replication capacity at 33ºC, the temperature of the mammalian URT, has also been found to promote airborne transmission[115-120]. Although not fully understood, the M gene has also been proposed to be involved in establishing high levels of airborne transmission in mammals, in association with NA, and possibly related to some aspects of virus morphology and shedding[121-123].

The *in vitro* acquisition of these features by induced mutagenesis has been suggested to be necessary but not always sufficient for an avian virus to be able to replicate efficiently and airborne transmit in the human host. However, *in vitro* reassortment experiments with the AIV A(H5), A(H7) and A(H9) subtypes and human viruses have proven that reassortants with this capacity can be naturally generated by classical reassortment, and serial passages in the ferret model have also been shown to promote adaptation of AIV towards acquiring this capacity[124-130]. In regards to

the A(H5) subtype, the lack of the glycosylation site has additionally been proposed to enhance the affinity to the α2,6-linked sialic acid receptors and an optimal pH threshold to induce membrane fusion have also been identified as crucial host range determinants[89, 131, 132]

Further research could clarify new aspects that impact the pandemic potential of zoonotic viruses and that could assist the risk assessment of emerging strains. The major public health concern is that influenza viruses evolve rapidly *in vivo*, acquire new traits and are introduced into new host species, enhancing the probability of encountering the human interface and evolving into efficient airborne transmission. The transmission efficiency would, ultimately, also be enhanced by host factors such as a more deficient immunologic competence to respond to the infection and prolonged virus shedding, and environmental factors that may contribute to a higher survival of the virus[52].

1.3 Viral fitness and evolution

1.3.1 Fitness

Viral fitness is a comprehensive concept. The Darwin's definition of fitness describes the ability of an organism to survive, to reproduce and to pass his genetic material to the descendants. It is now understood that the prevalence of genetic material over time in a viral population, can be affected by the replication capacity, transmissibility, ability to overcome the competition in co-infections and capacity to evade the hosts immune system[133]. These features are depended or consequential of one another and, for the purpose of this research, we will be referring to an overall viral fitness as an end result of a replicative, competitive and transmission fitness. The distribution and prevalence of particular genotypes overtime would be the end result of this overall viral fitness. The designation of robustness of a virus has been proposed to describe the preservation of the phenotype in the face of perturbations with overall virus fitness being the ultimate level of robustness[134].

The most robust viruses dominate the population and the main processes by which a set of genotypes become dominant are the positive and negative selection. The positive selection occurs when a set of genotypes is selected because a genetic feature translates into a phenotypic advantage in a given environment and, by contrast, the negative selection occurs when a set of

genotypes is eliminated or maintained at low frequency because of a negative impact of a genetic feature in the viral phenotype, in a given environment[135].

The more extreme epidemiologic situations, designated as selective sweeps, such as the interspecies transmission and major changes in geographic dispersion or virulence are generally associated with a positive natural selection. The phenotypic advantage may be conferred by a single genetic mutation, but generally it is given by the accumulation of genetic changes that constitute small fitness increases and which may be associated with each other at some level, such as compensatory mutations.

Animal viruses generally present a pattern of genetic evolution in which positive natural selection acts during the zoonotic transmission and the initial stage of transmission within the human host and, in late epidemic phases, the genetic evolution is generally characterized by negative purifying selection[136]. The dynamics of genetic evolution of the 2009 A(H1N1) pandemic virus is an example. The virus has a short evolutionary history in the human host but a long one in the swine population, suggesting that its emergence in pigs occurred years before the zoonotic transmission that lead to the 2009 pandemic[136, 137]. During the introduction of the virus in the human host, the genetic evolution was mainly driven by the need to adapt and characterized by a high ratio of non-synonymous to synonymous mutations (dN/dS), particularly in HA and the polymerase genes, consistent with a positive selection[136, 137]. In the post-pandemic period, since 2011, the ratio dN/dS decreased and the virus is characterized by little antigenic and genetic change until 2014, where the antigenic variation has begun to diversify probable due to some levels of immune driven selection[136, 137].

1.3.2 Population diversity of RNA viruses

The definition of quasispecies most commonly used in virology is a collection of related viral genomes subjected to a continuous process of genetic variation, competition and selection, that acts as a unit of selection[135].

RNA virus quasispecies populations are composed of mutants spectra which are naturally generated and facilitate the adaption of the virus to the environment[138]. The frequency of mutation has been proposed to be an evolutionary feature, in the sense that a more genetically

or phenotypically diverse population could better adapt to constraints in cell tropism, host range, selective pressure of host immune response or antivirals[135, 138, 139] Studies on lethal mutagenesis have however alerted to a threshold of mutation rate that can lead to the extinction of the viral population and that has been explored in antiviral research[140-142].

Interspecies transmission requires the most extreme challenging adaptation and a highly diverse quasispecies will increase the probability of an efficient transmission, where some mutants will be able to infect the new host, complete productive replication cycles and transmit between the members of the new host species[138].

1.3.3 Generation of diversity in Influenza A viruses

Genetic diversity is generated by accumulation of mutations and by reassortment events[1].

Influenza viral polymerase induces mutations at a rate of $1x10^{-4}$ and lacks proof-reading activity. Mutations can accumulate by natural selection and, if occurring in the antigenic sites of HA and NA, can cause antigenic *drifts* that facilitate the evasion from the host immune system and can compromise the immunity to circulating strains. Antigenic *drift* plays a crucial role in the evolution of seasonal influenza viruses[1, 143, 144].

The fact that the influenza viruses genome is segmented is a major evolutionary driver because it permits the reassortment between different strains that co-infect the same cell[6]. Studies based on full genome data sets show that reassortment is very frequent, but not all genomic constellations are viable probably because they can compromise the structural and functional compatibility of the viral proteins. When reassortment events place new antigenic proteins into a genomic background, an antigenic *shift* occurs which can severely compromise the host population immune defenses[1, 143, 144]. As opposed to reassortments within the antigenic gene segments HA and NA, that can result in changes in the antigenicity and have major impact in virus binding, entry, assembly, release or induction of the host's immune response, gene reassortments within the internal genes can have an impact in virulence and infectivity if they alter the virus capacity to replicate or to induce cellular apoptosis. Although the mechanisms by which specific genes promote virulence have not been fully clarified, the modification of genome transcription, replication and apoptosis kinetics are known determinants of virulence[145].

1.3.4 Gene segregation patterns in influenza A viruses

The gene segregation patterns are not completely understood and the mechanisms that drive them remain unclear, however, the implications in virulence and disease epidemiology have the potential to be severe.

The genetic analysis of swine influenza viruses shows that the genomic segments do not all share the same evolutionary history and that reassortment events are frequent[146].

The higher rates of reassortment are presented by HA and NA genes. The strongest genetic correlation occurs between the polymerase complex PB2, PB1 and PA genes, suggesting that these proteins tend to stay together, however, PB1 gene is reported to appear frequently in reassortment events among swine viruses[146].

In avian influenza viruses, it is also clear that the genomic segments have distinct evolutionary histories[147]. Within the internal genes, NS gene presents the higher reassortment rate and shows strong co-segregation pattern with antigenic proteins. This rate is immediately followed by the one of PB1 gene, which appears to be more frequently co-segregated with NA gene, followed by HA - NA together and at, a lower rate, with HA gene. The M gene also appears to present a high level of dependence to HA protein and the remaining internal genes present lower similar reassortment rates[147]. The strongest correlation is seen between PB2 and PA, even though a direct interaction between these proteins has not been reported[148].

In an historical perspective of the genetic evolution of human influenza virus, the acquisition of PB1 together with antigenic glycoproteins is also recurrent. In the reassortment events that lead to the emergence of the 1957 e 1968 pandemic viruses, the contemporary seasonal viruses acquired PB1 genomic segment from an avian virus together with antigenic glycoproteins[76, 149]. In 1957, the seasonal A(H1N1) descendent form the 1918 pandemic virus has acquired new HA, NA and PB1 genes from an avian A(H2N2) and, in 1968, this seasonal virus acquired new PB1 and HA segments from a reassortment event with an avian H3 virus, again causing a pandemic as previously referred[76, 149]. In both situations, viral fitness appears to have

benefitted from acquiring a PB1 genomic segment homologous to antigenic proteins rather than maintaining an homologous polymerase complex already adapted to the human host. A study by other authors shows that the *in vitro* incorporation of an avian PB1 gene into a background of human virus internal genes significantly increased mouse virulence[148]. Avian PB1 could have been implicated in the enhancement of replication and virulence of the pandemic reassortant viruses, that naturally overcame competition with seasonal influenza viruses[148].

A similar pattern was identified in 1947, where a reassortment event between seasonal viruses has altered the epidemiology of the infection to a near-pandemic geographic dispersion[76]. PB1, NA and M genes from a seasonal A(H1N1) virus with dominant circulation were combined with the remaining segments of a clade with minor circulation. The reassortant virus emerged causing high morbidity and mortality[76].

The profile of gene segregation in *in vivo* reassortment events appears to suggest that the functional compatibility between PB1 and antigenic glycoproteins can be determinant to virus fitness.

In pre-pandemic A(H5N1) seed virus research, viral growth has been proposed to improve with the inclusion of PB1 and M genes from a prototype immunogenic strain together with HA and NA[150]. In A(H1N1)pdm09 research, the neuraminidase and the polymerase activities have also been proposed to improve when PB1 is homologous to antigenic proteins[151].

Furthermore, a study on the retrospective evaluation of the internal genes content of seasonal high yield A(H3N2) vaccine viruses, used between 1995 and 2007, has revealed that PB1 was frequently incorporated with antigenic proteins[152]. However, in a subsequent study, an A(H3N2) reverse genetic seed virus with the same gene constellation presents a phenotype of reduced growth[153]. The apparently divergent results emphasize the knowledge gap and the need for further research.

1.4 Research context and objectives

The aim of this research was to contribute to clarify the role of PB1 as a molecular determinant of fitness and adaptation in influenza A viruses.

Different aspects of the generation of diversity and molecular epidemiology of influenza A viruses, such as the gene segregation patterns and its relation to viral fitness, suggested us that there may be a structural or functional dependence between PB1 and antigenic proteins, as described above.

In parallel, it has been found that a major limitation of vaccine seed virus production by reverse genetics is the growth deficit and that it could putatively be associated with a sub-optimal protein compatibility.

Vaccine research has been focused on inducing a broader and stronger immune response[20, 154]. The ultimate goal for influenza immunization would be the development of a universal vaccine which provided life-long or long-lasting immunity against a broad range of influenza subtypes, both seasonal and zoonotic. The different approaches would be to elicit antibody responses to conserved regions of HA or the M2 channel or to induce cross-protective T-cell responses against internal proteins like NP and M1[155]. Recent advances in this area include the identification of putative broadly neutralizing antibodies directed to the HA stem, the conserved stalk of HA[156].

However, influenza vaccination intents to prevent or mitigate the dispersion of cases and the development of severe complications in seasonal and pandemic epidemiologic contexts, and rapid improvements are urgent because current vaccines have serious limitations regarding the induced immune response and the manufacture process[20]. Over the past years, several improvements have been achieved with high-dose vaccines, the development of several adjuvants, the inclusion of a second influenza strain for B type virus and heterologous prime-boost regimens[20, 154, 157]. Also, the first recombinant vaccine and the first cell-culture-derived seasonal vaccine have been licensed[154, 157]. However, a major difficulty persists and is recurrent in the manufacture process. The time-frame in which it must occur, from the selection of the immunogenic strain to the large-scale production of the vaccine, must be as short as possible so that it does not compromise the timely and effective immunization of the population[154]. This has been extremely challenging for both seasonal and pandemic vaccines, essentially because of growth deficits of vaccine seeds viruses[158, 159]. The WHO and the National Institute of Allergy and Infectious Diseases, NIAID, have established as a goal for the

decade of vaccines 2011-2020, to develop and introduce new and improved technologies for vaccine manufacture that could accelerate the process and overcome this difficulty[159].

Despite the formulation, all vaccines require the production of seed viruses bearing the antigenic proteins of the immunogenic strain. The two most relevant characteristics of seed viruses are the growth kinetics and antigen yield[158, 160-162].

For seasonal influenza, seed viruses are commonly produced by classical reassortment, in which embryonated chicken eggs, traditionally, are co-infected with the circulating wild-type (wt) immunogenic strain and the apathogenic A/PuertoRico/08/34 (PR8). Progeny is screened for donor Hemagglutinin and Neuraminidase content and the strains with the higher growth rates are naturally selected by egg passage[150, 158].

Influenza seed virus production by reverse genetics is a more recent and precise technology, in which tissue cultures cells are transfected with cDNA clones from the apathogenic high growth PR8 backbone and the donor immunogenic antigenic proteins.

The advantages over classical reassortment are 1) seed viruses being generated with the desired gene constellation, obviating the screening for HA and NA content; 2) the possible insertion of cold adaptation and attenuation genetic features and removal of virulence determinants, such as the oligobasic cleavage site in HA, which is particularly relevant in the context of zoonotic transmission of avian viruses and pandemic preparedness; 3) rescue being performed directly from transfected cells, avoiding the use of eggs. The production time-frame for seed viruses is reduced when compared to the classical reassortment, and the process is more cost-effective[150, 154, 162].

Reverse genetic A(H1N1)pdm09 pandemic seed viruses have however presented low titers that have compromised the timely distribution of vaccines and the effective immunization and immunoprophylaxis of risk groups in the 2009 pandemics[163]. Current research on pre-pandemic reverse genetics vaccine seeds for A(H5N1) have also shown reduced growth kinetics and antigen yield[150, 164-166]. These findings have accentuated the need for improving the process of seed viruses production, as reported by the WHO and the NIAID[167].

Considering the putative impact of the compatibility between PB1 and antigenic proteins in viral fitness and the need for improved technologies in the production of vaccine seed virus, the main

questions we proposed to answer were if PB1 genomic segment is a molecular determinant of fitness and adaptation in influenza A virus and, particularly, if the functional compatibility between PB1 and antigenic proteins is a driver of the overall viral fitness and if it could be explored to improve seed virus production.

We propose to address these main research questions using the pandemic A(H1N1)pm09 virus, as a model. A(H1N1)pdm09 is a product of viral reassortment with an unprecedented genomic composition of segments from avian, swine and human seasonal viruses. These aspects make it an interesting model to evaluate a putative role of PB1 in the adaptation to new genomic backgrounds and new hosts. In this sense, we proposed to trace the genetic evolution of PB1 in the zoonotic transmission of influenza A(H1) viruses. Also, the 2009 pandemic vaccine presented severe growth deficits, as discussed above, and because the A(H1N1)pdm09 persisted in circulation with a seasonal epidemiologic profile, the demand for high yield A(H1N1)pdm09 vaccine seeds will be continuous and the need to adequate the immunogenic strain to the circulating viruses will be recurrent because of antigenic *drifts*.

The objectives of this research were defined as:

1) To evaluate the genetic evolution of PB1 in the zoonotic transmission of swine influenza virus and infer its putative contribution towards viral fitness and adaptation.

The general experimental approach to accomplishing this objective was to 1.1) select a study sample of PB1 nucleotide sequences from swine A(H1) virus that have infected the human host; 1.2) analyze phylogeny and mutation trends, and to; 1.3) search for putative markers for viral adaptation on the basis of viral molecular epidemiology, genomic location of the polymorphisms and amino-acid properties. The general experimental approach and major findings are described in chapters 2 and 3.

2) To determine if the functional or structural compatibility between PB1 and antigenic proteins is a molecular determinant of the overall virus fitness in the reverse genetics A(H1N1)pdm09 vaccine seed model.

The accomplishment of this objective followed the general approach of 2.1) determining the genetic profile of A(H1N1)pdm09 strains circulating in Portugal during the pandemic period and selecting a prototype immunogenic strain; 2.2) generating reassortant viruses with the genetic constellation of A(H1N1)pdm09 seed viruses prototypes bearing PB1 homologous and heterologous to antigenic proteins, and ; 2.3) evaluating *in vitro* viral growth and antigen yield. The general experimental approach and major findings are described in chapters 4 and 5.

Chapter 2

Genetic Evolution of PB1 in the zoonotic transmission of influenza A(H1) virus

Marta Gíria [a], Helena Rebelo de Andrade [ab]

[a] Host-Pathogen Interaction Unit, Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Lisboa, Portugal; [b] Departamento de Doenças Infeciosas, Instituto Nacional de Saúde Dr Ricardo Jorge IP, Lisboa, Portugal

Infect Genet Evol. 2014; 27: 234-243

2.1 Abstract

The epidemiology of human infection with swine-origin influenza A(H1) viruses suggests that the virus must adapt to replicate and transmit within the human host. PB1 is essential to the replication process. The objective of this study was to identify whether PB1 retains genetic traces of interspecies transmission and adaptation. We have found that the evolutionary history of PB1 is traceable. Lineage appears to be distinguished by amino acid changes between the conserved motifs of the viral polymerase, which can have major impact in PB1 protein folding, and by changes in the expression of the Mitochondrial Targeting Sequence and in the predicted helical region, that putatively affect induction of cellular apoptosis by PB1-F2. Furthermore, we found genomic markers that possibly relate to viral adaptation to new hosts and to new cellular environment and, additionally, to an enhanced compatibility with HA. We found no specific trend in the amino acid substitutions. Viral fitness appears to be favored by less reactive amino acids in some positions, while in others more reactive ones are fixed. Also, more flexible conformations appear associated with higher protein stability in general, although often more restrictive conformations appear to have favored protein folding and binding. Several aspects of PB1 mapping domains and the specific roles and interaction of PB1, PB1-F2 and N40 with each other and with other viral proteins and host cellular molecules remain unclear. Tracing the genetic evolution is critical to further understand the mechanisms by which PB1 affects vital fitness and adaptation. This analysis now permits putative adaptive related polymorphisms to be experimentally evaluated for phenotypic impact.

2.2 Introduction

The reservoir of aquatic birds sporadically introduces avian-origin viruses into mammalian hosts and interspecies transmission occurs between the swine and human populations. When crossing the species barrier, adaptation is mostly driven by natural selection and selective sweeps[136, 168]. Within the new host, adaptation mainly occurs by purifying selection[136, 168]. PB1, as an essential player in the replication process, undergoes genetic changes through the process of viral adaptation.

Human infections with swine influenza A(H1) virus resulting in un-sustained human-to-human transmission have been documented worldwide from 1970 to the 2009 pandemics[63, 64, 78]. These zoonotic viruses are designated as swine-origin influenza A(H1), SOIV A(H1). During their evolutionary history, reassortment events and interspecies transmission have placed PB1 into new viral genomic backgrounds and new host cellular environments. The extent to which PB1 retains genetic traces of interspecies transmission and adaptation is unknown. Questions arise as: Are there genetic markers that outline the lineage and host origin of PB1 segment, within a particular virus? Is the genetic evolution of PB1 towards viral adaptation traceable at the amino acid (aa) level? Is it possible to identify genetic markers for a) viral adaptation to new host cellular environments, and b) adaptation of PB1 to new genomic backgrounds, following reassortment events? In this study, we propose to trace the genetic evolution of PB1 of swine viruses that have infected the human host and infer its putative role in fitness and host adaptation, in view of the molecular epidemiology and evolutionary history of the viruses.

2.2.1 Role of PB1 genomic segment in viral fitness

The role of PB1 genomic segment is believed to be diversified and determinant in replication and induction of apoptosis. The segment encodes three proteins, PB1, PB1-F2 and N40. PB1 protein is responsible for the recognition of vRNA and initiation and elongation of cDNA and mRNA in viral transcription and replication. Interferences with the binding domains and the conserved motif of PB1 are specific targets for new antiviral research[169-171]. PB1-F2 is encoded in OFR+1 and exclusively found in infected cells. It has been associated with the induction of cellular apoptosis at a late stage of infection, which is supportive of viral replication and infectious particles release. Also, it is able to promote inflammation and it has been shown to up-regulate polymerase activity

by interacting with PB1 protein[172]. N40 is an N-terminal truncated form of PB1. It retains the ability to bind PB2 but is unable to bind PA. It has been reported as not essential to virus survival. However, polymerase activity is significantly reduced in the absence of N40, it and even further if PB1-F2 is also absent, although it seems not to be affected by the loss of PB1-F2 alone. On the other hand, the over-expression of N40 in the absence of PB1-F2 has been associated with a shift from transcription to replication and is thought to be regulated by the accumulation of the different RNA species[173]. Although new information regarding PB1, PB1-F2 and N40 is constantly being uncovered, several aspects of their specific roles and of their interaction with each other and with other viral proteins and host cellular molecules remain unclear. Namely, in the history of influenza virus classical reassortments, the acquisition of PB1 protein together with surface glicoproteins is a recurrent event and thereby thought to confer a biological advantage in natural selection by increasing the viral fitness[76, 150-152, 174]. Although it remains unclear as to how, the profile of gene segregation in reassortment events suggests that a functional compatibility between PB1 and HA enhances viral fitness.

2.2.2 Molecular epidemiology of human infections with swine-origin influenza A(H1) viruses

Sporadic cases and clusters of human infection with SOIV A(H1) have been identified in the past years, resulting in unsustained human to human transmission. From 1970-2000, over 50 cases of human infection with SOIV have been reported worldwide, mainly by A(H1N1) from the classic swine North American lineage. This was the predominant lineage isolated from pigs until the late nineties, with very little genetic change. It originates from avian A(H1N1) viruses, thought to be introduced in the swine population by interspecies transmission at the same time as they emerged in the human population in 1918 causing a pandemic. These SOIV A(H1) then share the genetic background of the avian A(H1N1) 1918 virus and the seasonal A(H1N1) descendents [63, 64]. In the swine population, multiple strains of Triple-Reassortant swine influenza A virus (TR-SIV) then emerged and became dominant in North America, as a result of a triple reassortment event between the classic swine North American, avian North American and seasonal A(H3N2) lineages[63]. The internal genes derive from swine (M, NS and NP), human (PB1) and avian viruses (PA, PB2) and this particular combination is designated as triple-reassortant internal

genes (TRIG) cassette. The TRIG cassette is very tolerant to antigenic glycoproteins and has been associated with other subtypes of swine virus (H3N2, H1N2). It is extremely stable and assumed to confer a selective advantage to the virus[175]. Since 2005, there have been 11 notifications of sporadic human infections with Triple-Reassortant swine-origin influenza A(H1), TR-SOIV A(H1).

A(H1N1)pdm09 then emerged in the human host, in 2009, and caused a pandemic. This emerging SOIV, although a A(H1N1)subtype, was genetically different from the previous swine A(H1N1) and TR-SIV A(H1) isolated from the human host. Its proposed origin was a reassortment event in which the backbone of TR-SIV acquired M and NA genomic segments from an Eurasian swine lineage[78]. The reassortment is presumed to have occurred in the swine population and to have suffered a long evolutionary process before the interspecies transmission to the human host. This period is phylogenetically estimated in up to 10 years and the introduction of the progeny virus occurred in single or multiple events of genetically closely related strains[136].

The epidemiology of human infections with swine influenza virus is dependent on environmental factors such as exposure, but it also reflects the genetic ability of the virus to infect the human host. Although sporadic infections have occurred, transmission among humans has been very limited until the 2009 pandemics, suggesting that the virus must adapt to replicate and transmit within the new host.

2.3 Methods

2.3.1 Study sample

This study was performed with PB1 nucleotide sequences accessed from the Influenza Virus Resource database at www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html and GISAID's EpiFlu™ database at www.gisaid.org.

The data set of SOIV included PB1 sequences from 8 isolates of SOIV A(H1) that have infected the human host and from 55 isolates of A(H1N1)pdm09 from the pandemic period, with worldwide distribution. The study sample of 8 SOIV A(H1) constitutes the entire set for which there are published PB1 sequences.

For the purpose of phylogeny and mutation trend analysis, the study sample additionally included 13 A(H1N1)pdm09 worldwide isolates from 2010/2011. For the putative adaptive mutation

analysis, SOIV and A(H1N1)pdm09 sequences were also evaluated against 19 seasonal A(H1N1) and 13 seasonal A(H3N2) isolates from 2009, with worldwide distribution, and their ancestors reference strains for the previous pandemics of 1918, 1957 and 1968 and reference strain for A(H1N1) reemergence in 1977. Strains designation and accession numbers are listed in table S2.1.

2.3.2 Phylogeny and mutation trend analysis

Nucleotide sequence alignment was performed by ClustalW, Mega5.2. Phylogeny was analyzed for the PB1 genomic segment exclusively in what concerns the PB1 protein coding region, since PB1-F2 protein is coded in different truncated forms that compromise the phylogenetic analysis. The phylogenetic tree of PB1 was constructed in PhyML, Seaview, using the model GTR+I selected by JModelTest software.

Within the branches of PB1 phylogeny, genetic analysis was performed for PB1 and PB1-F2 coding regions. For the purpose of this analysis, residues that were found exclusively in a particular lineage or host origin are indicated as putative genetic markers for that origin.

Viral RNA is used for non-coding functions such as packaging signals and promoter-related activities and, consequently, genetic mutations in the coding region of the protein may not be directly related to protein function. In our analysis of polymorphisms which have arisen and persisted in particular influenza virus lineages, however, residues considered putatively associated with viral adaptation on the basis of molecular epidemiology or amino acid substitution were identified as putative markers for adaptation.

2.4 Results and discussion

2.4.1 Phylogeny and evolutionary history

2.4.1.1 SOIV isolates of 1976 and 1988 are phylogenetically closely related to the 1918 pandemic reference strain and A(H1N1)seasonal virus

SOIV isolated in 1976 and 1988 are genetically divergent from Triple-Reassortment swine-origin influenza viruses (TR-SOIV) since they do not share the TRIG cassette. They have been isolated from humans previously to the emergence of TR-SOIV and present PB1 phylogenetically most closely related to the 1918 pandemic reference strain and seasonal A(H1N1) descendants (Figure

2.1). Based on the phylogenetic relation and the historical molecular epidemiology discussed above, we propose SOIV and seasonal A(H1N1) to have evolved in the swine and human hosts, respectively, from a common avian ancestor.

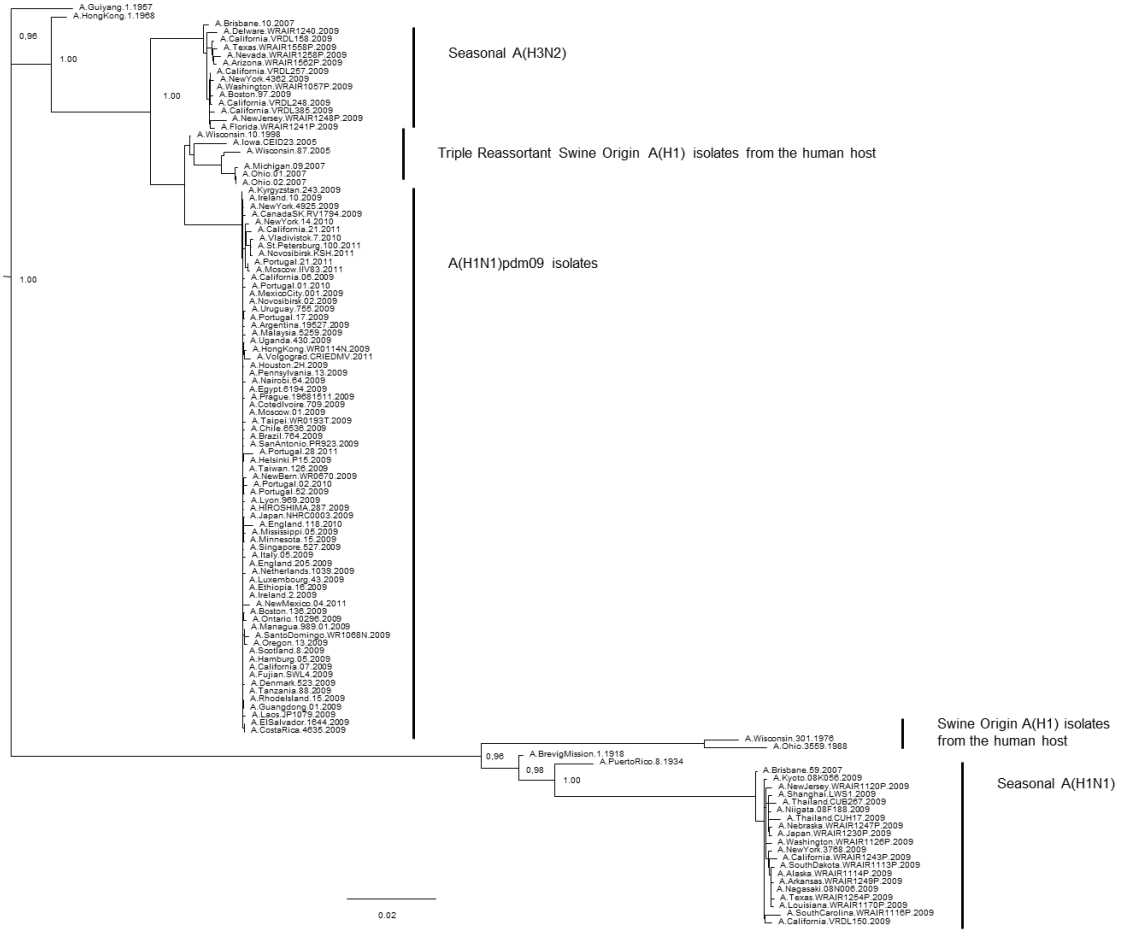


Figure 2.1: Phylogenetic tree of the PB1 coding region of swine-origin influenza viruses isolates from the human host. The phylogenetic tree was constructed in Seaview using the model GTR+I selected by JModelTest software. The accession numbers of the sequences used in this analysis are provided in Table S2.1.

Both SOIV strains code for a truncated non-functional form of 11aa PB1-F2 (Table 2.2), similar to the N-terminal 11aa of the 1918 pandemic virus and the human adapted progeny PR8 (Figure 2.2). Typically, all avian viruses code for a full length PB1-F2 and often this protein appears in a truncated form in swine and human adapted strains[176]. The adaptation of avian viruses to the mammalian host has in fact been proposed to include a truncation of PB1-F2, suggesting that it may not be crucial for effective transmission within these new hosts[172, 176]. We found that, in the human host, PR8 has evolved from its precedent 1918 pandemic virus to present a

Thryptophan-Stop change, resulting in a truncated functional form of 87aa (Figure 2.2). When A(H1N1) reemerged in 1977, however, PB1-F2 was coded in a shorter non-functional form of 57aa, resulting again from a Thryptophan-Stop change. Again based on phylogeny, genetics and the known historical molecular epidemiology, we presume this 1977 A(H1N1) to be the precursor for seasonal A(H1N1) viruses, that have then evolved in the human host to further accumulate genetic changes (Figure 2.2).

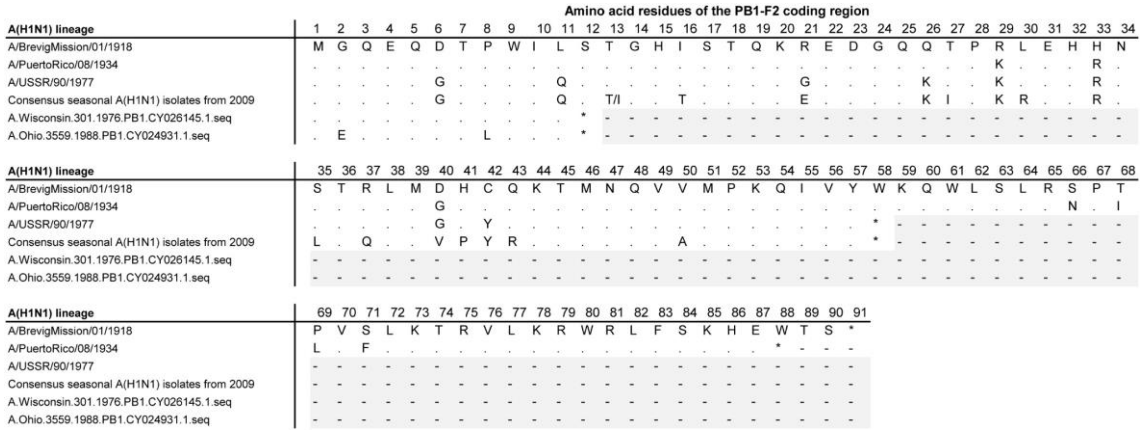


Figure 2.2: Amino acid alignment of PB1-F2 from A(H1N1) lineage viruses. Alignment of the PB1-F2 coding region was performed by ClustalW in Mega5.2 software. The accession numbers are provided in Table S2.1.

A particular Isoleucine/Leucine, I/L, pattern at positions 10/11 is found in PB1-F2 of 1976 and 1988 isolates, 1918 pandemic strain and PR8. (Table 2.2, Figure 2.2). Residue 10I appears to be a marker for classic North American A(H1N1) lineage, since it is also present in seasonal A(H1N1). Residue 11L, however, is exclusively present in these four strains and could refer to an early A(H1N1) residue, originated from the avian origin 1918 pandemic strain and transmitted to the swine and human hosts. In the human host, 11L is present in the 1934 PR8 although already substituted for Glutamine, Q, in the reemergent 1977 A(H1N1) subtype (Figure 2.2). The genomic position 11 of the PB1-F2 is not recognized, to date, as critical for protein function. However, since Glutamine is retained in 2009 seasonal A(H1N1), and is also present in seasonal A(H3N2) (Table 2.2), the function of the protein seems to have beneficiated from polarity over the hydrophobic and non-reactive properties of Leucine in this particular position.

The 1918 pandemic virus exclusively presents 33H, 66S, 69P and 90S and, together with PR8 further presents six residues characteristic of an early avian A(H1N1) lineage (positions 42, 43, 50, 65, 84, 86) (Table 2.2, Figure 2.2). Residue 66S is a known signature for virulence[172]. It would be interesting to further evaluate the association of the remaining residues with a phenotype of enhanced virulence. However, since all are located after the stop codon in 1976 and 1988 SOIV isolates, the aa are not encoded in the swine population and therefore could not have been implicated in the adaptation to the human host.

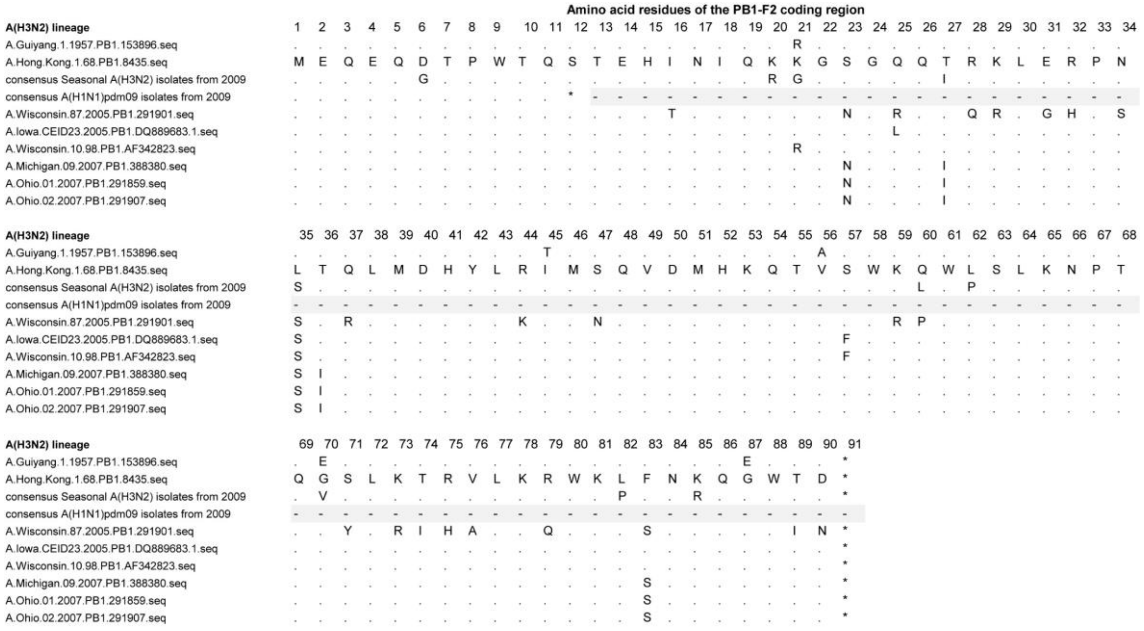


Figure 2.3: Amino acid alignment of PB1-F2 from A(H3N2) lineage viruses. Alignment of the PB1-F2 coding region was performed by ClustalW in Mega5.2 software. The accession numbers are provided in Table S2.1.

Table 2.1: Amino acid residues in PB1 protein of swine-origin influenza viruses isolates from the human host. Alignment of the PB1 coding region was performed by ClustalW in Mega5.2 software

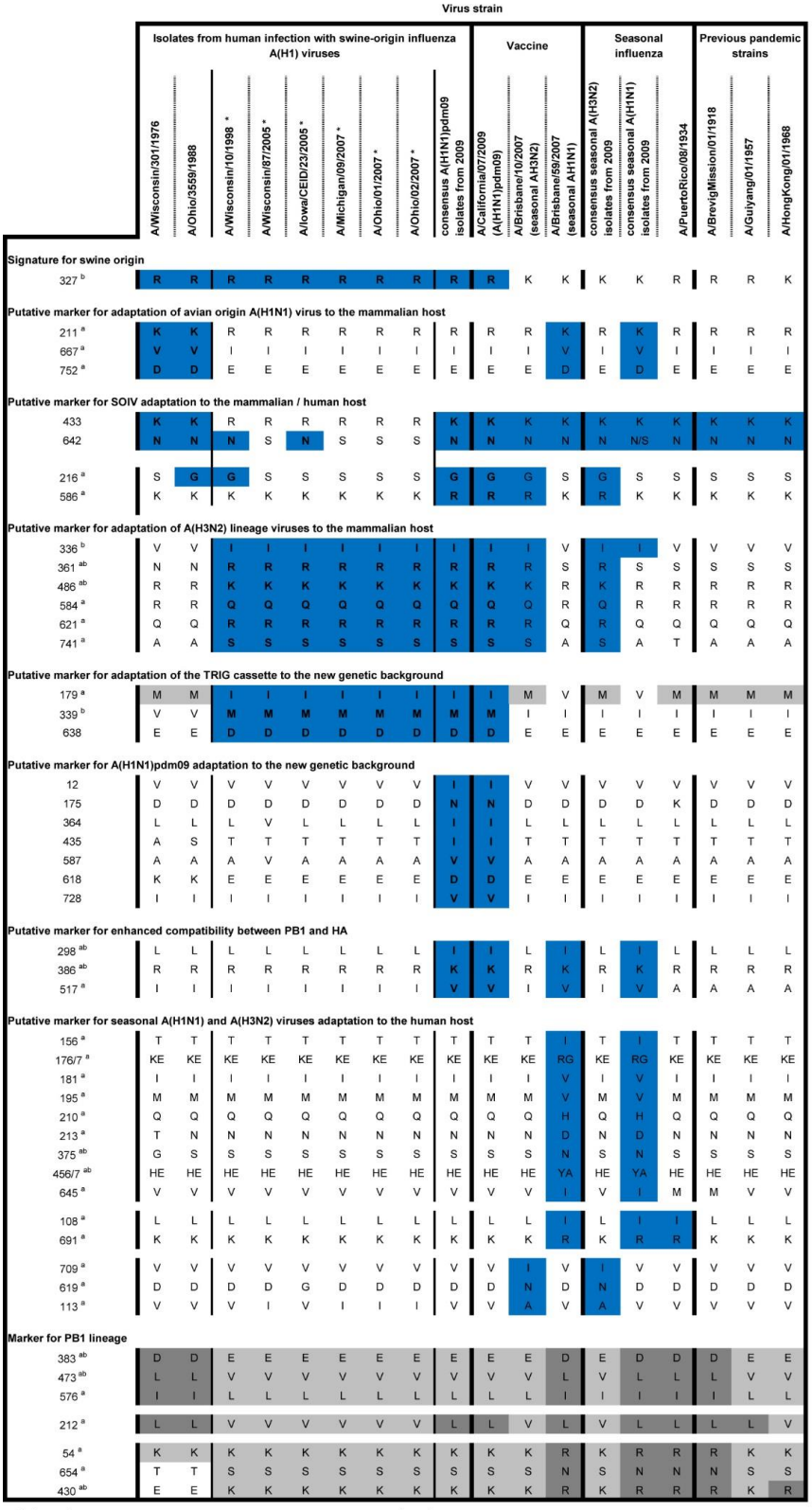
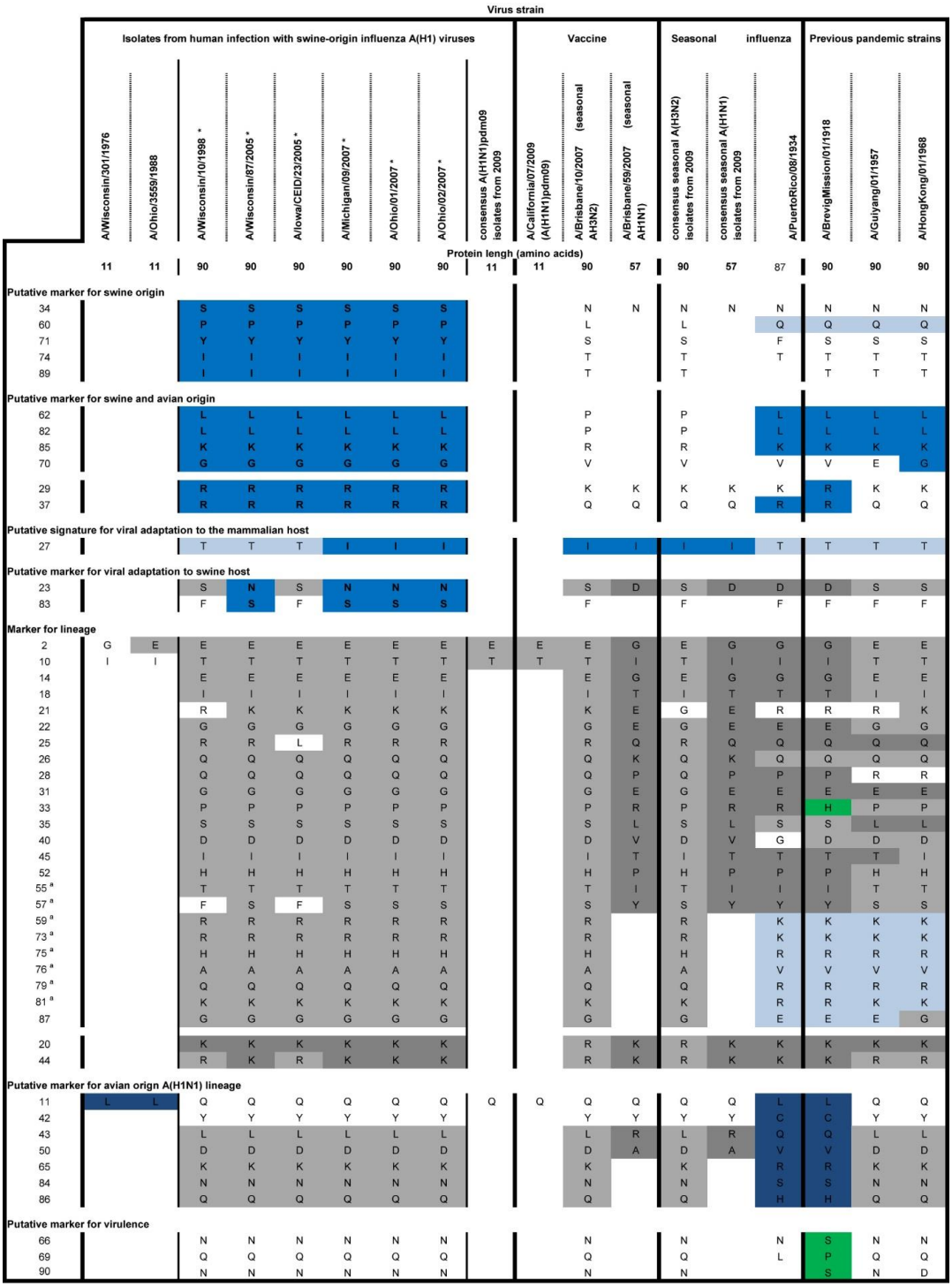


Table 2.2: Amino acid residues in PB1-F2 protein of swine-origin influenza viruses isolates from the human host. Alignment of the PB1 coding region was performed by ClustalW in Mega5.2 software.



color legend:
light blue fill: putative marker for avian origin
dark blue fill: putative marker for adaptation to mammalian host
green fill: putative marker for virulence
light grey fill: putative marker for seasonal A(H3N2) lineage
dark grey fill: putative marker for seasonal A(H1N1) lineage

legend:
a: residue located within the predicted helical region of the PB1-F2

2.4.1.2 TR-SOIV isolates are phylogenetically most closely related to the 1968 pandemic reference strain and A(H3N2)seasonal virus

As opposed to the 1976 and 1988 isolates, TR-SOIV and A(H1N1)pdm09 are phylogenetically more closely related to seasonal A(H3N2) and 1968 pandemic virus (Figure 2.1). Historically, in 1968, seasonal A(H2N2) virus acquired HA and PB1 from an A(H3) avian virus[64, 76]. The progeny A(H3N2) pandemic virus bearing an avian origin PB1 replaced the seasonal A(H2N2), as characteristically do all emerging virus. A(H3N2) then acquired a seasonal epidemiology and, unprecedented, remained in circulation after the reemergence of A(H1N1) in 1977. The PB1 segment of seasonal A(H3N2) was integrated into the TRIG cassette of TR-SOIV, in the swine population, and passed on to A(H1N1)pdm09. When A(H1N1)pdm09 emerged, the seasonal A(H3N2) was not eliminated. Both viruses co-circulate in the human host and share closely related PB1 proteins that originate from the same avian ancestor.

Within TR-SOIV isolates, the phylogenetic diversity of PB1 reflects the genetic evolution of the viruses in the swine host, from 1998–2007. A/Wisconsin/10/98 appears phylogenetically closest to the precursor of A(H1N1)pdm09, which is in agreement with A(H1N1)pdm09 emergence in the swine population around 1999 (Figure 2.1).

We have identified residue 327R, Arginine, as being present in all analyzed SOIV and TR-SOIV (Table 2.1). This is a swine and avian signature, as opposed to the human signature 327K, Lysine, [177]. Both are polar amino acids, positively charged and usually involved in salt bridges in active or binding sites, however, their structure is different. This position is located between the conserved motifs I and II of the viral polymerase[171]. Although putatively interfering with the folding of the protein, we consider that the presence of the swine signature Arginine in SOIV is probably not detrimental to virus fitness when infecting the human host, since the A(H1N1)pdm09 retains it.

All TR-SOIV analyzed code for a functional full length 90aa PB1-F2, as do seasonal A(H3N2) and 1968 pandemic virus (Table 2.2). It has been proposed that PB1-PB1-F2 interaction does not impact replication kinetics directly[176]. Nevertheless, A(H3N2) and TRIG cassette viruses characteristically have a high fitness and the mechanisms by which PB1-F2 may be contributing could involve mitochondrial targeting and the ability to cause cell death, the immunostimulatory properties or the interaction with the expression of N40[176]. TR-SOIV, seasonal A(H3N2) and

the reference strain for the 1968 pandemic present, however, amino acid divergences that we consider host related. Eleven residues were identified as putative swine origin markers (34, 60, 71, 74, and 89), six of which are concomitant putative avian markers (62, 82, 85, 29, 37 and 70) (Table 2.2). It is interesting to recognize that some residues are changed during adaptation to the human host and therefore no longer present in the descendent seasonal viruses (62, 82 85 and 70) and, in some circumstances, in the pandemic reference strain (29 and 37) (Table 2.2).
The A(H1N1)pdm09 codes for the exact same N-terminal amino acids as the TR-SOIV, but presents a Serine-Stop change resulting in a non-functional PB1-F2 of 11aa (Figure 2.3, Table 2.2). The reduced virulence of the 2009 pandemic virus has been attributed, in part, to the lack of a functional full length PB1-F2[172]. The mitochondrial localization of PB1-F2 is responsible for changes in morphology and membrane potential, ultimately associated with the initiation of the cellular intrinsic apoptotic pathway. PB1-F2 localizes to mitochondria through a Mitochondrial Targeting Sequence (MTS) comprised in 69-82 or 63-75aa[178]. The 11aa length protein does not encode this signal, will not localize to mitochondria and because of that may be less virulent. However, an A(H1N1)pdm09 virus genetic engineered to express PB1-F2, has been reported to maintain low level virulence [13]. It has even been suggested that PB1-F2 does not have an evolutionary importance to the virus and that it does not have major impact in virus fitness[179]. We can conclude that the impact of PB1-F2 in viral pathogenesis, virulence and fitness has to be strain dependent to some extent and that further research is crucial at this point.

2.4.1.3 Marker residues distinguish PB1 genomic segment from A(H1N1) and A(H3N2) lineages

We have identified marker residues that distinguish PB1 genomic segment from A(H1N1) and A(H3N2) lineages on the basis of protein structure and a phosphorylation site, and that we propose may contribute to the known differences in viral fitness. In PB1 coding region, 9 of the 35 markers for lineage (298, 361, 375, 383, 386, 430, 456/7 and 473) are located between the conserved motifs of the viral Polymerase (motif I 303-306, motif II 403-412, motif III 438-450, motif IV 474-484)[170, 171, 180] (Table 2.1). Both the composition and length of these inter-motifs sequences are critical to protein folding and could seriously alter the sites for binding and recognition. Regarding PB1-F2, a Serine-Leucine, substitution occurred in position 35 of 2009

seasonal A(H1N1). Based on our genetic analysis, and since the reemergent 1977 A(H1N1) presented 35S, we propose that the substitution has occurred between 1977-2009, during viral evolution in the human host. Residue 35S is a phosphorylation site associated the regulation of PB1-F2 interaction with PB1, which directly relates to virus titers[172]. The lack of this site has been reported to have a detrimental effect in fitness and we consider that it could have contributed to the phenotype of more reduced fitness in A(H1N1) seasonal viruses, when compared to A(H3N2). Particularly since within A(H3N2) seasonal lineage, the exact opposite event appears to have occurred (Table 2.2, Figure 2.3). The 1968 A(H3N2) avian origin pandemic virus lacking this phosphorylation site seems to have acquired it during its evolution in the human host and introduced it into the TRIG cassette, since it is now present in the 2009 seasonal isolates and TR-SOIV. Still within the markers for lineage in PB1-F2, 8 of the 32 (positions 55, 57, 59, 73, 75, 76, 79 and 81) are located within the predicted helical region (aa 55-85) which includes the MTS. The formation of an helical structure in this C-Terminus is essential because the positively charged amino acids have to be presented to the negatively charged mitochondrial membrane and other cellular compartments. From this interaction results the formation of pores that initiate the apoptotic process. Any changes in the folding of the protein or in the amino acids that are presented and interact with the membrane can affect the virus ability to control cellular apoptosis.

2.4.2 Genetic markers for viral adaptation

2.4.2.1 Residues found in PB1 and PB1-F2 proteins putatively associate with viral adaptation to the mammalian host, given the molecular epidemiology and the specific aa substitutions.

In SOIV isolates from 1976 and 1988, Lysine, K, Valine, V, and Aspartate, D, replaced Arginine, R, Isoleucine, I and Glutamate, E, in positions 211, 667 and 752 of PB1 (Table 2.1). Because the aa have similar properties, the consequences are not evident. However, seasonal A(H1N1) also present these substitutions and, although the swine and seasonal lineages appear to share a common avian ancestor, their genetic evolution paths occurred independently in different hosts. Also, 752 is located in the binding domain to PB2 (678-757aa)[170] and 667 is part of a broader ranged binding domain (600-757aa)[181], established prior to the more contemporary fine

mapping. We then propose that, within A(H1N1) subtype, avian origin viruses must have beneficiated from these substitutions during their adaptation to mammalian hosts.

In TR-SOIV, residues 433R, Arginine, and 642S, Serine, of the PB1 protein appear as a putative genetic marker and a dominant aa, respectively (Table 2.1). Arginine is frequently associated with active or binding sites because it is able to form multiple hydrogen bonds. The substitution of Arginine for Lysine is possible since both are positively charged. It is, however, very uncommon because Lysine is more limited in the number of hydrogen bonds it can establish. Although 433 has not been recognized as part of a binding domain, in the particular case of the TRIG cassette, a more flexible or less strong bonding capacity in this position appears to have favored A(H1N1)pdm09 fitness in the human host and putatively reflects an adaptation. Opposing, in position 642, a more reactive aa appears to have benefitted the fitness of A(H1N1)pdm09. Asparagine, N, is a polar aa, usually located at proteins surface in contact with the aqueous environment and predominantly involved in binding sites. This position is part of the broader ranged binding domain to PB2[181], referred above and although it may not be essential for binding, aa changes could alter the affinity. Given that NA and M were newly acquired by A(H1N1)pdm09, these could be compensatory mutations for adaptation to the new genetic background that promotes the structural interaction of PB1 and M proteins.

A different circumstance presents in positions 216 and 586. A(H1N1)pdm09 and seasonal A(H3N2) present Glycine, G, and Arginine, R, as opposed to Serine, S, and Lysine, K, respectively (Table 2.1). Molecular epidemiology suggests a phenotypic purpose to the substitutions. A Lysine-Arginine change is apparently neutral in terms of aa structure and function. Glycine, however, is not as reactive as Serine and is very particular because its structure allows the most flexible conformations. Under this circumstance, any substitution of Glycine, even for another small aa like Serine, most probably alters the structure of the protein.

In PB1-F2 coding region, residue 27T, Threonine, has been reported as phosphorylation site[172]. The phosphorylation status of PB1-F2 contributes to regulate the functionality of the protein in its direct interaction with PB1. TR-SIOV and seasonal A(H1N1) and A(H3N2), present 27I, Isoleucine (Table 2.2, Figure 2.3). Based on this genetics analysis, we propose Threonine to be a putative avian origin genetic marker that was introduced into the seasonal lineages A(H1N1) and A(H3N2), in 1918 and 1968 respectively. Since the reemergent 1977 A(H1N1) retained the

avian marker, we propose the substitution for Isoleucine to have occurred between 1977-2009. According to the known historical molecular epidemiology, within A(H3N2) lineage, the avian marker was probably introduced in the swine host in the late nineties, when the seasonal A(H3N2) PB1 genomic segment was acquired by the TRIG cassette, and, therefore, it was present in TR-SOIV 1998 and 2005 isolates. The T-I mutation must have, then, occurred independently in both hosts, strongly suggesting an adaptation of an avian origin virus to mammalian host cellular environments. We propose that this residue has probably been beneficial for avian viruses to infect mammalian hosts, but got lost in the subsequent circulation in swine and human populations. Another dramatic substitution occurs in position 83, where Phenylalanine, F, an hydrophobic aromatic aa usually involved with non-protein ligands is substituted by Serine, a small, polar, reactive aa (Table 2.2). In position 23, an apparently more neutral substitution occurred, in which Serine, S, was replaced by Asparagine, N (Table 2.2). Both are polar and reactive, usually exposed in the surface of the protein. Serine is, however, smaller in size and it could have altered the structure of the protein. According to our genetic analysis and to the known molecular epidemiology of the strains, both substitutions probably occurred in the swine population around 2005 and were fixed for being beneficial to the virus in the given environment.

2.4.2.2 Genomic positions putatively associate with the adaptation of PB1 from A(H3N2) lineage viruses to the mammalian host.

Position 584 has been previously reported as undergoing changes in selective pressure during host shifts from avian to human. Particularly, an Arginine, R, to Glutamine, Q, change is described as an adaptive mutation[182]. Here we found the avian marker Arginine present in pandemic strains of 1918, 1957 and 1968 (Table 2.1). Based on our genetic analysis and on the known molecular epidemiology data, we propose that seasonal A(H3N2) viruses have evolved to present R-Q change from the 1968 pandemic virus, subsequently incorporated in TR-SOIV and transferred to A(H1N1)pdm09. Since both 1976 and 1988 SOIV and seasonal A(H1N1) isolates retained the avian marker, we propose that this aa change was probably not essential for adaptation to the human or mammalian hosts, but instead was a genetic sweep that enhanced fitness. Both are polar aa, although Arginine is positively charged. It can be substituted by other positively charged aa but it has also been reported to tolerate changes to non-charged ones. In

fact, a more neutral side chain appears to have favored viral protein interactions of A(H3N2) viruses in the human cellular environment.

We propose that similar situations occur in 336, 361, 486, 621 and 741 (Table 2.1). None of the substitutions were neutral. Particularly, 336 and 361 are located between conserved motifs I and II of the viral Polymerase and 486 distances only 2 amino acids from the end of motif IV. All may have had an impact in protein folding. In position 336, the 2009 seasonal A(H1N1) isolates present the putative marker for adaptation Isoleucine, I. Isoleucine is similar to Valine, although more restrictive in the conformations it can adopt and which appears to be favored in the mammalian cell environment. In 486, Arginine is replaced by Lysine. Both are polar positively charged but this substitution is putatively prejudicial when interfering with structural sites, as discussed above. In position 361, a polar aa was replaced by an equally polar but positively charged one, and again in positions 621 and 741, less reactive or neutral amino acids were substituted for more reactive positively charged ones. In these cases, more reactive aa seem to benefit adaptation of avian TRIG cassette viruses to mammalian hosts, possibly by increasing the interaction of PB1 with host cellular proteins.

2.4.2.3 Residues putatively relate to the adaptation of PB1 to new genomic backgrounds on the basis of their molecular epidemiology.

Residues 179I, 339M and 638D are exclusive in PB1 of TRIG cassette viruses (Table 2.1). Position 638 is located in the broader range binding domain to PB2, referred above[181]. As opposed to Aspartate, D, in TRIG cassette viruses, both 1976 and 1988 SOIV, seasonal A(H1N1) and A(H3N2) and previous pandemic strains present Glutamate, E. These are polar amino acids, frequently exposed in the surface of the proteins and associated with active or binding sites. Aspartate confers a more rigid structure to the site, which is usually more advantageous for binding. In the TRIG cassette, PB1 and PB2 originate from seasonal A(H3N2) and avian lineages, respectively, and their interaction appears to have been favored by a more rigid structure.

In 179 and 339, Methionine, M, Isoleucine, I, and Valine, V, are present in specific patterns (Table 2.1). All are hydrophobic amino acids, not very reactive and not usually involved in protein function. Their roles are predominately associated with recognition or binding sites. Although Methionine is even more limited in the roles it can play in protein function because of its atomic

composition, Valine and Isoleucine are more restrictive in the conformations they can assume because of their structure. It is intriguing why these positions, apparently not corresponding to particular functions and not recognized as active or binding sites, present such defined patterns of aa. It is clear, however, that Methionine and Isoleucine must have been beneficial, to some extent, to PB1 activity, or were fixed as compensatory mutations for genetic changes in other viral proteins.

In the A(H1N1)pdm09, 12I, 175N, 364I, 435I, 587V, 618D and 728V are exclusive residues in PB1 coding region (Table 2.1). Position 12 marks the terminus of PB1 binding domain to PA[169]. In a previous study, this position has been assign a particular role in PB1-PA binding, and a Valine, V, to Aspartate, D, change was reported to decrease it by 40%[169]. Any change in the sequence of the domain could have a serious impact in the polymerase heterodimer formation. The V-D change previously reported replaces a hydrophobic amino acid for a polar negatively charged one. In our analysis, however, Valine is replaced by Isoleucine, I, which is similar in structure and function. Both are hydrophobic and usually located in protein cores, non-reactive and mostly involved in recognition sites. Because of their atomic composition, both are very restrictive regarding the conformations they can adopt. We propose that either Isoleucine appeared as a spontaneous mutation in the quasi-species of A(H1N1)pdm09 and prevailed since it was not detrimental to PB1-PA binding, or the substitution of has in fact a specific positive phenotypic translation, despite being so similar.

Two other unique residues in A(H1N1)pdm09, 618D, Aspartate, and 728V, Valine, are located , respectively, in the broader ranged and in the fine mapped PB1 binding regions to PB2[170, 181]. In 618, Aspartate replaces Glutamate that is present in the TR-SOIV isolates. The exact same substitution occurred in position 638 of TRIG cassette viruses, described above. A clear preference for a more rigid conformation seems to be occurring in the binding of PB1-PB2. In the A(H1N1)pdm09, the presence of 618D has probably further increased stability and fitness advantage over TR-SOIV when infecting the human host. Residue 728V distinguishes A(H1N1)pdm09 from all other SOIV analyzed, seasonal and previous pandemic strains, who present 728I, Isoleucine. Again, given the similarities between Valine and Isoleucine, it is unclear why the viruses present them in such a strict pattern. The remaining set of four exclusive signature

residues in A(H1N1)pdm09, 175N, 364I, 435I and 587V, is distributed along PB1 coding region. Their positions are close but not part of any mapped structural or functional sites and again aa changes do not seem to be consequential to a great extent. In 175N, Asparagine, and 435I, Isoleucine, a more neutral and less reactive amino acid appears to be favored in A(H1N1)pdm09, since all other SOIV, previous pandemic and seasonal strains present the polar more reactive amino acids Aspartate, D, and Threonine, T, respectively. In 364 and 587, Isoleucine, I, replaced Leucine, L, and Valine, V, replaced Alanine, A. All are hydrophobic aa, although Valine is slightly more than Alanine, and all are very similar in structure and function. We consider that all could reflect compensatory mutations in the adaptation of PB1 to the new genetic background or spontaneous mutations not detrimental to virus fitness.

2.4.2.4 Residues putatively relate to an enhanced compatibility between PB1 and HA, on the basis of their molecular epidemiology.

Residues in the A(H1N1)pdm09 were analyzed as putative markers for enhanced compatibility between PB1 and HA, in the human host. In position 386, an Arginine, R, is replaced by Lysine, K, in A(H1N1)pdm09 and 2009 seasonal A(H1N1) isolates (Table 2.1). The exact same substitution was discussed above as possibly neutral but also potentially prejudicial when interfering with structural sites. This particular position, 386, lays within the conserved polymerase motifs I and II where a less rigid aa composition appears to be favoring the folding of PB1 protein. The origin of PB1 in A(H1N1)pdm09 and seasonal A(H1N1) is distinct and, consequently, this mutation is not a product of a common path of evolution. These strains additionally present the same exclusive substitutions in positions 298 and 517. Although in 517 an apparently neutral substitution occurred, Isoleucine, I, to Valine, V, in position 298, which distances only 5 aa from the beginning of motif I, the substitution of Leucine, L, to Isoleucine, I, results in a more restricted conformation. The reassortment events in the history of PB1 genomic segment suggest that a genomic compatibility between HA and PB1 proteins enhances viral fitness, although the extent of the interaction remains unclear. We consider that a pattern where similar substitutions occur in PB1 proteins of different origins, near the conserved motifs where protein folding is critical and occurring in viruses that share the A(H1) subtype, suggests an adaptation of PB1 towards an enhanced compatibility with H1 in the human host.

2.5 Conclusions

We have found that PB1 does retain traces of viral interspecies transmission and adaptation. In the SOIV A(H1) that have infected the human host, the evolutionary history of PB1 and PB1-F2 proteins is traceable in term of lineage and host origin. PB1 from A(H1N1) and A(H3N2) lineages appear to be distinguished by aa changes in inter-motifs sequences of PB1 protein, that can have major impact in protein folding and heterotrimeric formation. Lineage distinction also appears to be related to the expression of the Mitochondrial Targeting Sequence in PB1-F2 and to genetic changes in the predicted helical region which can affect the virus ability to control apoptosis.

Moreover, specific genomic markers appear to be putatively related to viral adaptation. Substitutions between aa with different properties occur, in genomic positions that are critical to protein function or structure, and with molecular epidemiology data further supporting the assumption that they occur as part of the process of viral adaptation. We propose residues 27I in PB1-F2 and 336I, 361R, 486K and 584Q in PB1 as putative genetic markers for viral adaptation to the mammalian host. Also, residues 638D and 618D are proposed as putative genetic markers for viral adaptation to the new genomic background, in TRIG cassette viruses and in A(H1N1)pdm09, respectively. We additionally highlight residues 298I and 386K as putatively associated with the enhancement of PB1-HA compatibility. There is no apparent trend in the evolutionary process regarding aa reactivity or structure. The fitness of the viruses appears to have been favored by more neutral and less reactive aa in some positions, while in others more reactive ones were fixed. As regards to protein structure, more flexible conformations were also putatively associated with higher protein stability in general but, in some circumstances, more rigid or restrictive conformations appear to have favored the folding of the proteins and the binding to other polymerase subunits.

The evolutionary rates determined for internal proteins are typically lower than those for HA and NA, due to differences in the selective pressure exerted by the host immune system. The present analysis does, however, highlight that PB1 genomic segment of influenza A viruses evolves to divergent lineages and adapts to host and genetic backgrounds specificities. Tracing the genetic evolution is the basis to further understand the mechanisms by which PB1 affects vital fitness and, specifically, the identification of new residues and regions with putative roles in adaptation

can drive target research on antivirals. This analysis now permits putative adaptive related polymorphisms to be experimentally evaluated for phenotypic impact.

2.6 Acknowledgments

The authors would like to thank Vanessa Correia for providing expertise on phylogenetic and mutation trend analysis and both Vanessa Correia and Luis A Santos for having previously contributed with author Marta Gíria to the sequencing of the Portuguese strains. This study was sported by grants SFRH/BD/65211/2009 and PTDC/SAU-MIC/122780/2010 from the Fundação para a Ciência e Tecnologia, Portugal.

2.7 Statement of author's contribution

The phylogenetic and mutation trend analysis of the PB1 sequences, the research on evolutionary history of interspecies transmission of influenza virus and the manuscript preparation was performed by author Marta Gíria, The experimental work and manuscript were supervised by author Helena Rebelo de Andrade.

2.8 Supplementary data

Table S2.1: Accession numbers of the PB1 nucleotide sequences used in this study, accessed on Influenza Virus Resource database at www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html and GISAID's EpiFlu™ database at www.gisaid.org .

Strain		Acession number
Swine-origin Influenza virus	A/Wisconsin/301/1976	CY026145
	A/Ohio/3559/1988	CY024931.1
	A/Wisconsin/10/98	AF342823
	A/Wisconsin/87/2005	291901
	A/Iowa/CEID23/2005	DQ889683.1
	A/Michigan/09/2007	388380
	A/Ohio/01/2007	291859
	A/Ohio/02/2007	291907
Vaccine strains	A/California/07/2009	273607
	A/Brisbane/10/2007	165495
	A/Brisbane/59/2007	249528
Previous pandemic reference strains	A/BrevigMission/01/1918	99346
	A/Guiyang/01/1957	153896
	A/HongKong/01/1968	8433
Seasonal A(H3N2)	A/California/VRDL158/2009	267156
	A/California/VRDL248/2009	274112
	A/Boston/97/2009	316579
	A/California/VRDL385/2009	281813
	A/Delware/WRAIR1240/2009	275432
	A/Florida/WRAIR1241P/2009	325679
	A/Nevada/WRAIR1258P/2009	325751
	A/New_Jersey/WRAIR1248P/2009	325711
	A/New_York/4362/2009	309080
	A/Texas/WRAIR1558P/2009	325767
	A/Washington/WRAIR1057P/2009	275676
	A/California/VRDL257/2009	320427
	A/Arizona/WRAIR1562P/2009	325783
Pandemic A(H1N1) 2009 virus	A/Kyrgyzstan/243/2009	216955
	A/Ireland/10/2009	323794
	A/NewYork/4925/2009	246458
	A/CanadaSK/RV1794/2009	190993
	A/NewYork/14/2010	320110
	A/California/21/2011	330985
	A/Vladivistok/7/2010	306387
	A/St/Petersburg/100/2011	316433
	A/Novosibirsk/KSH/2011	350319
	A/Portugal/21/2011	EPI500762
	A/Moscow/IIV83/2011	324790
	A/California/06/2009	273599
	A/Portugal/01/2010	EPI500750
	A/MexicoCity/001/2009	217496
	A/Novosibirsk/02/2009	233326
	A/Uruguay/755/2009	194269
	A/Portugal/17/2009	EPI500694
	A/Argentina/19527/2009	267874
	A/Malaysia/5259/2009	244432
	A/Uganda/430/2009	217201
	A/HongKong/WR0114N/2009	277599
	A/Volgograd/CRIEDMV/2011	339334
	A/Houston/2H/2009	217469
	A/Pennsylvania/13/2009	227790
	A/Nairobi/64/2009	443112
	A/Egypt/6194/2009	216796
	A/Prague/19681511/2009	222977
	A/CotedIvoire/709/2009	191941
	A/Moscow/01/2009	182276
	A/Taipei/WR0193T/2009	277607
	A/Chile/6536/2009	211406
	A/Brazil/764/2009	194124
	A/SanAntonio/PR923/2009	190797
	A/Portugal/28/2011	EPI500767
	A/Helsinki/P15/2009	350288
	A/Taiwan/126/2009	213939
	A/NewBern/WR0670/2009	21537
	A/Portugal/02/2010	EPI500753
	A/Portugal/52/2009	EPI500725
	A/Lyon/969/2009	441677
	A/Hiroshima/287/2009	225950
	A/Japan/NHRC0003/2009	274991
	A/England/118/2010	294680
	A/Mississippi/05/2009	347027
	A/Minnesota/15/2009	227714
	A/Singapore/527/2009	277047
	A/Italy/05/2009	182420
	A/England/205/2009	348302
	A/Netherlands/1039/2009	269378
	A/Luxembourg/43/2009	184281
	A/Ethiopia/16/2009	189115
	A/Ireland/2/2009	323790
	A/NewMexico/04/2011	320075
	A/Boston/136/2009	267172
	A/Ontario/10296/2009	252312
	A/Managua/989/01/2009	249660
	A/SantoDomingo/WR1068N/2009	215460
	A/Oregon/13/2009	227783
	A/Scotland/8/2009	323766
	A/Hamburg/05/2009	350264
	A/Fujian/SWL4/2009	216591
	A/Denmark/523/2009	190196
	A/Tanzania/88/2009	194253
	A/Rhodelsland/15/2009	227801
	A/Guangdong/01/2009	276089
	A/Laos/JP1079/2009	226046
	A/ElSalvador/1644/2009	216835
	A/CostaRica/4635/2009	216773
Seasonal A(H1N1)	A/PuertoRico/08/1934	AF389116
	A/Kyoto/08K056/2009	243355
	A/NewJersey/WRAIR1120P/2009	275564
	A/Shanghai/LWS1/2009	213229
	A/Thailand/CUB267/2009	244661
	A/Niigata/08F188/2009	24344
	A/Thailand/CUH17/2009	230883
	A/Nebraska/WRAIR1247P/2009	275548
	A/Japan/WRAIR1230P/2009	275500
	A/Washington/WRAIR1126P/2009	275684
	A/NewYork/3768/2009	218013
	A/California/WRAIR1243P/2009	275416
	A/SouthDakota/WRAIR1113P/2009	275604
	A/Alaska/WRAIR1114P/2009	275353
	A/Arkansas/WRAIR1249P/2009	275400
	A/Nagasaki/08N006/2009	243395
	A/Texas/WRAIR1254P/2009	275628
	A/Louisiana/WRAIR1170P/2009	275540
	A/SouthCarolina/WRAIR1116P/2009	275596
	A/California/VRDL150/2009	277064

Chapter 3

Compatibility between influenza PB1 and HA: phenotypic evaluation of putative residues for genomic interaction and their relevance in the overall viral fitness

Marta Gíria[a], João Louro[a,b], Helena Rebelo de Andrade[a,b]

[a] Host-Pathogen Interaction Unit, Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Lisboa, Portugal; [b] Departamento de Doenças Infeciosas, Instituto Nacional de Saúde Dr Ricardo Jorge IP, Lisboa, Portugal

Manuscript in preparation

3.1 Abstract

Transmission of zoonotic swine-origin influenza viruses within the human host was very limited until the 2009 pandemics, suggesting that the virus must adapt. PB1 has been found to retain traces of interspecies transmission and adaptation and the functional compatibility between PB1 and antigenic proteins has been proposed to have an impact in viral fitness. Genomic mutations L298I, R386K and I517V in the PB1 have been previously found by the authors to have putatively enhanced the compatibility between PB1 and HA, in the A(H1N1)pdm09. In this study, the authors evaluated the phenotypic impact of acquiring these mutations in viral growth and infectivity, and inferred their putative role in the overall viral fitness. Acquiring this set of mutations appears to have resulted in a decrease in viral growth and infectivity of the A(H1N1)pdm09. However, L298I, R386K may impose some level of restriction in protein conformation and I517V appears to enlarge a span of complementary nucleotides possibly involved in the interaction between PB1 and HA at the genomic level. The authors propose these factors may be involved in establishing the biologic advantage *in vivo* by enhancing the compatibility between PB1 and HA, which is consistent with a selective packaging model for the assembling of infectious virus particles.

3.2 Short communication

Sporadic infections and small clusters of zoonotic swine influenza have been detected since 1970, but transmission within the human host has been very limited until the 2009 pandemics [51, 63, 65]. This suggests that, at a molecular level, the epidemiology of human infection with a swine virus is dependent on the virus capacity to adapt and efficiently infect and transmit within the human host.

In previous research, the authors have found that the PB1 gene segment from swine influenza viruses that have infected the human host (swine-origin influenza viruses-SOIV) retains traces of interspecies transmission and adaptation[183]. Specific genetic markers in PB1 protein, such 336I, 361R, 486K and 584Q, and in PB1-F2, such as 27I, have been proposed to putatively be associated with the adaptation of avian influenza viruses (AIV) to mammalian hosts. Additionally, residues 638D and 618D of the PB1 protein have been proposed to possibly be involved in the adaptation of this protein to the new genomic backgrounds after reassortment events, in Triple-reassortant SOIVs and in the 2009 pandemic virus, A(H1N1)pdm09, respectively[183]. Moreover, a specific set of mutations in the PB1 of A(H1N1)pdm09 and of the current seasonal A(H1N1), L298I, R386K and I/A517V, have been identified as possibly related to an enhanced compatibility between PB1 and the antigenic protein hemagglutinin (HA)[183]. The PB1 of the current seasonal A(H1N1) is human adapted, originating from an A(H1N1) avian ancestor introduced in the human population in 1918, eliminated in 1957 and reemergent in 1977 during a putative accidental laboratory release[64, 75]. On the other hand, PB1 from the A(H1N1)pdm09 has been acquired from the Triple-Reassortant A(H1N2) swine virus, which in turn has acquired it from the Triple-Reassortant swine A(H3N2)[64, 76]. The origin of PB1 in the Triple-Reassortant swine A(H3N2) was the human A(H3N2) lineage. In this human A(H3N2) ancestor, the PB1 segment was introduced by a reassortment event in which a seasonal A(H2N2) acquired an HA and PB1 from an avian virus[64, 76]. This analysis of the molecular epidemiology of PB1 from the current seasonal A(H1N1) and the A(H1N1)pdm09 viruses, shows that these segments have distinct phylogenetic histories and suggests that the set of mutations found may be related to their adaptation to an HA of the H1 subtype, in mammalian hosts[183].

The role of Polymerase Basic protein 1 (PB1) in viral fitness has been mostly related to the kinetics of viral replication, however, the functional compatibility between PB1 and antigenic proteins has

been suggested to have a role in the overall viral fitness within a host. In an historical perspective, the co-segregation of PB1 with antigenic proteins in *in vivo* reassortment events is recurrent[76]. Fitness appears to have been enhanced by acquiring PB1 homologous to antigenic proteins, over maintaining an homologous polymerase complex, even though the strongest genetic correlation was thought to occur between PB2, PB1 and PA[76, 147].

In the set of mutations L298I, R386K and I/A517V, previously identified as possibly related to an enhanced compatibility between PB1 and HA in the A(H1N1)pdm09 and the current seasonal A(H1N1), although in position 517 a more neutral substitution appears to have occurred, from Isoleucine (I) to Valine (V), the changes L298I, R386K are potentially more drastic. Position 386 is located between the polymerase conserved motifs I and II and the replacement of Arginine (R) by Lysine (K) is known to be possibly detrimental when occurring in structural sites. These are both polar, positively charged amino acids, but this replacement is uncommon because Lysine is more limited in the number of hydrogen bonds it can established. On the other hand, in position 298, the genetic change of Leucine (L) to Isoleucine (I) may result in a more restricted conformation and this position distances only a few nucleotides from the beginning of conserved motif I[183]. The structure of PB1 protein and the mapping of relevant domains for PB1 interaction with other viral proteins have not been fully resolved.

In this study, we proposed to evaluate the phenotypic expression of this set of mutations in the PB1 of A(H1N1)pdm09 and infer the putative role of these residues in the virus overall fitness and adaptation.

Our experimental strategy was to produce a plasmid-PB1 clone where these mutations have been reverted, by inducing K386R, V517I and I298L, to generate wild-type and mutated viral reassortants and to evaluate growth kinetics by determining the production of viral particles, hemagglutination capacity, infectivity and neuraminidase activity of the reassortants at 12 to 60 hours post-infection (hpi).

Madin Darby Canine Kidney cells (MDCK-SIAT1, European Collection of Cell Cultures - ECACC) were maintained in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco), 2 mM L-glutamin (Gibco), 1 x Non-Essential Amino Acids (NEAA, Gibco), 24 mM N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid (HEPES, Gibco), 1

mg/ml Geneticin® (Gibco). MDCK-SIAT1 cells were incubated at 37ºC. 293T cells were maintained in DMEM supplemented with 2,5 µg/ml Fungizone® (Gibco) and 1 x Penicillin-Streptomycin-Neomicin antibiotic mixture (PSN, Gibco). 293T cells were incubated at 37ºC with 5% CO2.

The prototype A(H1N1)pdm09 strain used in this study was previously selected by the authors, based on its genetic characterization and consensus towards the circulating strains during the pandemic period in Portugal[184]. The complete gene library for the prototype strain was assembled in pCIPolISapIT, an in-house dual promoter transcription plasmid kindly provided by the Centre for Disease Control and Prevention (CDC) – Atlanta.

The plasmid-PB1 clone bearing mutations K386R, V517I and I298L was produced by induced mutagenesis using a commercial kit (Stratagene), according to the manufacturer's protocol.

The recombinant wild type (wt) virus bearing wt PB1 and the reassortant virus bearing mutated PB1, were generated by plasmid-based reverse-genetics following an adaptation of the method by Hoffman et al[185].Transfection was performed in co-cultured 293T and MDCK-SIAT1 cells with Lipofectamine 2000 (Invitrogen) according to the manufacturers protocol and incubated at 37ºC with 5% CO2. At 4 hours post-transfection, the medium was changed and 2,5 µg/ml trypsin-TPCK (Worthington) was added. Viral stocks were assembled by inoculating MDCK-SIAT1 with the supernatants, 48 hours post-transfection, and incubating at 37ºC. First-passage stock titers were determined by Tissue Culture Infectious Dose (TCID50) following the method of Reed and Müench[186].

The recombinant wild-type and the mutated reassortant viruses were inoculated at the same multiplicity of infection (MOI), 0.001 TCID50/cell, in MDCK-SIAT1, in four independent assays. At 12, 24, 36, 48 and 60hpi supernatants were collected, centrifuged at 400g for 5 minutes, and stored at -30ºC. All experimental assays were performed in supernatants from each time point post-infection, from the four independent assays.

The Hemagglutination titer was determined in Guinea-Pig Red Blood Cells (GPRBC). Reassortants were tested in duplicates. Samples were two-fold serially diluted in DPBS to a 50µl final volume, mixed with 50ul GPRBC in a 0.75%v/v and incubated at 4ºC for 45 minutes.

The number of virus particles was determined by the number of vRNA copies of the Matrix (M) gene. Viral RNA was extracted using a commercial kit (Viral RNA mini Kit QIAGEN). The vRNA

copy number was determined by real time Reverse-Transcription Polymerase Chain Reaction (rtRT-PCR), following an adaptation of the protocols recommended by CDC and WHO[187, 188]. A TaqMan probe was design by the authors for the high specificity quantification of vRNA and is available upon request. Purified vRNAs were ran in duplicates and a set of serial dilutions of plasmid-cloned M gene were included in each run as a control for quantification.

The infectious virus titer was determined by Tissue Culture Infectious Dose (TCID50), using 10 replicates for each 10-fold dilution. TCID50 was calculated following the method by Reed and Müench[186]. Neuraminidase activity was determined by an in-house fluorescence MUNANA-based assay shared by Public Health England (PHE, London, UK)[189], as previously described[190]. A cutoff of 12500rfu was established and the dilution of each reassortant virus necessary to achieve NA activity was determined.

Statistical analysis was performed using the IBM Statistical Package for Social Sciences (SPSS) v21 software. Normality and homogeneity of variances were evaluated by Shapiro-Wilk and Lavene statistics, respectively. Mann-Whitney non-parametric test was used to evaluate association between variables. Values of $p<0.05$ were considered statistically significant.

When analyzing the growth kinetics of the wild-type recombinant A(H1N1)pdm09 and the PB1 mutated A(H1N1)pdm09 reassortant individually, major differences appear to occur in the dynamics of virus replication over time. The wild-type recombinant appears to have a slow growth kinetics as regards to production of virus particles from 12 to 24hpi, followed by a faster growth kinetics from 24 to 36hpi (Figure 3.1A). From 36 to 48hpi the growth kinetics actually decreases and from 48 to 60hpi higher values of production of virus particles are again evidenced, even surpassing the values achieved at 36hpi. The hemagglutination titer, neuraminidase activity and infectious virus titer show similar profiles (Figures 3.1B, 3.2A and 3.2B, respectively).

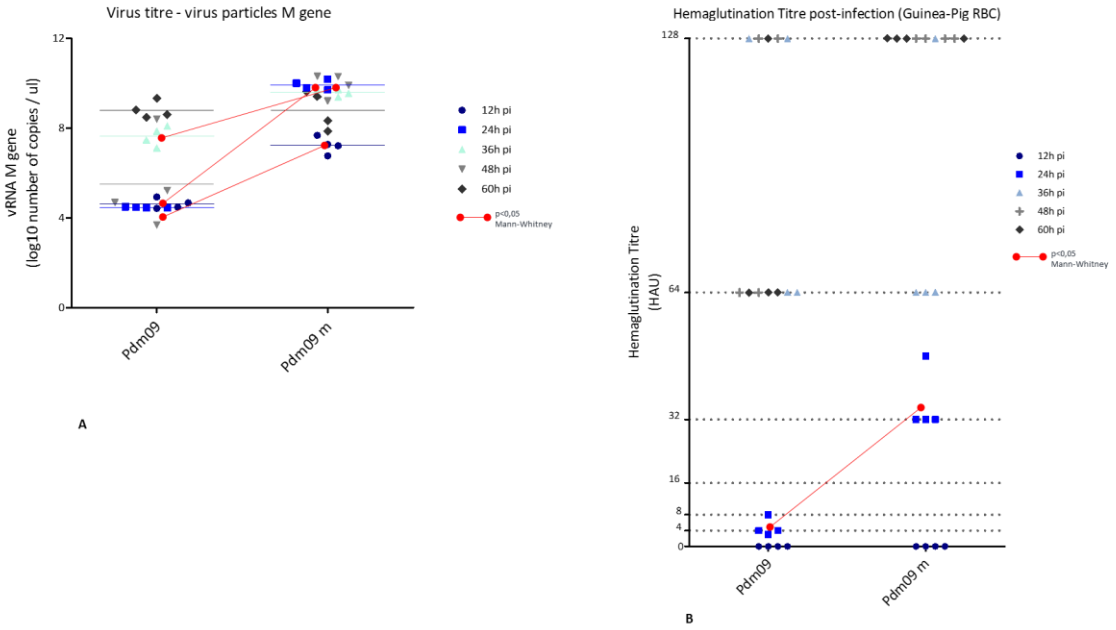


Figure 3.1: Evaluation of the viral growth and the hemagglutination capacity of a A(H1N1)pdm09 wild-type recombinant (pdm09) and a reverse genetic reassortant bearing a PB1 gene to which mutations K386R, V517I and I298L have been induced (pdm09 m). Figure 3.1A: Evaluation of the viral growth. The number of copies of vRNA from the M gene was determined by real time Reverse-Transcription Polymerase Chain Reaction (rtRT-PCR) at 12, 24, 36, 48 and 60h post-infection. Bars represent statistically significant differences between the reassortants, as determined by SPSS analysis. Figure 3.1B: Evaluation of the hemagglutination capacity. The hemagglutination titer was determined in Guinea-Pig Red-Blood Cells at 12, 24, 36, 48 and 60h post-infection. Bars represent statistically significant differences between the reassortants, as determined by SPSS analysis.

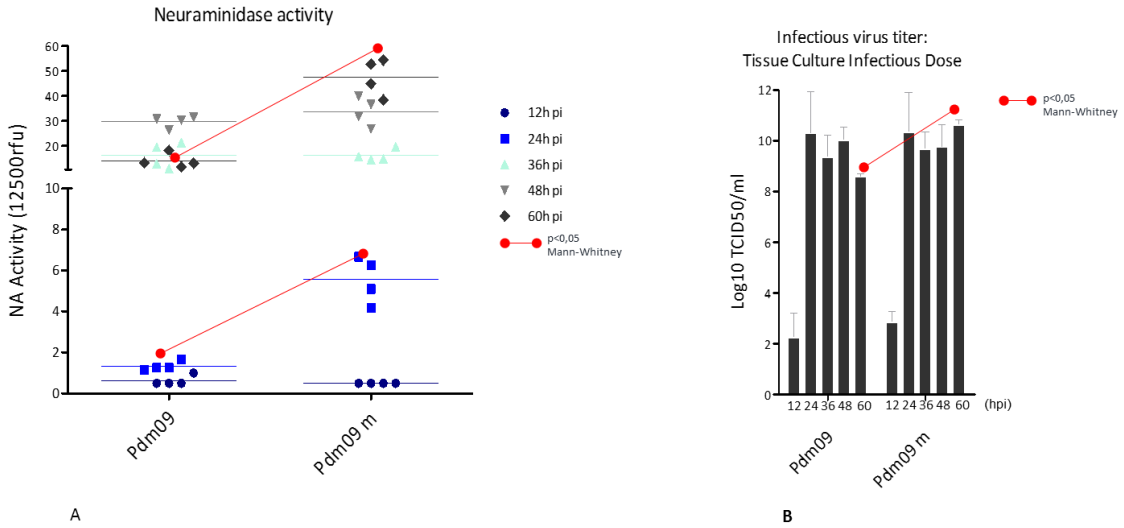


Figure 3.2: Evaluation of Neuraminidase activity and Infectious virus titer of an A(H1N1)pdm09 wild-type recombinant (pdm09) and a reverse genetic reassortant bearing a PB1 gene to which mutations K386R, V517I and I298L have been induced (pdm09 m). Figure 3.2A: Evaluation of Neuraminidase activity. Neuraminidase activity was determined by a fluorescence MUNANA-based assay, at 12, 24, 36, 48 and 60h post-infection. Bars represent statistically significant differences between the reassortants, as determined by SPSS analysis. Figure 3.2B: Evaluation of Infectious virus titer. The Tissue Culture Infectious Dose (TCID50) was determined following the method by Reed and Müench[186], at 12, 24, 36, 48 and 60h post-infection. Bars represent statistically significant differences between the reassortants, as determined by SPSS analysis.

In the PB1 mutated A(H1N1)pdm09 reassortant, the profile of growth kinetics is distinct. An accentuated increase in the production of virus particles occurs from 12 to 24hpi (Figure 3.1A), suggesting that the mutations that were induced translated into more dynamic first replication cycles. From 24 to 48hpi, growth kinetics appears to become more stationary and, from 48 to 60hpi, tends to decrease (Figure 3.1A) probably as a consequence of depletion of the media during the first replication cycles. This profile, where an accentuated increase in the growth kinetics occurs at early replication cycles, is similar when analyzing the hemagglutination titer, neuraminidase activity and infection virus titer (Figures 3.1B, 3.2A and 3.2B, respectively). In later time points post infection although the production of virus particles occurs with an apparent less accentuated kinetics (Figure 3.1A), the hemagglutination titer and the neuraminidase activity still show a gradual increase (Figures 3.1B and 3.2A, respectively).

As a consequence of these distinct viral growth dynamics, a significant increase in the production of virus particles is evidenced by the mutated A(H1N1)pdm09 reassortant at 12, 24 and 36hpi, when compared to the wild-type recombinant (Figure 3.1A). This increase is translated into a significantly higher hemagglutination capacity at 24hpi. In later time points, specifically from 48hpi to 60hpi, although the increase is not statistically significant, we can clearly see a tendency of the PB1 mutated A(H1N1)pdm09 reassortant to show a higher hemagglutination titer (Figure 3.1B). At 24hpi, the significant increase in the production of virus particles and hemagglutination titer is also translated into a significant increase in neuraminidase activity (Figure 3.2A). At 36hpi both reassortants present similar kinetics of viral growth but in later time points, although the production of virus particles is very similar, there is a clear tendency of the mutated A(H1N1)pdm09 reassortant to present higher hemagglutination capacity, neuraminidase activity and infectious virus titer (Figures 3.1A, 3.1B, 3.2A and 3.2B, respectively).

In summary, inducing mutations K386R, V517I and I298L, to revert the mutations acquired by A(H1N1)pdm09, clearly changed the dynamics of the growth kinetics. The mutated reassortant presents a significant increase in production of virus particles at 12, 24 and 36hpi, as evidenced in Figure 3.1A. This increase further translated into an increase in hemagglutination titer at 24hpi, as shown in Figure 3.1B, in neuraminidase activity at 24 and 60hpi, as shown in Figure 3.2A, and in infectious virus titer at 60hpi in MDCK-SIAT1, as seen in Figure 3.2B.

The apparent increase in the viral replicative fitness in the *in vitro* A(H1N1)pdm09 reassortant in which the mutations R386K, I517V and L298I have been reverted, raises the questions of if, and how, a growth impaired mutant dominated the competition *in vivo*.

From an evolutionary perspective, the viral quasispecies dynamics in a given host will result from the complex interplay between generation of diversity, competition between strains and the capacity to evade the host immune system. Fitness in a particular host environment is determined by this interplay, which means that a dominating strain may not necessary present the highest replicative capacity.

There is growing evidence that gene packaging in viral particles is selective, as opposed to random, which corroborates the concept of it not being driven uniquely by the highest replicative fitness[6, 191, 192].

We propose that the biological advantage of A(H1N1)pdm09 having acquired these mutations may be associated with the interaction between the PB1 and HA genome segments at the vRNA level, when the polymerase interacts independently with each segment for replication via their promoter replication signals, and/or at the vRNP level, when the 8 segments form the supramolecular structure that determines the selective packaging prior to budding. PB1 protein may have different affinities towards these interactions, when originating from different sources or if bearing distinct mutations in the specific regions where these interactions occur, and this could result in changes in the overall fitness of the viruses within the host and in their capacity to dominate the progeny.

In position 386 of the PB1 region of A(H1N1)pdm09, as discussed above, a more restricted conformation may have favored the structure and folding of the protein and a similar pattern was seen in position 298[183]. On the other hand, in residue 517, an apparently more neutral substitution has been shown to occur as regards to the properties of the amino acids involve and their putative impact in conformation and structure[183].

Mutation I517V is a result of a nucleotide change from guanine to adenine at position 1549. We have now found that this mutation occurs in a highly complementary region between PB1 and HA and this specific nucleotide change further enhances this complementarity by creating a span of 8 complementary nucleotides out of 10.

In recent research by other authors, PB1 has been shown to co-segregate with NA in the A(H3N2) subtype and small highly complementary regions of 7 to 22 nucleotides have been described to putatively promote an enhanced PB1-NA interaction, at the vRNA level[153, 191]. It has not yet been clarified if these segments interact at this level, or in this genomic position, but these findings were enlightening and suggestive of a possible mechanism for co-segregation. However, despite the fact that A(H3N2) strains bearing PB1 homologous to HA and NA appear to dominate the progeny in classical reassortment events, in reverse genetics vaccine seed prototypes, where the antigenic proteins of the A(H3N2) source are deliberately introduced in the A/PuertoRico/08/34 backbone, it has been shown to reduce virus growth[153, 193].

Further research is necessary to understand the interactions at the vRNA and vRNP levels and is critical to establish the biology of virus packaging and its contribution to the genome segregation patterns and to the overall viral fitness.

In the A(H1N1)pdm09, we propose that mutations K386R, I298L and V517I may have persisted in the progeny because the conformation restrictions that they appear to impose bring some biological advantage to the interaction between PB1 and HA and because the span of complementary nucleotides is enlarged. These findings may be extremely relevant to the identification and mapping of genomic regions where the interaction between these segments occur, and to further clarify their relevance in establishing the genomic composition of circulating viruses. We consider it would now be essential to distinguish the individual role of these mutations in the viral phenotype.

We propose that the mechanisms that drive virus packaging and genome segregation patterns are very determinant aspects in the selection of dominant strains in the quasispecies viral population, in parallel to the replication kinetics and the capacity to evade the host's immune response. This concept could contribute to explain how and why genetic variants that dominate the progeny *in vivo* can be found to be growth impaired *in vitro*.

3.3 Conflict of interest statement

The authors declare no conflict of interest.

3.4 Acknowledgments

The authors would like to thank the Centre for Disease Control and Prevention (CDC) – Atlanta, for the training on Reverse genetics for the construction of viral reassortants, which was applied for the design of the experimental strategy. Also, the authors would like to thank the CDC – Atlanta for kindly providing the pCIPolISapIT transcription plasmid.

This research has been funded by the Fundação para a Ciência e Tecnologia grants PTDC/SAU-MIC/122780/2010 and SFRH/BD/65211/2009.

Chapter 4

Genomic signatures and antiviral drug susceptibility profile of A(H1N1)pdm09

Marta Gíria[a], Helena Rebelo de Andrade[a,b], Luis Santos[a,b], Vanessa Coreia[b], Sónia Pedro[c], Madalena Almeida Santos[d]

[a] Host-Pathogen Interaction Unit, Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Lisboa, Portugal; [b] Departamento de Doenças Infeciosas, Instituto Nacional de Saúde Dr Ricardo Jorge IP, Lisboa, Portugal; [c] Departamento de Genética, Instituto Nacional de Saúde Dr Ricardo Jorge IP, Lisboa, Portugal [d] Serviço de Patologia Clínica/Biologia Molecular, Hospital de Curry Cabral, Lisboa, Portugal

J Clin Virol. 2012 Feb;53(2):140-4

4.1 Abstract

Background: Genetic changes in influenza surface and internal genes can alter viral fitness and virulence. Mutation trend analysis and antiviral drug susceptibility profiling of A(H1N1)pdm09 viruses is essential for risk assessment of emergent strains and disease management. Objective: To profile genomic signatures and antiviral drug resistance of A(H1N1)pdm09 viruses and to discuss the potential role of mutated residues in human host adaptation and virulence. Study design: A(H1N1)pdm09 viruses circulating in Portugal during pandemic and post-pandemic periods and 2009/2010 season. Viruses were isolated in MDCK-SIAT1 cell culture and subjected to mutation analysis of surface and internal proteins, and to antiviral drug susceptibility profiling. Results: The A(H1N1)pdm09 strains circulating during the epidemic period in Portugal were resistant to amantadine. The majority of the strains were found to be susceptible to oseltamivir and zanamivir, with five outliers to neuraminidase inhibitors (NAIs) identified. Specific mutation patterns were detected within the functional domains of internal proteins PB2, PB1, PA, NP, NS1, M1 and NS2/NEP, which were common to all isolates and also some cluster-specific. Discussion: Modification of viral genome transcription, replication and apoptosis kinetics, changes in antigenicity and antiviral drug susceptibility are known determinants of virulence. We report several point mutations with putative roles in viral fitness and virulence, and discuss their potential to result in more virulent phenotypes. Monitoring of specific mutations and genetic patterns in influenza viral genes is essential for risk assessing emergent strains, disease epidemiology and public health implications.

4.2 Background

The emergence of virulent influenza phenotypes is a consequence of genetic changes altering the function of individual proteins or their functional compatibility.

Genetic changes in the viral surface proteins can impact on virus binding, entry, assembly, release, induction of the host's immune response and antiviral drug resistance.

Within the internal proteins, genetic changes can impact on virulence and infectivity if they alter the ability of the virus to replicate or induce cellular apoptosis[145]. The viral replication process is multigenic and interactive, involving the replication complex and the structural and functionally associated M1 and NS2/NEP proteins[145, 148].

The gene segments coding for the proteins involved in replication, code for additional proteins associated with the induction of cellular apoptosis during viral infection. The (+1)ORF of PB1 encodes PB1-F2 with pro-apoptotic function in the host cells and the NS gene encodes NS1 protein associated with the activation of anti-apoptotic mechanisms, permitting viral replication to occur[11-13]. The cassette of internal genes therefore controls the major processes by which viral fitness is determined: replication and apoptosis. An increased replication rate or stimulation of anti-apoptotic mechanisms, produces higher viral loads that can overcome host immune response, enhance infectivity and result in a more severe and transmissible disease. The most severe epidemiological situations caused by seasonal influenza viruses occurred in 1947 and 1951 as a consequence of genetic changes within internal proteins involved in replication of A(H1N1) viruses. The strains spread worldwide and caused exceptionally high mortality[76].

Monitoring evolutionary trends in the genes encoding influenza virus internal proteins, and profiling antigenic proteins and antiviral drug susceptibility, are therefore essential to risk assessing influenza virulence, disease epidemiology and public health implications.

4.3 Objectives

With the aim of risk assessing emergent strains, the main objectives of the study were to profile genomic signatures and antiviral drug susceptibility of A(H1N1)pdm09 viruses and discuss the potential putative role of mutated residues in human host adaptation and virulence.

4.4 Study design

A total of 110 specimens from laboratory-confirmed cases of pandemic influenza were collected in Portugal from July 2009 to January 2010. Strains were isolated in MDCK-SIAT1 cell culture. Antigenic surface glycoproteins and M1 and M2 proteins were genetically characterized and antiviral drug susceptibility was profiled. Genetic characterization of internal proteins was performed on viruses isolated from subsets of pregnant women, outliers to NAIs and a 14 years old deceased patient with no risk factors. Cycle sequencing was performed based on adaptation of a protocol from CDC, recommended by WHO[194]. Genetic mutation analysis and phylogenetic analysis were performed with LasergeneV.4.05- DNASTAR and MEGA4.0, with comparison to sequences of the A(H1N1)pdm09 2010/2011 vaccine strain (A/California/7/2009) and published reference A(H1N1)pdm09 strains circulating worldwide within pandemic and post-pandemic periods. Phenotypic evaluation of antiviral drug susceptibility was performed by fluorescent assay with MUNANA substrate as previously described[195]. Minor and major phenotypic outliers were identified through the establishment of lower and upper IC50 cut offs values[195]. Genotypic evaluation of antiviral drug susceptibility to NAIs and to amantadine was performed by analyzing NA and M2 protein sequences, respectively. In the M2 sequence the 5 well-defined molecular markers of resistance to amantadine were analyzed: L26F/I; V27A/D; A30T; S31N and G34E.

4.5 Results

Virus isolation was performed for all 110 A(H1N1)pdm09 cases. Of these, 37 were genetically characterized for antigenic glycoproteins and for M protein. Antiviral drug susceptibility profile was phenotypically evaluated for 103 of the 110 isolates. Ninety-eight were further sequenced for genotypic evaluation of NAIs susceptibility and 96 for amantadine susceptibility. Genetic characterization of internal proteins PB2, PB1, PB1-F2, PA, NP, NS1 and NS2/NEP was performed for 13 of the 37 isolates, including viruses isolated from cases in the subsets described above.

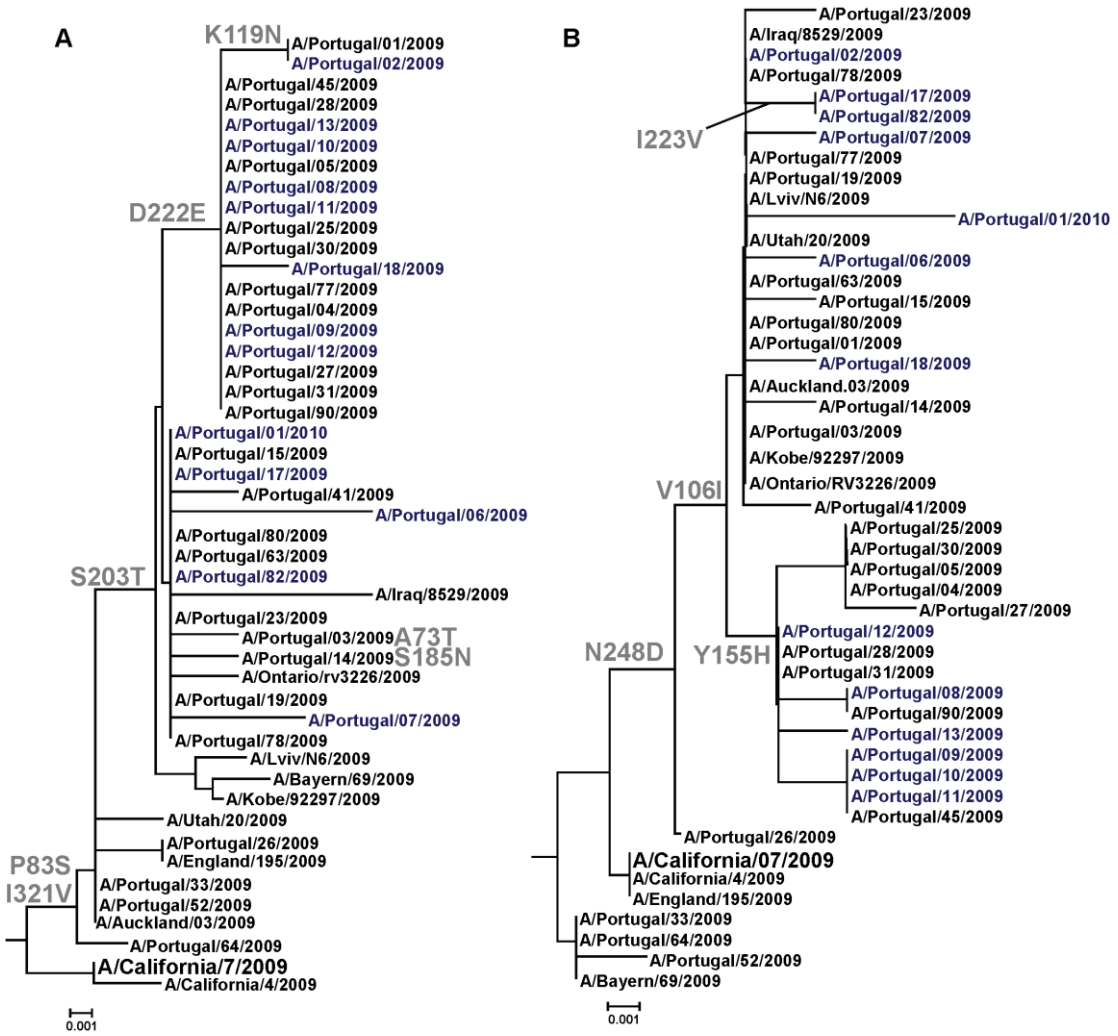


Figure 4.1: Phylogenetic analysis of the amino acid sequences of HA1 subunit of hemagglutinin (A) and neuraminidase (B) genes of Influenza A(H1N1)pdm09 strains circulating in Portugal.

Within the HA1 domain of the HA, all 37 strains were shown to contain the mutations P83S and I321V (Figure 4.1A). Thirty-three of the 37 strains had the substitution S203T, and 19 also contained D222E. Additional point mutations were observed, including A73T and S185N in individual strains, which are located in the putative antigenic sites Cb and Sb, respectively. The substitution K119N which results in the creation of an N–X–S motif, was observed in 2 strains. In the NA gene, 34 of the 37 strains analyzed contained N248D and V106I substitutions (Figure 4.1B). Fifteen strains had Y155H and two contained I223V in the NA gene. The genetic

characterization of MP of 37 strains identified M128L, E201D and M203I substitutions in individual isolates compared to the A(H1N1)pdm09 2010/2011 vaccine strain MP sequence, located within the M1 coding region.

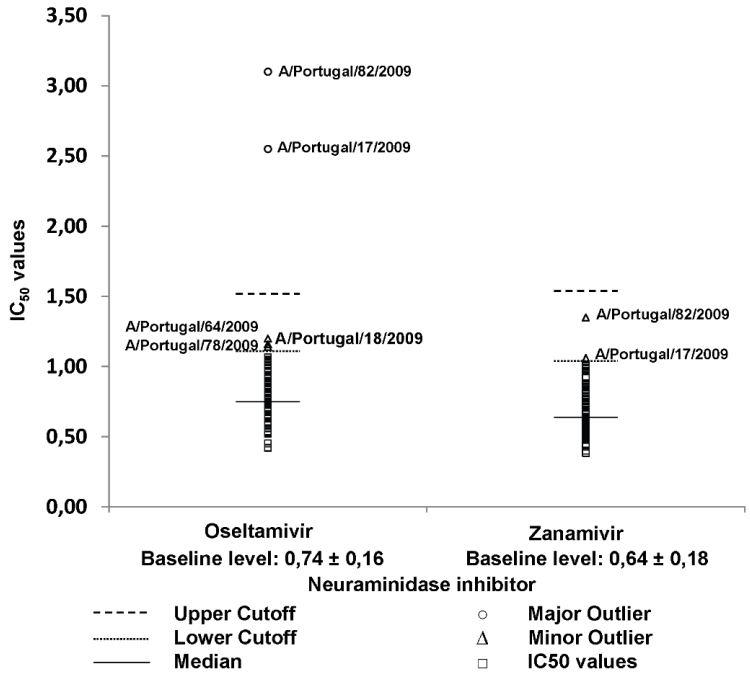


Figure 4.2: Phenotypic evaluation of NAIs susceptibility. IC50 values obtained by fluorescence assay for influenza A(H1N1)2009.

Phenotypic evaluation of NAIs susceptibility revealed three minor and two major outliers to oseltamivir, as shown in Figure 4.2. The three minor outliers exhibited reduced susceptibility of approximately two-fold and the two major outliers of approximately 3- and 4-fold, compared to baseline level. The two major outliers were also found to be minor outliers to zanamivir (MjOOsel/MnOZana), with reduced susceptibility of approximately two-fold compared to baseline level. In the NA sequence of these two MjOOselt/MnOZana strains, the mutation I223V was observed. Genotypic evaluation of susceptibility to amantadine revealed that all 96 strains analyzed had an asparagine (N) at position 31 in the M2 coding region. The genetic profile of the internal proteins of 13 A(H1N1)pdm09 isolates analyzed revealed substitutions P224S in PA, V100I and L122Q in NP and I123V in NS1, which were common to all 13 isolates (highlighted in bold and #, Figure 4.1), as shown in Table 4.1. PB1-F2 was present in the truncated non-

functional form of 11 amino acids. Viruses isolated from the subset of pregnant women, were observed to have R54K substitution in PB2 in 6 cases, 5 of which also contained D347N in PA. Of these, 4 also contained S384L in PB1 and a small cluster of 3 isolates was characterized by the additional substitutions A155V and R211K in NS1, and D54N in NS2/NEP. The mutation V480I was found only in the PB2 sequence of the two strains exhibiting a reduced susceptibility profile to both NAIs (MjOOselt/MnOZana).

Table 4.1. Non-synonymous mutations detected per gene/protein. Strains are ordered according to epidemiologic week of beginning of symptoms.

	Gene/Protein							
Strain	Epidemiological week	PB2	PB1	PA	NP	NS1	NS2/NEP	Observations
A/Portugal/17/2009	30/2009	V480I	-	P224S	V100I; L122Q	I123V	-	MjO Osel; MnO Zana
A/Portugal/82/2009	32/2009	V480I; D680E	-	P224S	V100I; L122Q; K400R	I123V	-	MjO Osel; MnO Zana
A/Portugal/2/2009	34/2009	-	-	P224S	V100I; L122Q	I123V; F103C	-	n.o.
A/Portugal/7/2009	34/2009	-	S633G	P224S	V100I; L122Q	I123V	-	Pregnancy
A/Portugal/8/2009	39/2009	R54K; I463M	-	P224S; D347N; M561I	V100I; L122Q	I123V	-	Pregnancy
A/Portugal/13/2009	39/2009	R54K	S384L	P224S; D347N; C95G; N409S	V100I; L122Q	I123V	S93G	Pregnancy
A/Portugal/9/2009	43/2009	R54K; I354V	S384L	P224S; D347N	V100I; L122Q	I123V; A155V; R211K	D54N	Pregnancy
A/Portugal/10/2009	44/2009	R54K	S384L	P224S; D347N	V100I; L122Q	A155V; R211K	D54N; G22E	Pregnancy
A/Portugal/11/2009	44/2009	R54K; E249G	S384L	P224S; D347N	V100I; L122Q	I123V; A155V; R211K	D54N	Pregnancy
A/Portugal/18/2009	45/2009	-	-	P224S; S296N	V100I; L122Q	I123V; P114H	-	MnO Osel; ICU
A/Portugal/12/2009	49/2009	R54K	-	P224S; S364N; N614S	V100I; L122Q	I123V	-	Deceased
A/Portugal/6/2009	50/2009	E191K	-	P224S; Q193K; P400L	V100I; L122Q	I123V	-	Children
A/Portugal/1/2010	04/2010	K340N	S678N	P224S; I94L; N675H	V100I; L122Q; V217I; V313I	I123V; A155V	-	Pregnancy

MjO – Major outlier; MnO – Minor outlier; Osel – Oseltamivir; Zana – Zanamivir; ICU – Intensive Care Unit; n.o. – no observations

4.6 Discussion

A(H1N1)pdm09 viruses characterized by the HA substitution S203T and both N248D and V106I in NA, were found to be the dominant circulating strain throughout the epidemic period (2009–2010) in Portugal. Previous studies have also reported dominant circulation of this strain worldwide and suggested this is a consequence of enhanced viral fitness[177, 196].

Within HA1, there was no evidence of evolutionary trends, in agreement with the genetic and antigenic homogeneity of A(H1N1)pdm09[197]. The location of amino acid substitutions observed in HA1 however, suggests these could result in phenotypic changes; the D222E substitution is located within a loop of the receptor-binding site, A73T and S185N within the putative antigenic

sites Cb and Sb, respectively, and K119N has been reported to result in creation of a potential N-glycosylation site[196, 198]. Correlation of these mutations with particular phenotypes, binding specificity or clinical outcome of infection however, needs further investigation.

In the internal genes, the V100I and L122Q substitutions observed in NP, I123V in NS1 and P224S in PA characterize the genetic profile of all isolates analyzed, reflecting the dominant circulation of this A(H1N1)pdm09 genetic variant in Portugal within the period analyzed. Residues 100 and 122 in NP are located within the body domain of the protein and are thought to be involved in PB2 and PB1–NP interaction, which is crucial for RNA replication. The amino acid substitution V100I was associated with the raising of the pandemic alert from phase 4 to 6, and is a human to avian signature change known to translate into enhanced viral fitness probably by increasing viral transmissibility or infectivity[177, 199].

I123V in NS1 is located in the effector domain, a functional region regulating cellular apoptosis, and therefore suspected to play a role in adaptation to the human host and to increase virulence[177]. P224S in PA is located in the N-terminal domain, which has endonuclease activity. Although no phenotypic outcome for this mutation has been established to date, endonuclease activity is critical for initiating transcription and therefore, genetic changes in this region might be expected to alter replication kinetics[200]. Although the phenotypic outcome of L122Q has not been established, this mutation has been previously reported as a virulence factor in H5N1 by greatly increasing replication[201].

In the viruses from pregnant women, in addition to the mutations common to all isolates analyzed, a unique profile of 6 genetic signatures in the functional domains of internal proteins defined a specific cluster. Considering their genetic location, they present a high potential for translating into phenotypic changes in fitness and virulence. Both R54K in PB2 and D347N in PA are located within functional domains for binding PB1[194, 202]. Since the function of the polymerase proteins depends on the formation of a heterodimer for optimal viral RNA replication and transcription, these substitutions together with P224S in PA, may have impacted on polymerase subunit binding to PB1 and have had an unpredictable effect on replication in these strains[169, 203]. Within NS1, the cluster is defined by A155V and R211K substitutions, both located within the C-terminal of the effector domain, as is I123V, which has previously been found to regulate cellular apoptosis[12, 202]. Replication and apoptosis are the major mechanisms that define viral fitness. The

occurrence of different point mutations as occurred in strains in this cluster, could be compensatory or have had a cumulative effect leading to a different phenotype. These mutations will continue to be monitored in non-pregnant patients to evaluate whether they are subset exclusive, resulting from enhanced function of individual proteins or increasing functional compatibility of proteins.

The virus isolate from the deceased patient contained a distinctive set of mutations which could have contributed to this more severe clinical outcome. In addition to N248D and V106I in NA, S203T in HA, V100I in NP and I123V in NS1, this isolate contained L122Q in NP, S364N and N614S in PA, and R54K in PB2, all located in PB1 binding domains of the respective proteins[169]. A direct association of PB1 binding to the other polymerase subunits and to NP with a more virulent phenotype has not been established to date. However, the formation of PB2–PB1–PA heterodimer and the binding of PB2 and PB1–NP into ribonucleoproteins both determine vRNA transcription and cRNA synthesis for genome replication and therefore, have the potential to greatly affect the viral replication cycle[169].

Within the subset of outliers to NAIs, the two MjOOsel/MnOZana were defined by I223V in NA and V480I in PB2. Amino acid substitutions at position 223 in NA have previously been shown to affect virus susceptibility to both oseltamivir and zanamivir, in recent A(H1N1)pdm09 and seasonal viruses[204, 205]. Residue 223 is located in the framework of the NA active site, and thus interacts with the catalytic residues to which antiviral drugs bind[206]. The contribution of V480I in PB2 to the reduced susceptibility profile observed is unclear. In addition, no phenotypic impact was observed in A(H1N1)pdm09 viruses containing NA mutation Y155H, previously observed in seasonal A(H1N1) viruses resistant to both NAIs[207].

Correlating specific mutations with virulent phenotypes in the human host is not straightforward. Animal models differ in their immune response to infection and multiple other factors interfere with clinical and epidemiology outcomes of infection. Also, the genetic background of specific mutations is divergent among influenza strains and can have different phenotypic outcomes. Several mutations have however, been recognized as virulence markers. Although some mechanisms remain unclear, the modification of genome transcription, replication, apoptosis and antigenicity are determinants of virulence. Additionally to known molecular markers of virulence,

we have placed selected residues within functional domains of internal proteins that may enhance viral fitness or increase virulence under genomic surveillance.

4.7 Funding

Calouste Gulbenkian Foundation (Fundação Calouste Gulbenkian) research grant funded the experimental work and the Portuguese Science and Technology Foundation the doctoral grants (SFRH/BD/65211/2009; SFRH/BD/62676/2009; SFRH/BD/48532/2008).

4.8 Competing interest

None declared.

4.9 Ethical approval

No human subjects were used in this research study. Specimens were accessed through anonymous banks, reviewed and exempt for specific issue document from the Ethic Commission of the National Institute of Health, Lisbon, Portugal.

4.10 Acknowledgments

The authors would like to thank Dr. Joanna Ellis from the Health Protection Agency in London for the revision of the manuscript.

Chapter 5

Reverse genetics vaccine seeds for influenza: proof of concept in the source of PB1 as a determinant factor in virus growth and antigen yield

Marta Gíria[a], Luís Santos[a,b], João Louro[a,b], Helena Rebelo de Andrade[a,b]

[a] Host-Pathogen Interaction Unit, Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Lisboa, Portugal; [b] Departamento de Doenças Infeciosas, Instituto Nacional de Saúde Dr Ricardo Jorge IP, Lisboa, Portugal

Virology. 2016; 486:21-27

5.1 Abstract

Growth deficits of reverse genetics vaccine seeds have compromised effective immunization. The impairment has been attributed to sub-optimal protein interactions. Some level of dependence may exist between PB1 and antigenic glycoproteins, however, further research is necessary to clarify the extent to which it can be used in favor of seed production. Our objective was to establish proof of concept on the phenotypic outcome of PB1 source in the PR8:A(H1N1)pdm09 reassortants. Reassortants were generated with the gene constellation of the classical 6:2 PR8:HA,NApdm09 seed prototype and the 5:3 reassortant PR8:HA,NA,PB1pdm09. Viral growth and antigen yield were evaluated 12-60h post-infection. The 5:3 reassortant presented statistically significant growth and antigen yield improvements when compared to the 6:2. We believe these findings to be of promising value to vaccine research towards an improvement of reverse genetic seeds, an overall more cost-effective vaccine manufacture and timely delivery.

5.2 Introduction

In influenza seed virus production by reverse genetics, tissue cultures cells are transfected with cDNA clones from the apathogenic high growth A/PuertoRico/08/34 (PR8) backbone and the donor antigenic proteins. The advantages over classical reassortment are 1) seed viruses being generated with the desired gene constellation, obviating the screening for Hemagglutinin (HA) and Neuraminidase (NA) content; 2) the possible insertion of cold adaptation and attenuation genetic features and removal of virulence determinants, such as the oligobasic cleavage site in HA which is particularly relevant in the context of pandemic preparedness, and; 3) rescue performed directly from transfected cells, avoiding the use of eggs. The production time-frame is reduced and the process is more cost-effective [150, 154, 208].

However, World Health Organization (WHO) has reported reduced titers of A(H1N1)pdm09 candidate pandemic vaccines, compromising the timely distribution of vaccines and the effective immunization and immunoprophylaxis of risk groups, which in turn has accentuated the need for improving the process of seed viruses production, as reported by the National Institute of Allergy and Infectious Diseases (NIAID) [150, 164, 165, 209-211].

Current research on pre-pandemic reverse genetics vaccine seeds for A(H5N1) and A(H7N9) have also shown reduced growth kinetics and antigen yield[150, 164-166, 212].

The features of reduced viral growth and antigenic yield have been attributed to a putative sub-optimal protein interaction [150].

Also, retrospective studies on the genetic content of high yield A(H3N2) vaccine seeds, produced by classical reassortment, unexpectedly revealed gene constellations other than the 6:2. The Polymerase Basic-Protein 1(PB1) gene from the donor strain was the most commonly incorporated in the PR8 backbone, together with HA and NA [152, 153, 213].

In an historical perspective of the genetic evolution of influenza virus, the acquisition of PB1 with antigenic glycoproteins is also recurrent. In the reassortment events that lead to the emergence of the 1957 and 1968 pandemic viruses A(H2N2) and A(H3N2), respectively, the contemporary seasonal viruses acquired PB1 genomic segment together with antigenic glycoproteins [76]. A similar pattern was identified in 1947, where a reassortment event between seasonal A(H1N1) viruses has altered the epidemiology of the infection to a near-pandemic geographic dispersion

[76]. Also, mutations in the PB1 gene of A(H1N1)pdm09 have been found to putatively enhance compatibility between PB1 and HA in the zoonotic transmission of influenza A(H1) viruses [183].

The profile of gene segregation in *in vivo* reassortment events and in the classical reassortment for seed virus production then appears to suggest that the functional compatibility between PB1 and antigenic glycoproteins can be determinant to virus fitness. However, the mechanisms of gene segregation in influenza viruses remain unclear.

Studies on A(H5) reverse genetics vaccine seed viruses suggest an enhancement of viral growth, but in the A(H3N2) vaccine seed model viral growth appears to decline with the introduction of PB1 homologous to antigenic proteins [150, 151, 153, 213, 214]. Nevertheless, in the A(H3N2) seed production by classical reassortment, the co-segregation of PB1 and antigenic proteins is frequent, suggesting an enhancement of viral fitness as described above. The functional compatibility may then be influenced by intra-subtype genetic differences.

In A(H1N1)pdm09 reverse genetic seeds, the neuraminidase and the polymerase activities have been proposed to be enhanced by the co-integration of PB1 with antigenic proteins [151].

The objective of this study was to establish proof of concept regarding the phenotypic outcome of the source of PB1 in PR8:A(H1N1)pdm09 reassortants bearing the gene constellation of reverse genetic vaccine seeds. We proposed to evaluate the phenotypic profile of the reassortants over time during infection by determining viral growth and infectivity, cellular apoptosis induction and antigen yield of a 5:3 reassortant bearing PB1 homologous to antigenic proteins, in comparison to the classical 6:2 gene constellation of the PR8:A(H1N1)pdm09 classical seed viruses.

5.3 Materials and methods

5.3.1 Cells, viruses and plasmids

Madin Darby Canine Kidney cells (MDCK-SIAT1, European Collection of Cell Cultures - ECACC) were maintained in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco), 2 mM L-glutamin (Gibco), 1 x Non-Essential Amino Acids (NEAA, Gibco), 24 mM N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid (HEPES, Gibco), 1

mg/ml Geneticin® (Gibco) and incubated at 37ºC. 293T cells were maintained in DMEM supplemented with 2,5 µg/ml Fungizone® (Gibco) and 1 x Penicillin-Streptomycin-Neomicin antibiotic mixture (PSN, Gibco) instead of 1 mg/ml Geneticin® (Gibco) and incubated at 37ºC with 5% CO2.

The A(H1N1)pdm09 strain used in this study was selected for being genetically similar to the consensus of 32 Portuguese strains from the 2009 pandemic period and the 2010/11 epidemic season and also to the profile of worldwide A(H1N1)pdm09 strains from the same period [215].

The cDNA clones of A/PuertoRico/8/34 and an in-house dual promoter transcription plasmid, pCIPolISapIT were kindly provided by the Centre for Disease Control and Prevention (CDC) – Atlanta.

5.3.2 Generation of reassortant viruses and titration of viral stocks

The complete gene libraries for A/Portugal/82/2009 and PR8 were assembled in pCIPolISapIT. Reassortant viruses were generated by plasmid-based reverse-genetics, following an adaptation of the method by Hoffman et al, to present the genomic constellation of both parental viruses A/Portugal/82/2009 and PR8, the reassortant PR8:HA,NApdm09 (6:2 seed prototype) and PR8:HA,NA,PB1pdm09 (5:3 reassortant)[216]. Plasmid-clones of the 8 genomic segments were transfected into co-cultured 293T and MDCK-SIAT1 cells with Lipofectamine 2000 (Invitrogen) according to the manufacturers protocol and incubated at 37ºC with 5% CO2. The medium, without Fetal Bovine Serum (FBS), was changed 4-5h post-transfection and 2,5 µg/ml trypsin-TPCK (Worthington) was added. At 48h post-transfection, supernatants were inoculated in MDCK-SIAT1 and incubated at 37ºC to assemble virus stocks. First passage viral stock titers were determined by Tissue Culture Infectious Dose (TCID50) following the method of Reed and Müench[217].

5.3.3 Evaluation of viral growth and antigen yield

At this phase of the research, considering the purpose of testing the hypothesis of proof of concept and in agreement with the experimental work performed by other authors, as in Abt et al 2011, Wanitchang et al 2010, Cobbin et al 2013, Zhang et al 2011, Pena et al 2011, Fulvini et al 2011, we have considered that methods commonly used in laboratory were applicable to evaluate the

virus growth and antigen yield of the reassortants. This, however, does not disregard the necessity of using standard methods such as single radial immunodiffusion for HA protein quantification and electron microscopy for counting virus particles for vaccine production. In this study, viral growth was evaluated by Hemagglutination, production of viral particles and infectious virus titer. The total cell death and induction of apoptosis were also determined to evaluate if the apoptotic role of PB1-F2 would be altered. The antigen yield was evaluated in terms of the production of vRNA of the Hemagglutinin gene and its translation into hemagglutination capacity. Additionally, the production of vRNA of the Neuraminidase gene and its translation into Neuraminidase activity were also determined.

5.3.4 Infection

Reassortant viruses were inoculated at the same multiplicity of infection (MOI), 0.001 TCID50/cell, in MDCK-SIAT1, in four independent assays. At 12, 24, 36, 48 and 60h post-infection supernatants were collected, centrifuged at 400g for 5 minutes, and stored at -30ºC.

.

5.3.5 Hemagglutination titer

The Hemagglutination titer was determined in parallel in Guinea-Pig Red Blood Cells (GPRBC) and in Human Red Blood Cells (HRBC). Reassortants were tested in duplicates, for each time point post-infection and for each independent assay (n=8). Samples were two-fold serially diluted in DPBS to a 50-µl final volume, mixed with 50ul GPRBC or 50ul HRBC in a 0.75%v/v and incubated at 4ºC for 45 minutes.

5.3.6 Number of virus particles

Supernatants collected at all time points post-infection, from the four independent assays, were processed for viral RNA extraction (QIAGEN). The number vRNA copies of M gene was determined by real time Reverse-Transcription Polymerase Chain Reaction (rtRT-PCR), following an adaptation of the protocols recommended by CDC and WHO [218, 219]. A TaqMan probe was design by the authors for the quantification of vRNA with high specificity, and primer-probe sets were also design for the additional detection and quantification of the PR8 strain. Purified vRNAs from the 4 independent assays were ran in duplicates (n=8), for each time point post-infection,

and a set of serial dilutions of plasmid-cloned M gene were included in each run as a control for quantification.

5.3.7 Infectious virus titer

The Infectious virus titer was determined by Tissue Culture Infectious Dose (TCID50), using 10 replicates for each 10-fold dilution. TCID50 of each reassortant was calculated following the method by Reed and Müench, for each of the four independent assays (n=4), at all time points post-infection[217].

5.3.8 Production of vRNA from the Hemagglutinin and Neuraminidase genes

The quantification of vRNA from the HA and NA genes was determined following the same procedure as described for the quantification of virus particles. Specific TaqMan probes were design by the authors for the quantification of the amount of HA and NA vRNA produced.

5.3.9 Neuraminidase activity

NA activity was determined by an in-house fluorescence MUNANA-based assay shared by Public Health England (PHE, London, UK)[220], as previously described, for each of the four independent assays (n=4) at all time points post-infection [195]. A cutoff of 12500rfu was established and the dilution of each reassortant virus, at each time point post-infection, necessary to achieve NA activity was determined.

5.3.10 Cell death and apoptosis

Cell death and apoptosis were determined by LDH, cells morphology in apoptosis and the activity of caspases 3/7 assays. The assays were performed at all time points post-infection and for the four independent assays (n=4).

The lactate dehydrogenase (LDH) release was determined by a commercial CytoTox96® Non-Radioactive Cytotoxicity assay and the activity of caspases 3/7 was measured by a luminescence assay in infected cultured cells using Promega Caspase-Glo® 3/7 kit, following the manufacturer protocol “in cell-based assays”.

The percentage of cells morphologically in apoptosis were determined as previously described [221]. For this, the cells were fixed with 4% paraformaldehyde in Dulbecco's Phosphate Buffered Saline (DPBS), stained with Hoescht dye 332358 (Sigma-Aldrich Corp.) at 5 mg/ml in DPBS and stored at 4ºC.

5.3.11 Statistical analysis

Statistical analysis was performed using the IBM Statistical Package for Social Sciences (SPSS) v21 software. Normality and homogeneity of variances were evaluated by Shapiro-Wilk and Lavene statistics, respectively. Mann-Whitney non-parametric test was used to evaluate association between variables. For association testing with factor categories with more than three classes, the Multiple Comparison tests were used to compare each category using Dunnet T3. Values of $p<0.05$ were considered statistically significant.

5.4 Results

5.4.1 The presence of PB1 homologous to antigenic proteins HA and NA significantly increased viral replication.

The Hemagglutination titer of the 6:2 seed prototype showed a significant decrease when compared to PR8, which was reverted by the inclusion of PB1 homologous to HA and NA in the 5:3 reassortant, at all time points post-infection in HRBC and at 12-36h and 60h post-infection in GPRBC ($p<0.008$ Mann-Whitney)(Figure 5.1 A and B).

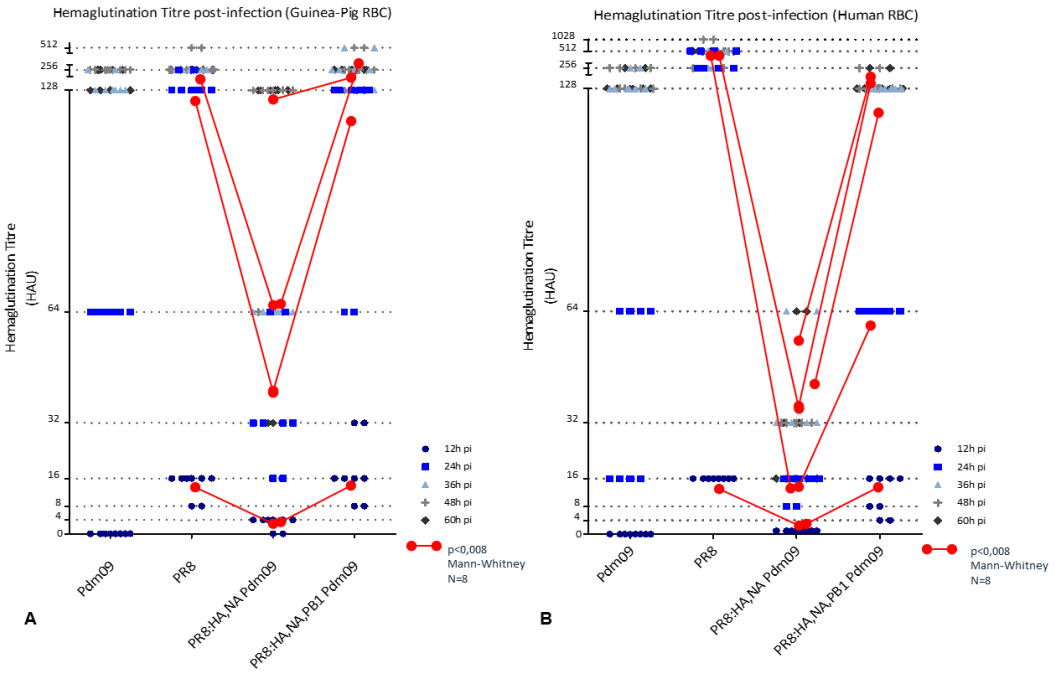


Figure 5.1: Hemagglutination titer of reverse genetic reassortants PR8, A(H1N1)pdm09 (pdm09), the classical 6:2 seed prototype PR8:HA,NApdm09 and the 5:3 reassortant PR8:HA,NA,PB1pdm09, at 12, 24, 36, 48 and 60h post-infection, in Guinea-Pig Red-Blood Cells (A) and Human Red-Blood Cells (B). The Hemagglutination titer was determined in duplicate for each reassortant, at each time point post-infection and for each independent assay (n=8). Bars represent statistically significant differences between the reassortants, as determined by SPSS analysis.

To establish if changes in HA virus titer were a consequence of viral replication kinetics, we analyzed the number of virus particles and the infectious virus titer.

At 12-36h post-infection, the number of virus particles of the 6:2 seed prototype is significantly lower than the one of PR8 and, again, the inclusion of PB1 homologous to HA and NA results in a significant increase in number of particles ($p<0.05$ Dunnett T3)(Figure 5.2 A).

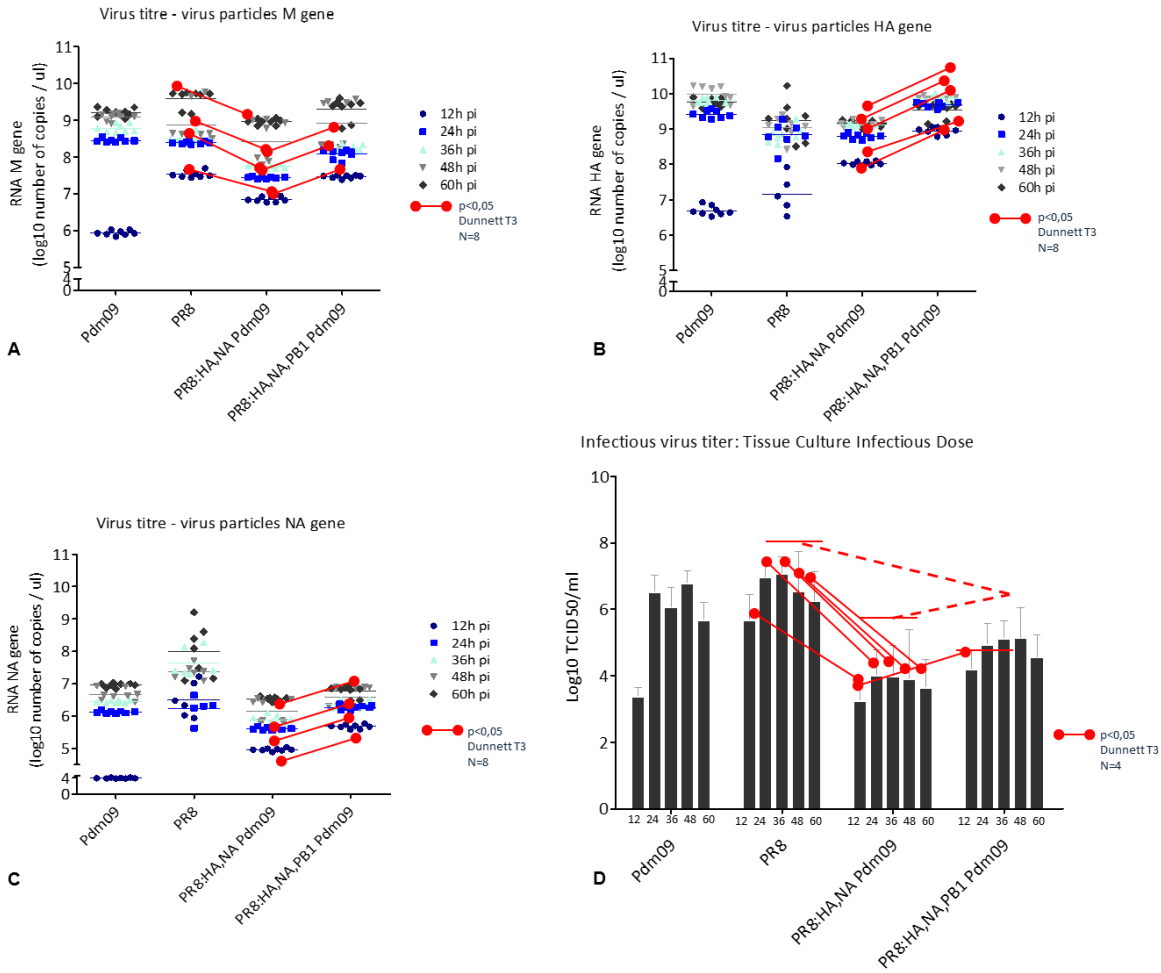


Figure 5.2: Evaluation of viral growth of reverse genetic reassortants PR8, A(H1N1)pdm09 (pdm09), the classical 6:2 seed prototype PR8:HA,NApdm09 and the 5:3 reassortant PR8:HA,NA,PB1pdm09. Number of copies of vRNA from the M (A), HA (B) and NA (C) genes were determined by real time Reverse-Transcription Polymerase Chain Reaction (rtRT-PCR) at 12, 24, 36, 48 and 60h post-infection. RNA was extracted from supernatants of the four independent assays and the number of copies of the M, HA and NA gene segments were determined in duplicates (n=8), at all time points post-infection. Plasmid-cloned M, HA and NA genes were included in each run as controls for the quantification. Infectious virus titer of reverse genetic reassortants was determined by Tissue Culture Infectious Dose (TCID50) using 10 replicates for each 10-fold dilution. The final titer was calculated following the method by Reed and Müench, for each of the four independent assays, at all time points post-infection (n=4). Bars represent statistically significant differences between the reassortants, as determined by SPSS analysis. Dotted bars are used to highlight a trend.

The infectious virus titer shows a similar profile at 12h post-infection. The significant decrease in infectious titer of the 6:2 seed prototype occurs throughout the remaining time points post-infection, when compared to PR8. In the 5:3 reassortant, TCID50 values follow the same trend as the Hemagglutination titer and number of particles and, although the increase is not significant when compared to the 6:2 prototype at later time points, it is also not significant when compared to PR8 ($p<0.05$ Dunnett T3)(Figure 5.2 D).

The significant increase in number of viral particles, Hemagglutination titer and infectious virus titer, are consistent with an enhancement of viral replication kinetics in the 5:3 reassortant when compared to the 6:2 classical seed prototype.

5.4.2 The presence of PB1 homologous to antigenic proteins HA and NA results in a significantly higher hemagglutination capacity

In the 5:3 reassortant virus, the synthesis of HA vRNA shows a significant increase when compared to the 6:2 classical seed prototype at all time points post-infection ($p<0.05$ Dunnett T3)(Figure 5.2 B).

The significant HA vRNA increase is also evident per particle number at 12, 24 and 36h post-infection, which means that more HA may be accumulating in the virions when compared to the classical 6:2 seed prototype ($p<0.05$ Dunnett T3)(Figure 5.3 A). Also, the ratio of HA titer per virus particle shows that there are significantly more hemagglutination units per virus particles in the 5:3 reassortant, at all time points post infection and in both GPRBC and HRBC ($p<0.05$ Mann-Whitney)(Figure 5.3 C and D). This means that not only more hemagglutinin gene segment is apparently being synthesized, but also that it translates into a higher hemagglutination capacity of the virion. The ratio of HA titer per infectious virus titer also shows that for a given infectious titer, the 5:3 reassortant presents a significantly higher hemagglutination capacity than the 6:2 seed prototype at 12-36h post-infections in GPRBC and at all time points post-infection in HRBC ($p<0.05$ Mann-Whitney)(Figure 5.3 E and F).

Differences in vRNA copy number of the M, HA and NA genes could be indicative of the presence of defective virus particles (Figure 5.2A, B and C), which it would be interesting to further explore.

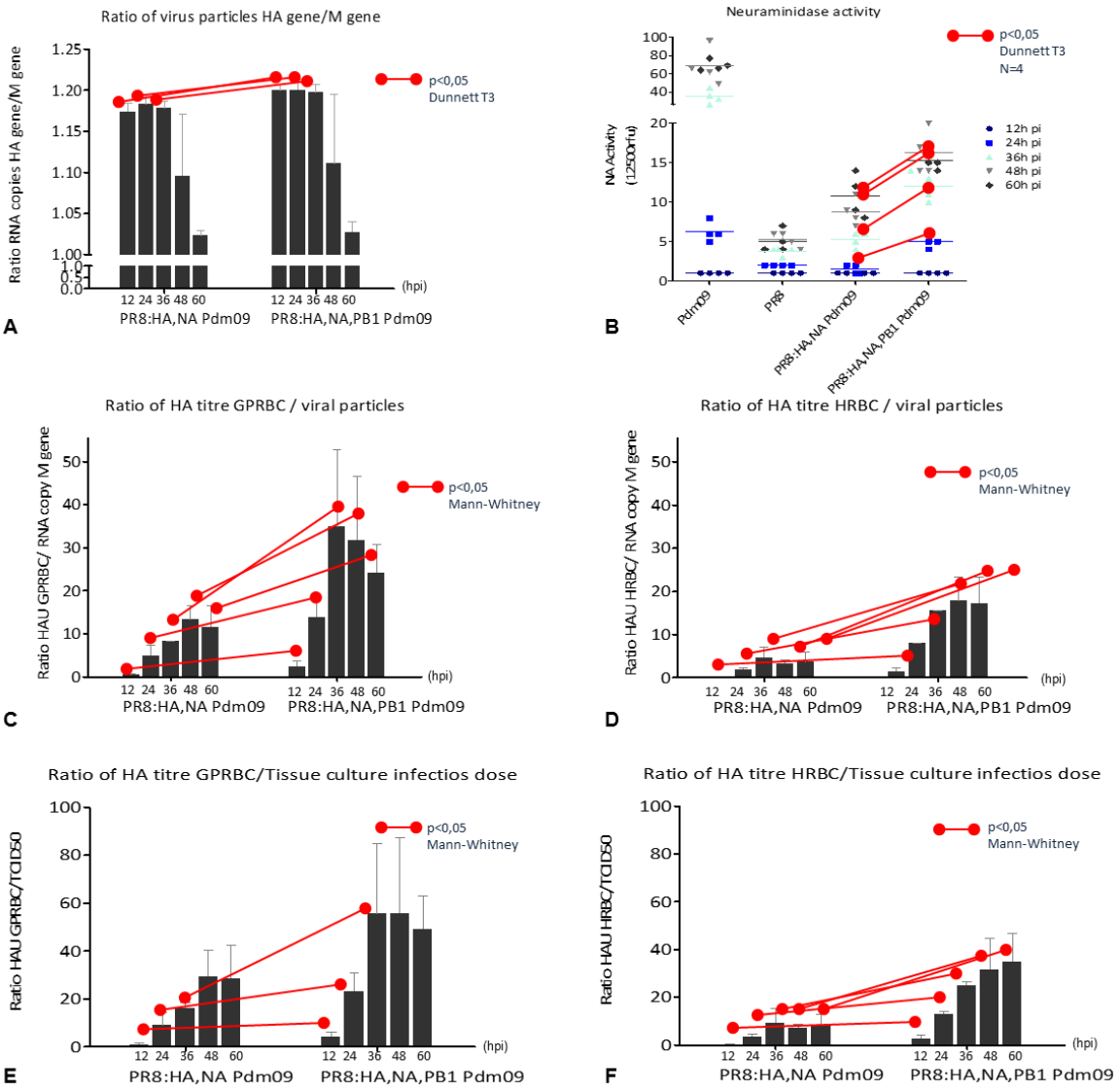


Figure 5.3: Evaluation of antigen yield of the reassortants PR8, A(H1N1)pdm09 (pdm09), the classical 6:2 seed prototype PR8:HA,NApdm09 and the 5:3 reassortant PR8:HA,NA,PB1pdm09, at 12, 24, 36, 48 and 60h post-infection. Figure 5.3A represents the ratio of number of particles of the HA gene per virus particle and 5.3B represents the NA activity, determined by a fluorescence MUNANA-based assay for each of the four independent assays (n=4). For this determination, a 12500rfu cutoff value was used as a reference and the dilution at which the reassortant viruses achieve this NA activity, at each time point post-infection, was then determined. Figures 5.3C and 5.3D represent the ratio of Hemagglutination titer in GPRBC and HRBC per virus particle and Figure 5.3E and 5.3F represent the ratio of Hemagglutination titer in

GPRBC and HRBC per infectious virus titer. Bars represent statistically significant differences between the reassortants, as determined by SPSS analysis.

5.4.3 The presence of PB1 homologous to antigenic proteins HA and NA results in a significantly higher Neuraminidase activity.

The NA vRNA synthesis pattern is similar to the HA, only lower by approximately 3 Log10 in the number of copies in the 5:3 reassortant and the 6:2 classical seed prototype (Figure 5.2 B and C). The 5:3 reassortant presents a significant increase in NA vRNA when compared to the 6:2 seed prototype ($p<0.05$ Dunnett T3)(Figure 5.2 C). This increase translates into a significantly higher NA activity at 24-60h post-infection ($p<0.05$ Dunnett T3)(Figure 5.3 B).

5.4.4 Changing the source of the PB1 gene did not significantly alter the profile of cell death and apoptosis in the 5:3 reassortant

Considering the differences in viral growth and antigen yield, we proceeded to analyze the impact of these parameters in the cellular environment, regarding total cell death and apoptosis induction kinetics.

We found that there is no significant difference in the measurements of LDH release ($p<0.05$ Dunnett T3), cells morphologically in apoptosis ($p<0.05$ Mann-Whitney) and caspase 3/7 activity ($p<0.05$ Dunnett T3) between PR8 and the 6:2 and 5:3 reassortants ($p<0.05$ Dunnett T3) (Figure 5.4 A, B and C).

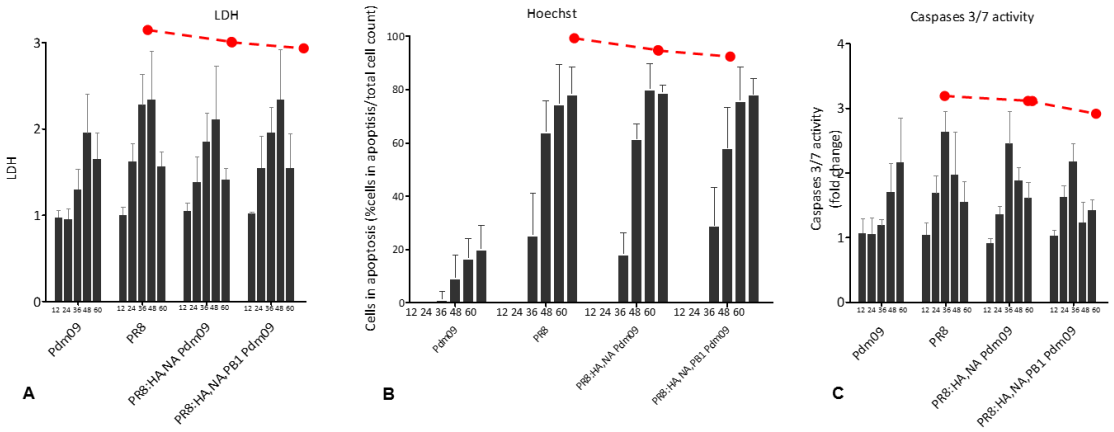


Figure 5.4: Evaluation of cell death and apoptosis of the reassortants PR8, A(H1N1)pdm09 (pdm09), the classical 6:2 seed prototype PR8:HA,NApdm09 and the 5:3 reassortant PR8:HA,NA,PB1pdm09. Dotted bars are used to highlight a trend. The lactate dehydrogenase (LDH) release (A) was determined by CytoTox96® Non-Radioactive Cytotoxicity assay, the percentage of cells morphologically in apoptosis (B) were determined as previously described [221] and the activity of caspases 3/7 (C) was measured by a luminescence assay in infected cultured cells using Promega Caspase-Glo® 3/7 kit, following the manufacturer protocol "in cell-based assays". Measurements were performed for each of the four independent assays (n=4), at all time points post-infection.

5.5 Discussion

In reverse genetics seed virus production, the most crucial factor is to dispose of a high yield, thermostable and attenuated backbone. Thermostability and attenuation features can be inserted in multiple segments of the high yield parental backbone, but the high growth capability can be lost if there is a decrease compatibility between this backbone and the newly inserted glycoproteins of the immunogenic strain.

Cell or egg passage could be a means to naturally increase compatibility and consequently to enhance growth rates, but in seed virus production it may not be an alternative because it could induce genetic changes to HA instead of the adaptation of the internal proteins and, also, it would be extensively time consuming. A different approach has been tested, inducing mutations to the PB1 protein of the PR8 backbone based on its genetic divergence to PB1 of the immunogenic strain, in order to increase similarity and hence putatively more compatible[222]. However, a set

of specific amino-acid changes to PB1 of PR8 with a positive effect in the replication of seed viruses for different immunogenic strains would be challenging to identify. Additionally, it would be necessary to consider the 5´ non-coding regions of the gene segments, since they contain signals that regulate translation. Genetic changes to these regions have been documented to alter the translation rate, potentially enhancing or inhibiting protein expression[223].

The structure–to–function information on the polymerase genes is limited and, to date, it is not possible to predict the replication efficiency or associated pathogenicity nor the effect of the gene constellation in the growth rates of vaccine seeds, on genetic sequence information alone.

Our strategy was to explore the phenotypic outcome of seed viruses bearing PB1 homologous to HA and NA, with a comprehensive analysis of the kinetics of viral growth, antigen yield and induction of cellular apoptosis.

Cell-based vaccines, in general have been produced using egg grown candidate viruses. It has been reported that it would be a major advantage to dispose of candidate vaccine viruses that are grown in cell culture [9, 20]. This is because the use of egg grown candidate viruses could be a constraint if the subsequent vaccine manufacture occurs in cell culture and, on the other hand, currently circulating H3N2 viruses are progressively more difficult to grow in eggs [20]. MDCK and Vero cells are promising platforms to grow vaccine candidate viruses. The experimental procedures in this study have been performed in MDCK-SIAT1 since previous research by other authors have placed this cell line as a preferred candidate for cell-based vaccine manufacture. A(H1N1) viruses, including PR8, have been proven to have similar growth characteristics in MDCIK-SIAT1 cells when compared to MDCK at lower MOI, which makes the process more cost-effective[224]. Additionally, MDCK-SIAT1has been shown to induce less mutations to the HA since the higher receptor levels results in a higher avidity of the interaction receptor-ligand, making this a preferred cell line for laboratory research[225].

We have established proof of concept that in the PR8:A(H1N1)pdm09 seed virus model, virus growth and antigen yield are significantly improved by the inclusion of PB1 of the immunogenic virus, together with HA and NA, when compared to the 6:2 classical seed prototype. Also, we have found that placing PB1 into a new genomic background did not translate into significant changes in cell death and apoptosis which could affect growth rates.

PB1-F2 has a known pro-apoptotic function in the host cells in a later stage of infection and it has been determined to increase polymerase activity *in vitro* by binding to the viral polymerase PB1 subunit. The expression of PB1-F2 is a known determinant of virulence, although the mechanisms behind the increased virulence remain unclear. Cell death and apoptosis were evaluated because changes to the expression of PB1-F2 could alter the balance between the antagonistic mechanisms of cell apoptosis initiated by PB1-F2 and NS1 and be reflected in cell death and replication.

The findings of this study may be of promising value to vaccine research towards an improvement of A(H1N1)pdm09 reverse genetics vaccine seed virus production and an overall more cost-effective vaccine manufacture and timely delivery.

We propose that the compatibility between HA, NA and PB1 may be established early in the infection when the polymerase interacts independently with each segment via their promoter replication signals and that different PB1 proteins have different affinities towards this interaction. We then find it would be very interesting to further explore this concept.

In a broader context, we consider these findings to contribute to a deeper knowledge of the molecular determinants of viral fitness and their role in the dynamics of the evolution of influenza virus. Also, to the awareness of the scientific community regarding the frequency of which natural genomic reassortants occur and it's potential to change viral fitness in the context of disease epidemiology, concerning both severity and geographical dispersion. These findings also increase awareness to the importance of including an evaluation of gene constellation in the risk assessment of emerging strains, as part of the virology surveillance of seasonal strains and of zoonotic transmission.

5.6 Conflict of interest statement

The authors declare no conflict of interest.

5.7 Acknowledgments

The authors would like to thank the Centre for Disease Control and Prevention (CDC) – Atlanta, for the training on Reverse genetics for the construction of viral reassortants, which was applied

for the design of the experimental strategy. Also, the authors would like to thank the CDC – Atlanta for kindly providing the cDNA clones of A/PuertoRico/08/34 and the pCIPolISapIT transcription plasmid.

This research has been funded by the Fundação para a Ciência e Tecnologia grants PTDC/SAU-MIC/122780/2010, SFRH/BD/65211/2009, SFRH/BD/62676/2009 and SFRH/BD/48532/2008.

Chapter 6

General discussion and future research perspectives

The present dissertation documents the rationale and the experimental approach followed to investigate PB1 as a determinant of fitness and adaptation in influenza A virus and, particularly, to contribute to clarify the functional or structural compatibility between PB1 and antigenic proteins as a driver of the overall virus fitness.

In the previous chapters we have discussed the specific findings of this research and its framing in the global influenza research achievements. In this chapter we further identify limitations of the individual studies, discuss three major concepts which we consider arise from this research and identify future research perspectives.

6.1 Limitations of the research

Our approach for evaluating the genetic evolution of PB1 in the zoonotic transmission of swine influenza virus and inferring its putative contribute towards viral fitness and adaptation, described in chapter 2, was a phylogenetically-based, *in silico* analyses focusing on influenza virus PB1 segments derived from avian, swine and human hosts. Based on the timing, frequency, persistence and host-dependence of the emergence of non-synonymous mutations in the PB1 gene, we have inferred the role and significance of the associated protein changes in terms of virus biology in general and host adaptation in particular.

The data collected for the comprehensive phylogenetic analysis of PB1 segment, permitted the detection of amino acid polymorphisms which have arisen and persisted in particular influenza virus lineages.

However, we consider that the hypotheses regarding the potential significance of amino acid mutations in PB1 and PB1-F2 proteins could be strengthened by experimental data. Due to the segmented nature of the virus genome and the use of RNA sequence for non-coding functions in the virus lifecycle such as packaging signals, changes in the protein sequence may or may not be related to advantageous changes in the protein function.

We have begun to address this limitation with complementary research on the phenotypic expression of mutations putatively related to the functional compatibility between PB1 and antigenic proteins, presented in chapter 3. However, additional experimental data would be required to clarify the role for individual mutations. Also, it would be relevant to further investigate

if the differences found in the experimental parameters tested have biologic significance in viruses other than the A(H1N1)pdm09, if and how often are these changes found in nature and what are the mechanisms underlying the altered activity.

Moreover, when analyzing the phylogenetic relatedness of virus strains, it is useful to consider the passage history. This information was not available to us and, particularly in the case of the PR8 strain, the egg or cell passage history could have biased the interpretation of host adaptation. Additionally, the phylogenetic tree was constructed with PhyML, which searches for the tree with the highest probability or likelihood without time-measured. Different algorithms can be used to reconstruct phylogenies that are also a framework for testing evolutionary hypotheses without conditioning on a single tree topology. They are orientated towards rooted, time-measured phylogenies inferred using molecular clock models. For the purpose of this particular analysis, we intended to contextualize the reports on the zoonotic transmissions with what is known about the epidemiology of the viruses in the reservoir or swine host. In this sense, our approach was to: a) analyze the phylogenetic relatedness between the lineages; b) search for the presence or absence of specific genetic changes in given strains, and; c) gather the published historical data on the epidemiology of zoonotic transmissions. We believe that this study could benefit from time-measured phylogenies algorithms since a more specific analysis on evolutionary hypotheses would require a comprehensive dating of the events.

Our approach to evaluate the functional compatibility between PB1 and antigenic proteins as a molecular determinant of virus growth in reverse genetics A(H1N1)pdm09 seed virus prototypes, presented in chapter 5, has been to determine the viral growth kinetics and antigen yield of reassortants. We have determined these parameters as a function of Hemagglutination titer, particle number, infectious virus titer, production of vRNA of the antigenic proteins and neuraminidase activity. We now believe this study could have benefited from additionally determining the Hemagglutinin density in the virion. The incorporation of a high amount of immunogenic HA and NA is very important for vaccine viruses to induce a protective immune response. We have found that more vRNA of the HA gene is produced per virus particle and that the hemagglutination capacity of the PR8:HA, NA, PB1 A(H1N1)pdm09 reassortant is enhanced.

Moreover, it would have been interesting to find if it translates into more HA protein being expressed at the surface of the virion as it would also be appealing to vaccine research.

6.2 Major discussion concepts and future research perspectives

The first concept which we consider arises from this research is that PB1 may be a molecular determinant of the overall virus fitness by establishing interactions with gene segments at the RNA level, both during packaging and to initiate replication. Different affinities towards the interactions during packaging would contribute to determine which segments are packed into the virus particles and, in this way, establish segregation patterns and the genomic compositions that become dominant. On the other hand, different affinities in the interactions between the polymerase and the individual gene segments in the context of viral replication, may result in changes in the replication kinetics.

In the A(H1N1)pdm09 reverse genetic seeds model, virus growth and antigen yield have been found to significantly improve by incorporating PB1 genomic segment from the immunogenic strain together with antigenic glycoproteins and so, the compatibility between PB1 and antigenic proteins appears to favor the overall virus fitness at the replicative level.

However, this has proven not to be the case in the A(H3N2) subtype, where reverse genetic reassortant viruses bearing PB1 from the A(H3N2), homologous to HA and NA, have been reported by other authors to present reduced virus growth, as referred previously[193]. This has occurred even though homologous PB1, HA and NA was the dominant genomic composition in the classical reassortment of A(H3N2) viruses[153].

We consider our findings regarding the A(H1N1)pdm09 seed viruses model to be a major contribute to vaccine research towards a more cost-effective process of manufacture. However, these outcomes make us aware of the fact that the biological advantage of PB1-HA interaction does not necessarily have to be associated with a higher replicative fitness.

Also, we have proposed that specific mutations acquired by the A(H1N1)pdm09 in its evolutionary history may have been involved in enhancing the compatibility between PB1 and HA, by promoting conformational changes and enhancing the span of complementary nucleotides putatively involved in the RNA interaction. The fact that these changes have proven detrimental to growth kinetics *in vitro*, further lead us to suggest that the interaction between genomic

segments may be a determinant of co-segregation but the mechanisms that drive this process are probably not dependent on a replicative advantage.

The interactions that occur during packaging appear to determine the genomic composition of the virus before the interaction between PB1 and the individual genome segments have an impact the replication kinetics. In this case, the dominant strains would be determined by the affinity of the interaction between PB1 and segments from the antigenic proteins during packaging. Subsequently, the higher or lower affinity of the interaction between PB1 and antigenic proteins at the RNA level to initiate replication could then have an impact in the replication kinetics of the virus. We found that in the A(H1N)pdm09 the replication kinetics was enhanced, which makes it appealing to vaccine research, contrary to the A(H3N2) subtype as reported by other authors.

This concept is also in line with a selective packaging model proposed by other authors, in which the genomic composition of dominant strains may not be driven uniquely by the highest replicative fitness *in vivo*, as would occur with a random genome packaging. In this model, the incorporation of a complete set of the eight vRNAs into a virus particle is determined by packaging signals, in such a way that only one vRNA copy from each segment is incorporated and cellular RNA is discarded[6]. We consider that these interactions may also be responsible for the co-segregation of genome segments.

It has been suggested by other authors that the 3´ and 5´ terminal sequences are essential for packaging but do not confer specificity to the process of selecting the vRNAs to pack[192]. Interactions at the vRNA level between NS gene segment and PB1, for example, have been demonstrated to occur in conserved codons in the central coding region of these gene segments, and have been proposed to directly contribute to the selection of the vRNAs that are packed. Disrupting these interactions has been shown to alter the RNA ratios in the virion, suggesting its involvement in the packaging process[192].

One other relevant aspect is that a potentially important region of interaction between PB1 and NA in the A(H3N2) subtype has been identified by other authors and is suggestive of a possible mechanisms for the co-segregation[191]. These authors have also identified *in vitro* interactions between the eight gene segments of A(H3N2) and A(H5N2), predominantly located in distinct regions of the gene segments. In the A(H3N2) these were mostly located in the terminal region while in the A(H5N2) they predominantly occur in the central coding region[191].

These findings lead us to suggest that although a similar mechanism for selective packaging may occur in different influenza subtypes, the specificity of selecting which segments to incorporate may be driven by interactions occurring in distinct regions for each subtype. It would now be interesting to further investigate which packaging sequences are involved in establishing PB1-HA compatibility in the A(H1N1)pdm09 and confirm if the mutations identified in this research as putatively involved are in fact located in these regions. A possible approach for this investigation would be the creation of chimera PB1 genes from the A(H1N1)pdm09, with specific regions originating from the PR8 source, and setting up competitive transfection systems where these chimera PB1 could compete to be incorporated in the genomic backbone.

Further research is necessary to clarify the mechanisms of viral genome packaging, the role of interactions at the RNA level in establishing the co-segregation patterns and the specificities of this interactions at the subtype level. However, it becomes clear that the PB1 genomic segment has a determinant role in establishing its co-incorporation with antigenic proteins, appears to be able to genetically change and adapt to improve this process and in this context, is a determinant factor in the molecular epidemiology of the viruses.

The second concept that we consider arises from this research is the fact that a major improvement may be achieved in the production of reverse genetics seed viruses, by exploring the perception that the compatibility between gene segments or proteins is a determinant factor in the overall viral fitness.

The challenge associated with influenza seed virus production by reverse genetics is that since reassortants are generated with a particular genomic composition, as opposed to the classical reassortment, there will not be a pool from which to choose the high yield progeny. To use and fully explore the advantages of reverse genetics over classical reassortment, it is essential to overcome the growth deficits of some reverse genetic seeds and understand how the growth kinetics and overall fitness can be improved.

Since the enhancement of virus growth by creating seed viruses with PB1 homologous to antigenic proteins may not necessarily be universally applicable for all subtypes, we question if it is possible to develop similar strategies for other subtypes by combining other gene segments with the antigenic proteins for the improvement of vaccine seeds.

To approach this question, we would propose to set up *in vitro* classical reassortment systems between PB8 and different influenza subtypes, identify the gene constellations with a fitness advantage and confirm their impact in growth kinetics by reverse genetics.

Also, it has been proposed by other authors that the introduction of the M gene from the immunogenic strain in A(H5N1) pre-pandemic vaccines could have a positive effect in viral growth[150]. The M1 protein lays beneath the membrane and interacts with antigenic proteins and with the vRNPs and together with M2 are thought to contribute to the production of virus particles, assembly and packaging. It would be interesting to evaluate the impact in viral fitness of introducing the M gene of the immunogenic strain in PR8 reverse genetic seed prototypes of different influenza subtypes. This could be done with a similar experimental approach as the one we used in this study.

Exploring these aspects is a research priority because of the continuous need to produce seed viruses and because of the constraints that currently exist in the large scale cost-effective and timely production.

The third concept we highlight from this research is that the genomic composition of a virus strain is a determinant of virulence, even though the mechanisms that drive the selective packaging remain unclear, and its inclusion in the risk assessment of influenza strains would be extremely relevant.

The segmented nature of influenza genome makes it possible for individual genomic segments to have different origins and this can translate into different levels of compatibility and different phenotypes, as discussed previously. Currently, the risk assessment of influenza strains does not include an evaluation of the viral genomic composition as a determinant of virulence. When analyzing the historical evolution of influenza viruses, reviewed in section *1.2.2.Historical perspective the emergence of influenza A viruses in the human host and current seasonal epidemiology*, of the General Introduction chapter, it becomes apparent that changes in the genomic composition have the potential to impact public health, by changing the epidemiology of the disease to near pandemic dispersions or by enhancing the severity of the illness. Additionally to inter-subtype reassortment events that have resulted in pandemics, intra-subtype reassortment between contemporary strains have resulted in episodes of enhanced geographic dispersion,

vaccine failure or increased severity, as described previously. The results of the present research further strengthen the perception that changes in the genomic composition can translate into distinct phenotypes. At the same time as this may explored as an advantage to virus growth in the context of vaccine research, for example, it must also be recognized as a potential risk to public health.

As a final consideration, we believe our present achievements and future research perspectives to be fully framed within the WHO public health research agenda for influenza, which focuses on limiting the spread and minimizing the impact of pandemic, zoonotic and seasonal epidemic influenza[226].

References

1 Rebelo-de-Andrade H, ed. *Orthomyxovirus*: Lidel - edições técnicas, 2014.
2 Bouvier NM, Palese P. The biology of influenza viruses. *Vaccine*. 2008; **26 Suppl 4**: D49-53.
http://www.ncbi.nlm.nih.gov/pubmed/19230160
3 CDC. Types of Influenza Viruses. 2014.
http://www.cdc.gov/flu/about/viruses/types.htm
4 CDC. Types of inlfuenza viruses. 2016.
http://www.cdc.gov/flu/about/viruses/types.htm
5 Hartge P. A Dictionary of Epidemiology, Sixth Edition. *Am J Epidemiol*. 2015.
http://www.ncbi.nlm.nih.gov/pubmed/25776012
6 Gerber M, Isel C, Moules V, Marquet R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol*. 2014; **22**: 446-455.
http://www.ncbi.nlm.nih.gov/pubmed/24798745
7 Jagger BW, Wise HM, Kash JC, et al. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science*. 2012; **337**: 199-204.
http://www.ncbi.nlm.nih.gov/pubmed/22745253
8 Muramoto Y, Noda T, Kawakami E, Akkina R, Kawaoka Y. Identification of novel influenza A virus proteins translated from PA mRNA. *J Virol*. 2013; **87**: 2455-2462.
http://www.ncbi.nlm.nih.gov/pubmed/23236060
9 Fodor E. The RNA polymerase of influenza a virus: mechanisms of viral transcription and replication. *Acta Virol*. 2013; **57**: 113-122.
http://www.ncbi.nlm.nih.gov/pubmed/23600869
10 Matsuoka Y, Matsumae H, Katoh M, et al. A comprehensive map of the influenza A virus replication cycle. *BMC Syst Biol*. 2013; **7**: 97.
http://www.ncbi.nlm.nih.gov/pubmed/24088197
11 Hale BG, Randall RE, Ortin J, Jackson D. The multifunctional NS1 protein of influenza A viruses. *J Gen Virol*. 2008; **89**: 2359-2376.
http://www.ncbi.nlm.nih.gov/pubmed/18796704
12 Zhirnov OP, Konakova TE, Wolff T, Klenk HD. NS1 protein of influenza A virus down-regulates apoptosis. *J Virol*. 2002; **76**: 1617-1625.
http://www.ncbi.nlm.nih.gov/pubmed/11799156
13 Hai R, Schmolke M, Varga ZT, et al. PB1-F2 expression by the 2009 pandemic H1N1 influenza virus has minimal impact on virulence in animal models. *J Virol*. 2010; **84**: 4442-4450.
http://www.ncbi.nlm.nih.gov/pubmed/20181699
14 CDC. *http://www.cdc.gov/flu/about/disease/complications.htm*. 2016.

15 Biggerstaff M, Cauchemez S, Reed C, Gambhir M, Finelli L. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. *BMC Infect Dis*. 2014; **14**: 480.
http://www.ncbi.nlm.nih.gov/pubmed/25186370
16 WHO. Vaccines against influenza WHO position paper - November 2012. *Wkly Epidemiol Rec*. 2012; **87**: 461-476.
http://www.ncbi.nlm.nih.gov/pubmed/23210147
17 CDC. Prevention and Control of Seasonal Influenza with Vaccines Recommendations of the Advisory Committee on Immunization Practices — United States, 2016–17 Influenza Season. *Morbidity and Mortality Weekly Report (MMWR)*. 2016.

18 Fiore AE, Bridges CB, Cox NJ. Seasonal influenza vaccines. *Curr Top Microbiol Immunol*. 2009; **333**: 43-82.
http://www.ncbi.nlm.nih.gov/pubmed/19768400
19 WHO. *http://www.who.int/biologicals/vaccines/influenza/en/*. 2016.

20 Wiersma LC, Rimmelzwaan GF, de Vries RD. Developing Universal Influenza Vaccines: Hitting the Nail, Not Just on the Head. *Vaccines (Basel)*. 2015; **3**: 239-262.
http://www.ncbi.nlm.nih.gov/pubmed/26343187
21 Jin H, Chen Z. Production of live attenuated influenza vaccines against seasonal and potential pandemic influenza viruses. *Curr Opin Virol*. 2014; **6**: 34-39.
http://www.ncbi.nlm.nih.gov/pubmed/24705137
22 WHO. *Key Facts About Seasonal Flu Vaccine.* 2016.

23 EMEA. *European public assessment reports.* 2016.
http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp&mid=WC0b01ac058001d124&source=homeMedSearch&keyword=influenza&category=human&isNewQuery=true
24 DGS. *Gripe 2016/2017.* 2016.
http://www.dgs.pt/pagina.aspx?f=1&lws=1&mcna=0&lnc=&mid=5005&codigoms=0&codigono=683368347070AAAAAAAAAAAA
25 ECDC. *Influenza vaccination.* 2016.

http://ecdc.europa.eu/en/healthtopics/seasonal_influenza/vaccines/Pages/influenza_vaccination.aspx#vaccinationstrategies
26 FDA. *Influenza (Flu) Antiviral Drugs.* 2016.

27 DGS. Terapêutica e quimioprofilaxia da gripe sazonal. 2015.

28 ECDC. *Antivirals.* 2016.
http://ecdc.europa.eu/en/healthtopics/seasonal_influenza/antivirals/Pages/antivirals.aspx
29 McKimm-Breschkin JL. Influenza neuraminidase inhibitors: antiviral action and mechanisms of resistance. *Influenza Other Respir Viruses*. 2013; **7 Suppl 1**: 25-36.
http://www.ncbi.nlm.nih.gov/pubmed/23279894
30 Pozo F, Lina B, Andrade HR, et al. Guidance for clinical and public health laboratories testing for influenza virus antiviral drug susceptibility in Europe. *J Clin Virol*. 2013; **57**: 5-12.
http://www.ncbi.nlm.nih.gov/pubmed/23375738
31 CDC. *Influenza Antiviral Drug Resistance.* 2016.

32 CDC. *Treatment - Antiviral Drugs.* 2016.

33 Joseph U, Su YC, Vijaykrishna D, Smith GJ. The ecology and adaptive evolution of influenza A interspecies transmission. *Influenza Other Respir Viruses*. 2016.
http://www.ncbi.nlm.nih.gov/pubmed/27426214
34 Su S, Bi Y, Wong G, Gray GC, Gao GF, Li S. Epidemiology, Evolution, and Recent Outbreaks of Avian Influenza Virus in China. *J Virol*. 2015; **89**: 8671-8676.
http://www.ncbi.nlm.nih.gov/pubmed/26063419
35 Schrauwen EJ, Fouchier RA. Host adaptation and transmission of influenza A viruses in mammals. *Emerg Microbes Infect*. 2014; **3**: e9.
http://www.ncbi.nlm.nih.gov/pubmed/26038511
36 Short KR, Richard M, Verhagen JH, et al. One health, multiple challenges: The inter-species transmission of influenza A virus. *One Health*. 2015; **1**: 1-13.
http://www.ncbi.nlm.nih.gov/pubmed/26309905
37 Du Ry van Beest Holle M, Meijer A, Koopmans M, de Jager CM. Human-to-human transmission of avian influenza A/H7N7, The Netherlands, 2003. *Euro Surveill*. 2005; **10**: 264-268.
http://www.ncbi.nlm.nih.gov/pubmed/16371696
38 Wang H, Feng Z, Shu Y, et al. Probable limited person-to-person transmission of highly pathogenic avian influenza A (H5N1) virus in China. *Lancet*. 2008; **371**: 1427-1434.
http://www.ncbi.nlm.nih.gov/pubmed/18400288
39 Ungchusak K, Auewarakul P, Dowell SF, et al. Probable person-to-person transmission of avian influenza A (H5N1). *N Engl J Med*. 2005; **352**: 333-340.
http://www.ncbi.nlm.nih.gov/pubmed/15668219
40 Kandun IN, Wibisono H, Sedyaningsih ER, et al. Three Indonesian clusters of H5N1 virus infection in 2005. *N Engl J Med*. 2006; **355**: 2186-2194.
http://www.ncbi.nlm.nih.gov/pubmed/17124016
41 WHO. Human cases of avian influenza A (H5N1) in North-West Frontier Province, Pakistan, October-November 2007. *Wkly Epidemiol Rec*. 2008; **83**: 359-364.
http://www.ncbi.nlm.nih.gov/pubmed/18833663
42 Li Q, Zhou L, Zhou M, et al. Epidemiology of human infections with avian influenza A(H7N9) virus in China. *N Engl J Med*. 2014; **370**: 520-532.
http://www.ncbi.nlm.nih.gov/pubmed/23614499
43 Nguyen-Van-Tam JS, Nair P, Acheson P, et al. Outbreak of low pathogenicity H7N3 avian influenza in UK, including associated case of human conjunctivitis. *Euro Surveill*. 2006; **11**: E060504 060502.
http://www.ncbi.nlm.nih.gov/pubmed/16816456
44 CDC. Notes from the field: Highly pathogenic avian influenza A (H7N3) virus infection in two poultry workers--Jalisco, Mexico, July 2012. *MMWR Morb Mortal Wkly Rep*. 2012; **61**: 726-727.
http://www.ncbi.nlm.nih.gov/pubmed/22971746
45 WHO. Influenza at the human-animal interface. Summary and assessment, 20 July to 3 October 2016. 2016.

46 WHO. WHO RISK ASSESSMENT of Human infections with avian influenza A(H7N9) virus. 2015.

47 WHO. *http://www.who.int/influenza/human_animal_interface/H5N1_cumulative_table_archives/en/*. 2016.

48 WHO. *Influenza at the human-animal interface.* 2016.
http://www.who.int/influenza/human_animal_interface/Influenza_Summary_IRA_HA_interface_10_03_2016.pdf?ua=1

49 Parrish CR, Murcia PR, Holmes EC. Influenza virus reservoirs and intermediate hosts: dogs, horses, and new possibilities for influenza virus exposure of humans. *J Virol*. 2015; **89**: 2990-2994.
http://www.ncbi.nlm.nih.gov/pubmed/25540375
50 Chan PK. Outbreak of avian influenza A(H5N1) virus infection in Hong Kong in 1997. *Clin Infect Dis*. 2002; **34 Suppl 2**: S58-64.
http://www.ncbi.nlm.nih.gov/pubmed/11938498
51 Freidl GS, Meijer A, de Bruin E, et al. Influenza at the animal-human interface: a review of the literature for virological evidence of human infection with swine or avian influenza viruses other than A(H5N1). *Euro Surveill*. 2014; **19**.
http://www.ncbi.nlm.nih.gov/pubmed/24832117
52 Richard M, Fouchier RA. Influenza A virus transmission via respiratory aerosols or droplets as it relates to pandemic potential. *FEMS Microbiol Rev*. 2016; **40**: 68-85.
http://www.ncbi.nlm.nih.gov/pubmed/26385895
53 Zhang T, Bi Y, Tian H, et al. Human infection with influenza virus A(H10N8) from live poultry markets, China, 2014. *Emerg Infect Dis*. 2014; **20**: 2076-2079.
http://www.ncbi.nlm.nih.gov/pubmed/25425075
54 Fu W, Hu M, Liu X, Xie Y, Xiong Y, Pan H. [Retrospective analysis of human avian influenza A (H10N8) virus infection in Jiangxi province]. *Zhonghua Liu Xing Bing Xue Za Zhi*. 2014; **35**: 1131-1134.
http://www.ncbi.nlm.nih.gov/pubmed/25567019
55 Arzey GG, Kirkland PD, Arzey KE, et al. Influenza virus A (H10N7) in chickens and poultry abattoir workers, Australia. *Emerg Infect Dis*. 2012; **18**: 814-816.
http://www.ncbi.nlm.nih.gov/pubmed/22516302
56 CDC. *http://www.cdc.gov/flu/pandemic-resources/monitoring/viruses-concern.html*. 2016.

57 WHO. Zoonotic flu: Influenza viruses at the human-animal interface. 2016.
http://www.euro.who.int/en/health-topics/communicable-diseases/influenza/zoonotic-influenza
58 Nelli RK, Kuchipudi SV, White GA, Perez BB, Dunham SP, Chang KC. Comparative distribution of human and avian type sialic acid influenza receptors in the pig. *BMC Vet Res*. 2010; **6**: 4.
http://www.ncbi.nlm.nih.gov/pubmed/20105300
59 Van Poucke SG, Nicholls JM, Nauwynck HJ, Van Reeth K. Replication of avian, human and swine influenza viruses in porcine respiratory explants and association with sialic acid distribution. *Virol J*. 2010; **7**: 38.
http://www.ncbi.nlm.nih.gov/pubmed/20158900
60 CDC. Update – Swine-origin triple reassortant influenza A(H3N2) variant viruses in North America 2012.

61 Shu B, Garten R, Emery S, et al. Genetic analysis and antigenic characterization of swine origin influenza viruses isolated from humans in the United States, 1990-2010. *Virology*. 2012; **422**: 151-160.
http://www.ncbi.nlm.nih.gov/pubmed/22078166
62 CDC. *http://www.cdc.gov/flu/pandemic-resources/monitoring/irat-virus-summaries.htm*. 2016.

63 Shinde V, Bridges CB, Uyeki TM, et al. Triple-reassortant swine influenza A (H1) in humans in the United States, 2005-2009. *N Engl J Med*. 2009; **360**: 2616-2625.
http://www.ncbi.nlm.nih.gov/pubmed/19423871
64 Zimmer SM, Burke DS. Historical perspective--Emergence of influenza A (H1N1) viruses. *N Engl J Med*. 2009; **361**: 279-285.
http://www.ncbi.nlm.nih.gov/pubmed/19564632
65 Nelson MI, Vincent AL, Kitikoon P, Holmes EC, Gramer MR. Evolution of novel reassortant A/H3N2 influenza viruses in North American swine and humans, 2009-2011. *J Virol*. 2012; **86**: 8872-8878.
http://www.ncbi.nlm.nih.gov/pubmed/22696653
66 Finelli L, Swerdlow DL. The emergence of influenza A (H3N2)v virus: what we learned from the first wave. *Clin Infect Dis*. 2013; **57 Suppl 1**: S1-3.
http://www.ncbi.nlm.nih.gov/pubmed/23794725
67 CDC. *http://www.cdc.gov/flu/about/season/flu-season-2015-2016.htm*. 2016.

68 CDC. *http://www.cdc.gov/flu/swineflu/variant.htm*. 2016.

69 Peiris JS, Guan Y, Markwell D, Ghose P, Webster RG, Shortridge KF. Cocirculation of avian H9N2 and contemporary "human" H3N2 influenza A viruses in pigs in southeastern China: potential for genetic reassortment? *J Virol*. 2001; **75**: 9679-9686.
http://www.ncbi.nlm.nih.gov/pubmed/11559800
70 Karasin AI, Olsen CW, Brown IH, Carman S, Stalker M, Josephson G. H4N6 influenza virus isolated from pigs in Ontario. *Can Vet J*. 2000; **41**: 938-939.
http://www.ncbi.nlm.nih.gov/pubmed/11143928

71 Nelson MI, Wentworth DE, Culhane MR, et al. Introductions and evolution of human-origin seasonal influenza a viruses in multinational swine populations. *J Virol*. 2014; **88**: 10110-10119.
http://www.ncbi.nlm.nih.gov/pubmed/24965467
72 Nelson MI, Gramer MR, Vincent AL, Holmes EC. Global transmission of influenza viruses from humans to swine. *J Gen Virol*. 2012; **93**: 2195-2203.
http://www.ncbi.nlm.nih.gov/pubmed/22791604
73 Kirkland PD, Finlaison DS, Crispe E, Hurt AC. Influenza virus transmission from horses to dogs, Australia. *Emerg Infect Dis*. 2010; **16**: 699-702.
http://www.ncbi.nlm.nih.gov/pubmed/20350392
74 Anthony SJ, St Leger JA, Pugliares K, et al. Emergence of fatal avian influenza in New England harbor seals. *MBio*. 2012; **3**: e00166-00112.
http://www.ncbi.nlm.nih.gov/pubmed/22851656
75 Worobey M, Han GZ, Rambaut A. Genesis and pathogenesis of the 1918 pandemic H1N1 influenza A virus. *Proc Natl Acad Sci U S A*. 2014; **111**: 8107-8112.
http://www.ncbi.nlm.nih.gov/pubmed/24778238
76 Nelson MI, Viboud C, Simonsen L, et al. Multiple reassortment events in the evolutionary history of H1N1 influenza A virus since 1918. *PLoS Pathog*. 2008; **4**: e1000012.
http://www.ncbi.nlm.nih.gov/pubmed/18463694
77 Wertheim JO. The re-emergence of H1N1 influenza virus in 1977: a cautionary tale for estimating divergence times using biologically unrealistic sampling dates. *PLoS One*. 2010; **5**: e11184.
http://www.ncbi.nlm.nih.gov/pubmed/20567599
78 Dawood FS, Jain S, Finelli L, et al. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med*. 2009; **360**: 2605-2615.
http://www.ncbi.nlm.nih.gov/pubmed/19423869
79 Dawood FS, Iuliano AD, Reed C, et al. Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. *Lancet Infect Dis*. 2012; **12**: 687-695.
http://www.ncbi.nlm.nih.gov/pubmed/22738893
80 Simonsen L, Spreeuwenberg P, Lustig R, et al. Global mortality estimates for the 2009 Influenza Pandemic from the GLaMOR project: a modeling study. *PLoS Med*. 2013; **10**: e1001558.
http://www.ncbi.nlm.nih.gov/pubmed/24302890
81 WHO. Influenza (Seasonal). 2014.
http://who.int/mediacentre/factsheets/fs211/en/
82 WHO. A manual for estimating disease burden associated with seasonal influenza. 2015.

83 Schrauwen EJ, de Graaf M, Herfst S, Rimmelzwaan GF, Osterhaus AD, Fouchier RA. Determinants of virulence of influenza A virus. *Eur J Clin Microbiol Infect Dis*. 2014; **33**: 479-490.
http://www.ncbi.nlm.nih.gov/pubmed/24078062
84 Chutinimitkul S, van Riel D, Munster VJ, et al. In vitro assessment of attachment pattern and replication efficiency of H5N1 influenza A viruses with altered receptor specificity. *J Virol*. 2010; **84**: 6825-6833.
http://www.ncbi.nlm.nih.gov/pubmed/20392847
85 Chutinimitkul S, Herfst S, Steel J, et al. Virulence-associated substitution D222G in the hemagglutinin of 2009 pandemic influenza A(H1N1) virus affects receptor binding. *J Virol*. 2010; **84**: 11802-11813.
http://www.ncbi.nlm.nih.gov/pubmed/20844044
86 Watanabe Y, Ibrahim MS, Ellakany HF, et al. Acquisition of human-type receptor binding specificity by new H5N1 influenza virus sublineages during their emergence in birds in Egypt. *PLoS Pathog*. 2011; **7**: e1002068.
http://www.ncbi.nlm.nih.gov/pubmed/21637809
87 Chen H, Bright RA, Subbarao K, et al. Polygenic virulence factors involved in pathogenesis of 1997 Hong Kong H5N1 influenza viruses in mice. *Virus Res*. 2007; **128**: 159-163.
http://www.ncbi.nlm.nih.gov/pubmed/17521765
88 Zhang Y, Zhu J, Li Y, et al. Glycosylation on hemagglutinin affects the virulence and pathogenicity of pandemic H1N1/2009 influenza A virus in mice. *PLoS One*. 2013; **8**: e61397.
http://www.ncbi.nlm.nih.gov/pubmed/23637827
89 Wang W, Lu B, Zhou H, et al. Glycosylation at 158N of the hemagglutinin protein and receptor binding specificity synergistically affect the antigenicity and immunogenicity of a live attenuated H5N1 A/Vietnam/1203/2004 vaccine virus in ferrets. *J Virol*. 2010; **84**: 6570-6577.
http://www.ncbi.nlm.nih.gov/pubmed/20427525
90 Wagner R, Matrosovich M, Klenk HD. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Rev Med Virol*. 2002; **12**: 159-166.
http://www.ncbi.nlm.nih.gov/pubmed/11987141
91 de Wit E, Munster VJ, van Riel D, et al. Molecular determinants of adaptation of highly pathogenic avian influenza H7N7 viruses to efficient replication in the human host. *J Virol*. 2010; **84**: 1597-1606.
http://www.ncbi.nlm.nih.gov/pubmed/19939933

92 Fouchier RA, Schneeberger PM, Rozendaal FW, et al. Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. *Proc Natl Acad Sci U S A*. 2004; **101**: 1356-1361.
http://www.ncbi.nlm.nih.gov/pubmed/14745020
93 Sorrell EM, Song H, Pena L, Perez DR. A 27-amino-acid deletion in the neuraminidase stalk supports replication of an avian H2N2 influenza A virus in the respiratory tract of chickens. *J Virol*. 2010; **84**: 11831-11840.
http://www.ncbi.nlm.nih.gov/pubmed/20826691
94 Munier S, Larcher T, Cormier-Aline F, et al. A genetically engineered waterfowl influenza virus with a deletion in the stalk of the neuraminidase has increased virulence for chickens. *J Virol*. 2010; **84**: 940-952.
http://www.ncbi.nlm.nih.gov/pubmed/19889765
95 Gabriel G, Klingel K, Otte A, et al. Differential use of importin-alpha isoforms governs cell tropism and host adaptation of influenza virus. *Nat Commun*. 2011; **2**: 156.
http://www.ncbi.nlm.nih.gov/pubmed/21245837
96 Varga ZT, Ramos I, Hai R, et al. The influenza virus protein PB1-F2 inhibits the induction of type I interferon at the level of the MAVS adaptor protein. *PLoS Pathog*. 2011; **7**: e1002067.
http://www.ncbi.nlm.nih.gov/pubmed/21695240
97 Conenello GM, Zamarin D, Perrone LA, Tumpey T, Palese P. A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 influenza A viruses contributes to increased virulence. *PLoS Pathog*. 2007; **3**: 1414-1421.
http://www.ncbi.nlm.nih.gov/pubmed/17922571
98 McAuley JL, Chipuk JE, Boyd KL, Van De Velde N, Green DR, McCullers JA. PB1-F2 proteins from H5N1 and 20 century pandemic influenza viruses cause immunopathology. *PLoS Pathog*. 2010; **6**: e1001014.
http://www.ncbi.nlm.nih.gov/pubmed/20661425
99 Smith AM, McCullers JA. Molecular signatures of virulence in the PB1-F2 proteins of H5N1 influenza viruses. *Virus Res*. 2013; **178**: 146-150.
http://www.ncbi.nlm.nih.gov/pubmed/23499672
100 McAuley JL, Hornung F, Boyd KL, et al. Expression of the 1918 influenza A virus PB1-F2 enhances the pathogenesis of viral and secondary bacterial pneumonia. *Cell Host Microbe*. 2007; **2**: 240-249.
http://www.ncbi.nlm.nih.gov/pubmed/18005742
101 Marc D. Influenza virus non-structural protein NS1: interferon antagonism and beyond. *J Gen Virol*. 2014; **95**: 2594-2611.
http://www.ncbi.nlm.nih.gov/pubmed/25182164
102 Li W, Wang G, Zhang H, et al. Effects of NS1 variants of H5N1 influenza virus on interferon induction, TNFalpha response and p53 activity. *Cell Mol Immunol*. 2010; **7**: 235-242.
http://www.ncbi.nlm.nih.gov/pubmed/20228833
103 Seo SH, Hoffmann E, Webster RG. Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. *Nat Med*. 2002; **8**: 950-954.
http://www.ncbi.nlm.nih.gov/pubmed/12195436
104 WHO. Pandemic influenza. 2016.
http://www.euro.who.int/en/health-topics/communicable-diseases/influenza/pandemic-influenza
105 Yuan J, Lau EH, Li K, et al. Effect of Live Poultry Market Closure on Avian Influenza A(H7N9) Virus Activity in Guangzhou, China, 2014. *Emerg Infect Dis*. 2015; **21**: 1784-1793.
http://www.ncbi.nlm.nih.gov/pubmed/26402310
106 Bouvier NM, Lowen AC. Animal Models for Influenza Virus Pathogenesis and Transmission. *Viruses*. 2010; **2**: 1530-1563.
http://www.ncbi.nlm.nih.gov/pubmed/21442033
107 Lowen AC, Mubareka S, Tumpey TM, Garcia-Sastre A, Palese P. The guinea pig as a transmission model for human influenza viruses. *Proc Natl Acad Sci U S A*. 2006; **103**: 9988-9992.
http://www.ncbi.nlm.nih.gov/pubmed/16785447
108 Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog*. 2007; **3**: 1470-1476.
http://www.ncbi.nlm.nih.gov/pubmed/17953482
109 Lowen AC, Steel J, Mubareka S, Palese P. High temperature (30 degrees C) blocks aerosol but not contact transmission of influenza virus. *J Virol*. 2008; **82**: 5650-5652.
http://www.ncbi.nlm.nih.gov/pubmed/18367530
110 Steel J, Palese P, Lowen AC. Transmission of a 2009 pandemic influenza virus shows a sensitivity to temperature and humidity similar to that of an H3N2 seasonal strain. *J Virol*. 2011; **85**: 1400-1402.
http://www.ncbi.nlm.nih.gov/pubmed/21084485
111 Tumpey TM, Maines TR, Van Hoeven N, et al. A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. *Science*. 2007; **315**: 655-659.
http://www.ncbi.nlm.nih.gov/pubmed/17272724

112 Roberts KL, Shelton H, Scull M, Pickles R, Barclay WS. Lack of transmission of a human influenza virus with avian receptor specificity between ferrets is not due to decreased virus shedding but rather a lower infectivity in vivo. *J Gen Virol*. 2011; **92**: 1822-1831.
http://www.ncbi.nlm.nih.gov/pubmed/21508186
113 Pappas C, Viswanathan K, Chandrasekaran A, et al. Receptor specificity and transmission of H2N2 subtype viruses isolated from the pandemic of 1957. *PLoS One*. 2010; **5**: e11158.
http://www.ncbi.nlm.nih.gov/pubmed/20574518
114 Matrosovich M, Tuzikov A, Bovin N, et al. Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. *J Virol*. 2000; **74**: 8502-8512.
http://www.ncbi.nlm.nih.gov/pubmed/10954551
115 Manz B, Schwemmle M, Brunotte L. Adaptation of avian influenza A virus polymerase in mammals to overcome the host species barrier. *J Virol*. 2013; **87**: 7200-7209.
http://www.ncbi.nlm.nih.gov/pubmed/23616660
116 Arai Y, Kawashita N, Daidoji T, et al. Novel Polymerase Gene Mutations for Human Adaptation in Clinical Isolates of Avian H5N1 Influenza Viruses. *PLoS Pathog*. 2016; **12**: e1005583.
http://www.ncbi.nlm.nih.gov/pubmed/27097026
117 Gabriel G, Czudai-Matwich V, Klenk HD. Adaptive mutations in the H5N1 polymerase complex. *Virus Res*. 2013; **178**: 53-62.
http://www.ncbi.nlm.nih.gov/pubmed/23732876
118 Steel J, Lowen AC, Mubareka S, Palese P. Transmission of influenza virus in a mammalian host is increased by PB2 amino acids 627K or 627E/701N. *PLoS Pathog*. 2009; **5**: e1000252.
http://www.ncbi.nlm.nih.gov/pubmed/19119420
119 Van Hoeven N, Pappas C, Belser JA, et al. Human HA and polymerase subunit PB2 proteins confer transmission of an avian influenza virus through the air. *Proc Natl Acad Sci U S A*. 2009; **106**: 3366-3371.
http://www.ncbi.nlm.nih.gov/pubmed/19211790
120 Liu Q, Qiao C, Marjuki H, et al. Combination of PB2 271A and SR polymorphism at positions 590/591 is critical for viral replication and virulence of swine influenza virus in cultured cells and in vivo. *J Virol*. 2012; **86**: 1233-1237.
http://www.ncbi.nlm.nih.gov/pubmed/22072752
121 Chou YY, Albrecht RA, Pica N, et al. The M segment of the 2009 new pandemic H1N1 influenza virus is critical for its high transmission efficiency in the guinea pig model. *J Virol*. 2011; **85**: 11235-11241.
http://www.ncbi.nlm.nih.gov/pubmed/21880744
122 Campbell PJ, Danzy S, Kyriakis CS, Deymier MJ, Lowen AC, Steel J. The M segment of the 2009 pandemic influenza virus confers increased neuraminidase activity, filamentous morphology, and efficient contact transmissibility to A/Puerto Rico/8/1934-based reassortant viruses. *J Virol*. 2014; **88**: 3802-3814.
http://www.ncbi.nlm.nih.gov/pubmed/24429367
123 Ma W, Liu Q, Bawa B, et al. The neuraminidase and matrix genes of the 2009 pandemic influenza H1N1 virus cooperate functionally to facilitate efficient replication and transmissibility in pigs. *J Gen Virol*. 2012; **93**: 1261-1268.
http://www.ncbi.nlm.nih.gov/pubmed/22337640
124 Kimble JB, Angel M, Wan H, Sutton TC, Finch C, Perez DR. Alternative reassortment events leading to transmissible H9N1 influenza viruses in the ferret model. *J Virol*. 2014; **88**: 66-71.
http://www.ncbi.nlm.nih.gov/pubmed/24131710
125 Wan H, Sorrell EM, Song H, et al. Replication and transmission of H9N2 influenza viruses in ferrets: evaluation of pandemic potential. *PLoS One*. 2008; **3**: e2923.
http://www.ncbi.nlm.nih.gov/pubmed/18698430
126 Sorrell EM, Wan H, Araya Y, Song H, Perez DR. Minimal molecular constraints for respiratory droplet transmission of an avian-human H9N2 influenza A virus. *Proc Natl Acad Sci U S A*. 2009; **106**: 7565-7570.
http://www.ncbi.nlm.nih.gov/pubmed/19380727
127 Chen LM, Blixt O, Stevens J, et al. In vitro evolution of H5N1 avian influenza virus toward human-type receptor specificity. *Virology*. 2012; **422**: 105-113.
http://www.ncbi.nlm.nih.gov/pubmed/22056389
128 Imai M, Watanabe T, Hatta M, et al. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature*. 2012; **486**: 420-428.
http://www.ncbi.nlm.nih.gov/pubmed/22722205
129 Zhang W, Shi Y, Lu X, Shu Y, Qi J, Gao GF. An airborne transmissible avian influenza H5 hemagglutinin seen at the atomic level. *Science*. 2013; **340**: 1463-1467.
http://www.ncbi.nlm.nih.gov/pubmed/23641058
130 Sutton TC, Finch C, Shao H, et al. Airborne transmission of highly pathogenic H7N1 influenza virus in ferrets. *J Virol*. 2014; **88**: 6623-6635.
http://www.ncbi.nlm.nih.gov/pubmed/24696487

131 Gao Y, Zhang Y, Shinya K, et al. Identification of amino acids in HA and PB2 critical for the transmission of H5N1 avian influenza viruses in a mammalian host. *PLoS Pathog*. 2009; **5**: e1000709.
http://www.ncbi.nlm.nih.gov/pubmed/20041223
132 Galloway SE, Reed ML, Russell CJ, Steinhauer DA. Influenza HA subtypes demonstrate divergent phenotypes for cleavage activation and pH of fusion: implications for host range and adaptation. *PLoS Pathog*. 2013; **9**: e1003151.
http://www.ncbi.nlm.nih.gov/pubmed/23459660
133 Wargo AR, Kurath G. Viral fitness: definitions, measurement, and current insights. *Curr Opin Virol*. 2012; **2**: 538-545.
http://www.ncbi.nlm.nih.gov/pubmed/22986085
134 Santiago FE. RNA virus robustness: possible causes and some consequences. *current opinion in virology*. 2012; **2**: 525-530.

135 Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiol Mol Biol Rev*. 2012; **76**: 159-216.
http://www.ncbi.nlm.nih.gov/pubmed/22688811
136 Ding N, Wu N, Xu Q, Chen K, Zhang C. Molecular evolution of novel swine-origin A/H1N1 influenza viruses among and before human. *Virus Genes*. 2009; **39**: 293-300.
http://www.ncbi.nlm.nih.gov/pubmed/19693661
137 Su YC, Bahl J, Joseph U, et al. Phylodynamics of H1N1/2009 influenza reveals the transition from host adaptation to immune-driven selection. *Nat Commun*. 2015; **6**: 7952.
http://www.ncbi.nlm.nih.gov/pubmed/26245473
138 Borderia AV, Stapleford KA, Vignuzzi M. RNA virus population diversity: implications for inter-species transmission. *Curr Opin Virol*. 2011; **1**: 643-648.
http://www.ncbi.nlm.nih.gov/pubmed/22440922
139 Manrubia SC. Modelling viral evolution and adaptation: challenges and rewards. *Curr Opin Virol*. 2012; **2**: 531-537.
http://www.ncbi.nlm.nih.gov/pubmed/22796325
140 Bull JJ, Sanjuan R, Wilke CO. Theory of lethal mutagenesis for viruses. *J Virol*. 2007; **81**: 2930-2939.
http://www.ncbi.nlm.nih.gov/pubmed/17202214
141 Paff ML, Stolte SP, Bull JJ. Lethal mutagenesis failure may augment viral adaptation. *Mol Biol Evol*. 2014; **31**: 96-105.
http://www.ncbi.nlm.nih.gov/pubmed/24092771
142 Martin G, Gandon S. Lethal mutagenesis and evolutionary epidemiology. *Philos Trans R Soc Lond B Biol Sci*. 2010; **365**: 1953-1963.
http://www.ncbi.nlm.nih.gov/pubmed/20478890
143 Ghedin E, Fitch A, Boyne A, et al. Mixed infection and the genesis of influenza virus diversity. *J Virol*. 2009; **83**: 8832-8841.
http://www.ncbi.nlm.nih.gov/pubmed/19553313
144 Cheung PP, Watson SJ, Choy KT, et al. Generation and characterization of influenza A viruses with altered polymerase fidelity. *Nat Commun*. 2014; **5**: 4794.
http://www.ncbi.nlm.nih.gov/pubmed/25183443
145 Basler CF, Aguilar PV. Progress in identifying virulence determinants of the 1918 H1N1 and the Southeast Asian H5N1 influenza A viruses. *Antiviral Res*. 2008; **79**: 166-178.
http://www.ncbi.nlm.nih.gov/pubmed/18547656
146 Khiabanian H, Trifonov V, Rabadan R. Reassortment patterns in Swine influenza viruses. *PLoS Curr*. 2009; **1**: RRN1008.
http://www.ncbi.nlm.nih.gov/pubmed/20029610
147 Lu L, Lycett SJ, Leigh Brown AJ. Reassortment patterns of avian influenza virus internal segments among different subtypes. *BMC Evol Biol*. 2014; **14**: 16.
http://www.ncbi.nlm.nih.gov/pubmed/24456010
148 Chen LM, Davis CT, Zhou H, Cox NJ, Donis RO. Genetic compatibility and virulence of reassortants derived from contemporary avian H5N1 and human H3N2 influenza A viruses. *PLoS Pathog*. 2008; **4**: e1000072.
http://www.ncbi.nlm.nih.gov/pubmed/18497857
149 Holmes EC, Ghedin E, Miller N, et al. Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. *PLoS Biol*. 2005; **3**: e300.
http://www.ncbi.nlm.nih.gov/pubmed/16026181
150 Abt M, de Jonge J, Laue M, Wolff T. Improvement of H5N1 influenza vaccine viruses: influence of internal gene segments of avian and human origin on production and hemagglutinin content. *Vaccine*. 2011; **29**: 5153-5162.
http://www.ncbi.nlm.nih.gov/pubmed/21624413

151 Wanitchang A, Kramyu J, Jongkaewwattana A. Enhancement of reverse genetics-derived swine-origin H1N1 influenza virus seed vaccine growth by inclusion of indigenous polymerase PB1 protein. *Virus Res*. 2010; **147**: 145-148.
http://www.ncbi.nlm.nih.gov/pubmed/19879908
152 Bergeron C, Valette M, Lina B, Ottmann M. Genetic content of Influenza H3N2 vaccine seeds. *PLoS Curr*. 2010; **2**: RRN1165.
http://www.ncbi.nlm.nih.gov/pubmed/20802842
153 Cobbin JC, Verity EE, Gilbertson BP, Rockman SP, Brown LE. The source of the PB1 gene in influenza vaccine reassortants selectively alters the hemagglutinin content of the resulting seed virus. *J Virol*. 2013; **87**: 5577-5585.
http://www.ncbi.nlm.nih.gov/pubmed/23468502
154 Krammer F, Palese P. Advances in the development of influenza virus vaccines. *Nat Rev Drug Discov*. 2015; **14**: 167-182.
http://www.ncbi.nlm.nih.gov/pubmed/25722244
155 Berlanda Scorza F, Tsvetnitsky V, Donnelly JJ. Universal influenza vaccines: Shifting to better vaccines. *Vaccine*. 2016; **34**: 2926-2933.
http://www.ncbi.nlm.nih.gov/pubmed/27038130
156 Joyce MG, Wheatley AK, Thomas PV, et al. Vaccine-Induced Antibodies that Neutralize Group 1 and Group 2 Influenza A Viruses. *Cell*. 2016; **166**: 609-623.
http://www.ncbi.nlm.nih.gov/pubmed/27453470
157 CDC. *http://www.cdc.gov/flu/about/qa/advances.htm#vaccine-advances*. 2016.

158 WHO. Pandemic influenza vaccine manufacturing process and timeline. 2009.
http://www.who.int/csr/disease/swineflu/notes/h1n1_vaccine_20090806/en/
159 Fauci. Vaccine Research and Development: Challenges and Opportunities. *First Global Vaccine and Immunization Research Forum (GVIRF)*: USA 2014.
http://www.who.int/immunization/research/forums_and_initiatives/gvirf_bethesda_march2014/en/
160 WHO. Influenza Vaccine Virus Selection And Development. 2007.
http://apps.who.int/gb/pip/pdf_files/Fluvaccvirusselection.pdf?ua=1
161 Krammer F, Palese P, Steel J. Advances in universal influenza virus vaccine design and antibody mediated therapies based on conserved regions of the hemagglutinin. *Curr Top Microbiol Immunol*. 2015; **386**: 301-321.
http://www.ncbi.nlm.nih.gov/pubmed/25007847
162 Plant. Gene Constellation of Influenza Vaccine Seed Viruses. In: InTech, ed. 2013.
163 Fineberg HV. Pandemic preparedness and response--lessons from the H1N1 influenza of 2009. *N Engl J Med*. 2014; **370**: 1335-1342.
http://www.ncbi.nlm.nih.gov/pubmed/24693893
164 Harvey R, Wheeler JX, Wallis CL, Robertson JS, Engelhardt OG. Quantitation of haemagglutinin in H5N1 influenza viruses reveals low haemagglutinin content of vaccine virus NIBRG-14 (H5N1). *Vaccine*. 2008; **26**: 6550-6554.
http://www.ncbi.nlm.nih.gov/pubmed/18840494
165 Stephenson I, Gust I, Pervikov Y, Kieny MP. Development of vaccines against influenza H5. *Lancet Infect Dis*. 2006; **6**: 458-460.
http://www.ncbi.nlm.nih.gov/pubmed/16870521
166 Zhang W, Xue T, Wu X, et al. Increase in viral yield in eggs and MDCK cells of reassortant H5N1 vaccine candidate viruses caused by insertion of 38 amino acids into the NA stalk. *Vaccine*. 2011; **29**: 8032-8041.
http://www.ncbi.nlm.nih.gov/pubmed/21864614
167 NIAID. Research on new vaccine technologies. 2015.
http://www.niaid.nih.gov/topics/Flu/Research/vaccineResearch/pages/technologies.aspx
168 Bhatt S, Lam TT, Lycett SJ, et al. The evolutionary dynamics of influenza A virus adaptation to mammalian hosts. *Philos Trans R Soc Lond B Biol Sci*. 2013; **368**: 20120382.
http://www.ncbi.nlm.nih.gov/pubmed/23382435
169 Perez DR, Donis RO. Functional analysis of PA binding by influenza a virus PB1: effects on polymerase activity and viral infectivity. *J Virol*. 2001; **75**: 8127-8136.
http://www.ncbi.nlm.nih.gov/pubmed/11483758
170 Reuther P, Manz B, Brunotte L, Schwemmle M, Wunderlich K. Targeting of the influenza A virus polymerase PB1-PB2 interface indicates strain-specific assembly differences. *J Virol*. 2011; **85**: 13298-13309.
http://www.ncbi.nlm.nih.gov/pubmed/21957294
171 Chu C, Fan S, Li C, et al. Functional analysis of conserved motifs in influenza virus PB1 protein. *PLoS One*. 2012; **7**: e36113.
http://www.ncbi.nlm.nih.gov/pubmed/22615752
172 Krumbholz A, Philipps A, Oehring H, et al. Current knowledge on PB1-F2 of influenza A viruses. *Med Microbiol Immunol*. 2011; **200**: 69-75.
http://www.ncbi.nlm.nih.gov/pubmed/20953627
173 Vater S. Studies on Influenza A PB1-F2 Protein http://hdl.handle.net/10023/2075. 2011.

174 Khiabanian H, Trifonov V, Rabadan R. Reassortment patterns in Swine influenza viruses. *PLoS One*. 2009; **4**: e7366.
http://www.ncbi.nlm.nih.gov/pubmed/19809504
175 Ma W, Lager KM, Lekcharoensuk P, et al. Viral reassortment and transmission after co-infection of pigs with classical H1N1 and triple-reassortant H3N2 swine influenza viruses. *J Gen Virol*. 2010; **91**: 2314-2321.
http://www.ncbi.nlm.nih.gov/pubmed/20484565
176 McAuley JL, Zhang K, McCullers JA. The effects of influenza A virus PB1-F2 protein on polymerase activity are strain specific and do not impact pathogenesis. *J Virol*. 2010; **84**: 558-564.
http://www.ncbi.nlm.nih.gov/pubmed/19828614
177 Pan C, Cheung B, Tan S, et al. Genomic signature and mutation trend analysis of pandemic (H1N1) 2009 influenza A virus. *PLoS One*. 2010; **5**: e9549.
http://www.ncbi.nlm.nih.gov/pubmed/20221396
178 Dundon WG. Variability among the neuraminidase, non-structural 1 and PB1-F2 proteins in the influenza A virus genome. *Virus Genes*. 2012; **44**: 363-373.
http://www.ncbi.nlm.nih.gov/pubmed/22261818
179 Trifonov V, Racaniello V, Rabadan R. The Contribution of the PB1-F2 Protein to the Fitness of Influenza A Viruses and its Recent Evolution in the 2009 Influenza A (H1N1) Pandemic Virus. *PLoS Curr*. 2009; **1**: RRN1006.
http://www.ncbi.nlm.nih.gov/pubmed/20029605
180 Biswas SK, Nayak DP. Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. *J Virol*. 1994; **68**: 1819-1826.
http://www.ncbi.nlm.nih.gov/pubmed/8107244
181 Ohtsu Y, Honda Y, Sakata Y, Kato H, Toyoda T. Fine mapping of the subunit binding sites of influenza virus RNA polymerase. *Microbiol Immunol*. 2002; **46**: 167-175.
http://www.ncbi.nlm.nih.gov/pubmed/12008925
182 Tamuri AU, Dos Reis M, Hay AJ, Goldstein RA. Identifying changes in selective constraints: host shifts in influenza. *PLoS Comput Biol*. 2009; **5**: e1000564.
http://www.ncbi.nlm.nih.gov/pubmed/19911053
183 Giria M, Rebelo de Andrade H. Genetic evolution of PB1 in the zoonotic transmission of influenza A(H1) virus. *Infect Genet Evol*. 2014; **27**: 234-243.
http://www.ncbi.nlm.nih.gov/pubmed/25092558
184 Giria M, Santos L, Louro J, Rebelo de Andrade H. Reverse genetics vaccine seeds for influenza: Proof of concept in the source of PB1 as a determinant factor in virus growth and antigen yield. *Virology*. 2016; **496**: 21-27.
http://www.ncbi.nlm.nih.gov/pubmed/27240145
185 Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc Natl Acad Sci U S A*. 2000; **97**: 6108-6113.
http://www.ncbi.nlm.nih.gov/pubmed/10801978
186 Reed LJM, H. A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT ENDPOINTS. *Am J Epidemiol*. 1938; **27**: 493-497.

187 CDC. CDC. CDCprotocolofrealtimeRTPCRforinfluenza A(H1N1). 2009.

188 WHO. WHO. WHOInformationforLaboratoryDiagnosisofPandemic(H1N1)2009(virus)in humans. 2009.

189 HPE. AgencyStandardoperatingProcedureforInfluenza MUNANANeuraminidase ActivityandInhibitionAssayLondon. 2006.
ohttp://www.isirv.org/site/ima
ges/stories/avg_documents/Methodology/munana_ic50_sop_for_external.
pdf4
190 Correia V, Santos LA, Giria M, Almeida-Santos MM, Rebelo-de-Andrade H. Influenza A(H1N1)pdm09 resistance and cross-decreased susceptibility to oseltamivir and zanamivir antiviral drugs. *J Med Virol*. 2015; **87**: 45-56.
http://www.ncbi.nlm.nih.gov/pubmed/25042157
191 Gilbertson B, Zheng T, Gerber M, et al. Influenza NA and PB1 Gene Segments Interact during the Formation of Viral Progeny: Localization of the Binding Region within the PB1 Gene. *Viruses*. 2016; **8**.
http://www.ncbi.nlm.nih.gov/pubmed/27556479
192 Gavazzi C, Yver M, Isel C, et al. A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proc Natl Acad Sci U S A*. 2013; **110**: 16604-16609.
http://www.ncbi.nlm.nih.gov/pubmed/24067651

193 Cobbin JC, Ong C, Verity E, Gilbertson BP, Rockman SP, Brown LE. Influenza virus PB1 and neuraminidase gene segments can cosegregate during vaccine reassortment driven by interactions in the PB1 coding region. *J Virol*. 2014; **88**: 8971-8980.
http://www.ncbi.nlm.nih.gov/pubmed/24872588
194 WHO. *Influenza sequencing primers and protocol.* 2009.
http://www.who.int/influenza/resources/documents/molecular_diagnosis_influenza_virus_humans_update_201108.pdf
195 Correia V, de Andrade HR, Santos LA, Lackenby A, Zambon M. Antiviral drug profile of seasonal influenza viruses circulating in Portugal from 2004/2005 to 2008/2009 winter seasons. *Antiviral Res*. 2010; **86**: 128-136.
http://www.ncbi.nlm.nih.gov/pubmed/20083142
196 Igarashi M, Ito K, Yoshida R, Tomabechi D, Kida H, Takada A. Predicting the antigenic structure of the pandemic (H1N1) 2009 influenza virus hemagglutinin. *PLoS One*. 2010; **5**: e8553.
http://www.ncbi.nlm.nih.gov/pubmed/20049332
197 WHO. *Recommended viruses for influenza vaccines for use in the 2011 southern hemisphere influenza season.* 2010.
http://www.who.int/csr/disease/influenza/201009 Recommendation.pdf
198 Kilander A, Rykkvin R, Dudman SG, Hungnes O. Observed association between the HA1 mutation D222G in the 2009 pandemic influenza A(H1N1) virus and severe clinical outcome, Norway 2009-2010. *Euro Surveill*. 2010; **15**.
http://www.ncbi.nlm.nih.gov/pubmed/20214869
199 Wu R, Zhang H, Yang K, et al. Multiple amino acid substitutions are involved in the adaptation of H9N2 avian influenza virus to mice. *Vet Microbiol*. 2009; **138**: 85-91.
http://www.ncbi.nlm.nih.gov/pubmed/19342184
200 Liu Y, Lou Z, Bartlam M, Rao Z. Structure-function studies of the influenza virus RNA polymerase PA subunit. *Sci China C Life Sci*. 2009; **52**: 450-458.
http://www.ncbi.nlm.nih.gov/pubmed/19471867
201 Tada T, Suzuki K, Sakurai Y, et al. NP body domain and PB2 contribute to increased virulence of H5N1 highly pathogenic avian influenza viruses in chickens. *J Virol*. 2011; **85**: 1834-1846.
http://www.ncbi.nlm.nih.gov/pubmed/21123376
202 Salahuddin P, Khan AU. Structural and functional analysis of NS1 and NS2 proteins of H1N1 subtype. *Genomics Proteomics Bioinformatics*. 2010; **8**: 190-199.
http://www.ncbi.nlm.nih.gov/pubmed/20970747
203 Guu TS, Dong L, Wittung-Stafshede P, Tao YJ. Mapping the domain structure of the influenza A virus polymerase acidic protein (PA) and its interaction with the basic protein 1 (PB1) subunit. *Virology*. 2008; **379**: 135-142.
http://www.ncbi.nlm.nih.gov/pubmed/18657841
204 Hayden FG, de Jong MD. Emerging influenza antiviral resistance threats. *J Infect Dis*. 2011; **203**: 6-10.
http://www.ncbi.nlm.nih.gov/pubmed/21148489
205 Deyde VM, Sheu TG, Trujillo AA, et al. Detection of molecular markers of drug resistance in 2009 pandemic influenza A (H1N1) viruses by pyrosequencing. *Antimicrob Agents Chemother*. 2010; **54**: 1102-1110.
http://www.ncbi.nlm.nih.gov/pubmed/20028826
206 Yen HL, Hoffmann E, Taylor G, et al. Importance of neuraminidase active-site residues to the neuraminidase inhibitor resistance of influenza viruses. *J Virol*. 2006; **80**: 8787-8795.
http://www.ncbi.nlm.nih.gov/pubmed/16912325
207 Monto AS, McKimm-Breschkin JL, Macken C, et al. Detection of influenza viruses resistant to neuraminidase inhibitors in global surveillance during the first 3 years of their use. *Antimicrob Agents Chemother*. 2006; **50**: 2395-2402.
http://www.ncbi.nlm.nih.gov/pubmed/16801417
208 Plant E. Gene Constellation of Influenza Vaccine Seed Viruses. In: Victor Romanowski, ed.: InTech 2013.

209 Partridge J, Kieny MP. Global production of seasonal and pandemic (H1N1) influenza vaccines in 2009-2010 and comparison with previous estimates and global action plan targets. *Vaccine*. 2010; **28**: 4709-4712.
http://www.ncbi.nlm.nih.gov/pubmed/20488262
210 Zhang W. Candidate pandemic H1N1 vaccine viruses and vaccine potency reagents - an update of the development and availability. . 2009.
http://www.who.int/immunization/sage/4.Zhang_SAGE_vaccine_virus_reagents_-_Final.pdf
211 NIAID. *Research on new vaccine technologies.* 2015.
http://www.niaid.nih.gov/topics/Flu/Research/vaccineResearch/pages/technologies.aspx
212 Ridenour C, Johnson A, Winne E, et al. Development of influenza A(H7N9) candidate vaccine viruses with improved hemagglutinin antigen yield in eggs. *Influenza Other Respir Viruses*. 2015; **9**: 263-270.
http://www.ncbi.nlm.nih.gov/pubmed/25962412

213 Fulvini AA, Ramanunninair M, Le J, et al. Gene constellation of influenza A virus reassortants with high growth phenotype prepared as seed candidates for vaccine production. *PLoS One*. 2011; **6**: e20823.
http://www.ncbi.nlm.nih.gov/pubmed/21695145
214 Rudneva IA, Timofeeva TA, Shilov AA, et al. Effect of gene constellation and postreassortment amino acid change on the phenotypic features of H5 influenza virus reassortants. *Arch Virol*. 2007; **152**: 1139-1145.
http://www.ncbi.nlm.nih.gov/pubmed/17294090
215 Giria MT, Rebelo de Andrade H, Santos LA, Correia VM, Pedro SV, Santos MA. Genomic signatures and antiviral drug susceptibility profile of A(H1N1)pdm09. *J Clin Virol*. 2012; **53**: 140-144.
http://www.ncbi.nlm.nih.gov/pubmed/22177272
216 Hoffmann E, Krauss S, Perez D, Webby R, Webster RG. Eight-plasmid system for rapid generation of influenza virus vaccines. *Vaccine*. 2002; **20**: 3165-3170.
http://www.ncbi.nlm.nih.gov/pubmed/12163268
217 Reed LJ MH. A symple method of estimating fifty percent end points. *American Journal of Epidemiology*. 1938; **27**: 493-497.

218 WHO. WHO information for laboratory diagnosis of pandemic (H1N1) 2009 virus in humans. 2009.

219 CDC. CDC protocol of realtime RTPCR for influenza A(H1N1). 2009.

220 Agency HP. Standard operating procedure for influenza MUNANA neuraminidase activity and inhibition assay London 2006.
http://www.isirv.org/site/images/stories/avg_documents/Methodology/munana_ic50_sop_for_external.pdf
221 Santos LA, Sola S, Rodrigues CM, Rebelo-de-Andrade H. Distinct kinetics and pathways of apoptosis in influenza A and B virus infection. *Virus Res*. 2015; **205**: 33-40.
http://www.ncbi.nlm.nih.gov/pubmed/26002021
222 Plant EP, Liu TM, Xie H, Ye Z. Mutations to A/Puerto Rico/8/34 PB1 gene improves seasonal reassortant influenza A virus growth kinetics. *Vaccine*. 2012; **31**: 207-212.
http://www.ncbi.nlm.nih.gov/pubmed/23116694
223 Wang L, Lee CW. Sequencing and mutational analysis of the non-coding regions of influenza A virus. *Vet Microbiol*. 2009; **135**: 239-247.
http://www.ncbi.nlm.nih.gov/pubmed/18986781
224 Abdoli A, Soleimanjahi H, Jamali A, et al. Comparison between MDCK and MDCK-SIAT1 cell lines as preferred host for cell culture-based influenza vaccine production. *Biotechnol Lett*. 2016; **38**: 941-948.
http://www.ncbi.nlm.nih.gov/pubmed/26945752
225 YP L. Evolution of the receptor binding properties of the
influenza A(H3N2) hemagglutinin. *PNAS*. 2012; **109**.

226 WHO. WHO PUBLIC HEALTH RESEARCH AGENDA FOR INFLUENZA Biannual Progress Review and Report 2010–2011. 2013.
http://www.who.int/influenza/resources/research/RA_Progress_Report_2010_2011.pdf?ua=1