Final degree project

ERYTHROPOIETIN GENE DOPING

Iolanda Mitjans Suriol

Main field:

Biochemistry and Molecular Biology

Secondary fields:

Physiology and Pathophysiology, Legislation and History

March 2018

Faculty of Pharmacy and Food Sciences

University of Barcelona

ACKNOWLEDGEMENTS

I absolutely appreciate the help from Jordi Segura Noguera and Sergi Coll Camenforte of Catalonian Anti-Doping Laboratory belonging to Doping Control Research Group, from the Institute of medical investigation (IMIM). I would really like to thank the Department of Biochemistry and Molecular Biology for accepting my proposal of gene doping. In particular, I am really grateful to my tutor for guide me during the project.

INDEX

1. ABSTRACT 1
2. INTEGRATION OF THE DIFFERENT SCOPES 2
3. INTRODUCTION 3
4. OBJECTIVES 4
5. MATERIALS AND METHODS 4
6. RESULTS 5
6.1 HISTORY AND LEGISLATION 5
6.1.1 History of gene therapy 5
6.1.2 History of blood and erythropoietin manipulation 6
6.1.3 Legislation of gene doping 7
6.2 ERYTHROPOIETIN 8
6.2.1 Erythropoiesis 8
6.2.2 Erythropoietin gene expression 9
6.2.3 EpoR 10
6.2.4 Adverse side effects of erythropoietin 11
6.3 EPO GENE DOPING 12
6.3.1 Types of Epo gene doping 12
6.3.1.1 Viral gene transfer 12
6.3.1.2 Regulation of Epo transgene expression 13
6.3.1.3 Non-viral gene transfer 14
6.3.1.4 Ex vivo gene transfer 15
6.3.2 Safety concerns related to gene doping 16
6.4 DETECTION OF EPO GENE DOPING IN BLOOD 17
6.4.1 Screening for blood parameters 17
6.4.2 Detection of transgenic Epo protein 18
6.4.3 Detection of immune response 18
6.4.4 Detection using transcriptomics 19
6.4.5 Detection of Epo cDNA 19
6.5 ADVANCES IN DETECTING EPO GENE DOPING 21
6.5.1 Design of the RM sequence 21
6.5.2 Contributions of this study 22
7. DISCUSSION: Potentiality of detection of Epo gene doping 23
8. CONCLUSIONS 25

9. BIBLIOGRAPHY..... 26
10. LIST OF ABBREVIATIONS..... 28

1. ABSTRACT

In the last 30 years, gene therapy has constituted a significant area of biomedical research, focused on improving the direct introduction of genetic material into the body by regulation of the activity of some deficient or harmful genes. Therefore, this method allows a nearly physiological and continuous production of some proteins, avoiding the periodic administration of external recombinant proteins. However, members of the international sports community are aware of the fact that elite sportsmen are seeking a physical advantage in competition misusing gene therapy, the following called gene doping.

According to World Anti-Doping Agency (WADA), gene doping includes "The use of polymers of nucleic acids or nucleic acid analogues, the use of gene editing agents designed to alter genome sequences and/or the transcriptional or epigenetic regulation of gene expression and the use of cells, whether they are genetically modified cells or not". Consequently, gene doping could potentially be used to increase muscle size and strength, enhance endurance, promote more rapid healing of sports injuries and reduce its associated pain. In this project, it has been selected erythropoietin for gene doping as it is considered an ideal candidate for its role in increasing endurance and enhance capacity to deliver oxygen to various tissues.

RESUM

En els últims 30 anys, la teràpia genètica ha constituït una àrea significant en la recerca biomèdica, centrada en millorar la introducció directa de material genètic en el cos per regular l'activitat de gens deficients o nocius. Per tant, aquest mètode permet una producció relativament fisiològica i continua de determinades proteïnes, evitant l'administració periòdica de proteïnes recombinants. Tanmateix, els membres de la comunitat internacional esportiva són conscients de que actualment els atletes d'elit estan buscant un avantatge físic en les competicions utilitzant la teràpia genètica, també anomenada dopatge genètic.

Segons l'agència antidopatge mundial (WADA), el dopatge genètic inclou "L'ús de polímers d'àcids nucleics o àcids nucleics anàlegs, l'ús d'agents que modifiquen els gens per tal d'alterar les seqüències genòmiques i/o la regulació de l'expressió genètica transcripcional o epigenèticament i també, l'ús de cèl·lules indistintivament si estan modificades genèticament com no". Per tant, el dopatge genètic es podria emprar potencialment per incrementar la massa muscular i la força, prolongar la resistència, promoure la cura més ràpida de lesions esportives i reduir-ne el dolor associat. En aquest treball, s'ha escollit l'eritropoetina per al dopatge genètic ja que es considera un candidat ideal per al seu rol en incrementar la resistència i la capacitat de distribuir oxigen a varis teixits.

2. INTEGRATION OF THE DIFFERENT SCOPES

This work integrated different educational fields. The main one was Biochemistry and Molecular Biology. Since the principal issue of this project was gene doping, it was very significant to study and understand interactions with erythropoietin receptor and transcriptional factors that influence its gene expression. It was also important to go into detail about the principal types of gene doping and the actual situation of detecting gene doping. Afterwards, this subject was vital to propose a model to standardise gene doping detection that could withstand legal scrutiny.

As secondary educational fields, Physiology and Pathophysiology, History and Legislation were included. Physiology and Pathophysiology was essential to understand the erythropoietin physiological role such as the pathway of erythropoiesis, from stem cells to erythrocytes. History was used to comprehend how important gene doping has been getting in the last thirty years and to realise the implicated difficulties in getting a success in gene doping. Regarding Legislation, WADA was the main focus as it is the anti-doping international agency. Furthermore, it was emphasised its pending success in implementing a standard, viable and economical method to detect gene doping.

3. INTRODUCTION

After only a short history of three decades, advances in gene technology and therapy have shown to have potential to treat human diseases. Despite this ability, gene therapy remains in the area of experimental medicine, and much clinical study is necessary to demonstrate efficacy and safety. However, progress in gene transfer technology could be misused to enhance athletic performance in sports, called gene doping. Its main attraction, as against traditional drug-based doping, lies in the apparent difficulty in detecting its use. It is assumed that both the transgene and the expressed protein would be indistinguishable from their endogenous equivalent (1). As a matter of fact, gene doping represents a threat to the integrity of sport, weaken the principles of fair play and can involve a potential harm in athletes, society and the environment (2).

To date, plenty of genes have been identified whose products may affect physical performance and, therefore, they become potential candidates for gene doping. In fact, many of these gene products have also been discovered to be linked to diseases in humans. This intensive research has been focused on developing gene therapy approaches for their treatment.

Additionally, most of genes develop different roles in the body. Given their complex biological functions, it would be essential to consider what other changes might take place when modulating one single gene, as it could be altered the function of other essential genes. Among the most relevant candidate genes for gene doping, erythropoietin is particularly interesting due to its role in rising endurance. The expressed protein stimulates erythropoiesis, increases blood oxygenation and oxygen delivery to tissues. Indeed, these pathways are beneficial to treat particular anemias using gene therapy (1).

As a matter of fact, it is required an updated review of the technologies of gene transfer, the genes with potential to influence physical performance, and advancements in legitimate gene therapy. In order to deal with this situation, it will be essential to study deeply the different techniques used, as they are getting more sophisticated and, consequently, harder to detect (2).

4. OBJECTIVES

The main aim of this project is to evaluate the current situation of gene doping in a global context. The key purposes could be classified in three:

1. Perform a bibliographical research of erythropoietin physiological role, such as its gene expression's regulation and interactions with erythropoietin receptor.

2. Define the concept of gene doping according to WADA, evaluate different gene doping methods and their detection, as well as describe adverse side effects related to its misuse.

3. Choose the most promising method to detect erythropoietin gene doping, as there is no available an official system yet. This analysis should be robust, viable and economic in order to be implemented by WADA.

5. MATERIALS AND METHODS

The first approach of gene doping was the book *Genetically modified athletes: biomedical ethics, gene doping and sport.* It has been consulted dozens of bibliographical sources found since the active search in databases like PubMed or Scopus. While developing this project it has been used scientific reviews and also original articles to redact. It has not been applied any restriction to publication date, but it has been given preference to more present papers. Occasionally, it has been researched some web pages, like U.S. National library of medicine, to clarify some concepts and to look up the definition of gene doping by WADA. Furthermore, Jordi Segura Noguera and Sergi Coll Camenforte from IMIM aided me in selecting the most promising method of detecting erythropoietin gene doping and stressed the importance of standardization.

6. RESULTS

6.1 HISTORY AND LEGISLATION

6.1.1 History of gene therapy

The first direct human gene therapy trial took place in 1974 to treat two patients suffering from hyperargininemia, an urea cycle disorder. It was believed that the wild-type Shope papilloma virus encoded the gene for arginase activity so the gene could be transferred by administering the virus intravenously. Unfortunately, the trial was unsuccessful as there was not a change in the arginine levels.

The first human gene therapy trial approved by the Food and Drug Administration (FDA) took place in 1990. It was treated a patient with adenosine deaminase (ADA) deficiency associated with severe combined immunodeficiency (SCID). It was administered an autologous *ex vivo* therapy with white blood cells, which became the first clearly successful gene therapy.

The result in *ex vivo* gene therapy has been genetic and phenotypic improvement of these diseases with normalization or significant improvement in immune parameters. Furthermore, the latest approach in gene therapy *ex vivo* is partial correction of the inborn genetic condition.

Unfortunately, these clinical trials have demonstrated potential risk, which may be inherently associated with the use of integrating retroviral vectors. The main viral vectors used in gene doping are explained in *6.3.1.1 Viral gene transfer*. Three children with X-linked inheritance combined immunodeficiency developed T-cell leukemia, because retrovirus inserted near to a proto-oncogene and one of them died. In another gene therapy trial, while using a recombinant adenovirus (rAd) to *in vivo* deliver a gene encoding for ornithine transcarbamilase, one patient died of an acute and uncontrollable reaction to the vector. The administration was direct to the liver through the hepatic artery, and caused intravascular coagulation and multi-organ failure.

Gendicine™ was the first gene therapy product for clinical use in humans. It was approved by the Chinese State Food and Drug Administration in 2003 for the treatment of head and neck squamous cell carcinoma. This treatment was based on an adenoviral gene delivery system that inserts p53 gene into tumor cells stimulating cell death. European Medicines Agency (EMA) recommended for the first time gene therapy product Glybera®, alipogene tiparvovec, for approval in the European Union in 2012. It was based on a recombinant adeno-associated viral (rAAV) vector and it was approved for the deficit of lipoprotein lipase (3).

Strimvelis® was the first *ex vivo* stem-cell gene therapy to be approved in Europe. This treatment consisted in hematopoietic progenitor CD34+ cells transduced with retroviral vector that encoded for the human adenosine deaminase cDNA sequence. This was used in patients with adenosine deaminase deficiency associated to severe combined immunodeficiency, that were not suitable to undergo bone-marrow transplant due to lack of matching donor. Nowadays, most gene clinical trials are currently in phase I or II with less than 3% of all trials in clinical development *(Table 1)* (3).

1974	• First human gene therapy: hyperarginemia
1990	• First gene therapy approved by FDA: ADA- SCID
2003	• First gene therapy product for clinical use: Gendicine®
2012	• First EMA recommendation of gene therapy: Glybera®
2016	• First *ex vivo* stem-cell gene therapy to be approved in Europe: Strimvelis®

Table 1: Timeline of gene therapy

6.1.2 History of blood and erythropoietin manipulation

The first evidence linking erythropoietin (Epo) genetic alteration to enhanced athletic performance was suspected of a skier in 1964. Later studies concluded that he had a natural mutation in Epo receptor gene that resulted in the generation of a greater number of erythrocytes and enhanced capacity to deliver oxygen to various tissues including muscles (4).

During the Olympic Games held at moderate altitude, 2250 m at sea level, altitude-induced blood adaptations, such as an increase in haemoglobin concentration, were considered responsible for athletes living at altitude winning most of the endurance races in 1968. Consequently, it was demonstrated that elevation in haemoglobin concentration and oxygen delivery improved performance. It was speculated that the first accounts of blood transfusions in sport were in the 1970s by athletes. The next form of Epo doping was recombinant human Epo (rHuEpo) in the 1990s. In 2000, it was demonstrated that rHuEpo administration also increased haematocrit as well as time to exhaustion. Nowadays, it is known that some athletes are using rHuEpo in combination with blood transfusions.

In 1990, erythropoietin was included on the list of prohibited substances by the International Olympic Committee as it was suspected its misuse, although no approved

test existed. As rHuEpo was directly undetectable at the time and as well as to protect the health of athletes, it was introduced the 50% haematocrit rule in 1997 by the International Cycling Union. So, any racing cyclist who had a haematocrit above 50% was declared ineligible and was excluded from the race. Despite this rule, WADA was created in 1999, an independent international agency. His mission was based on setting standards for doping detection and coordinating sports organisations and public authorities.

Although rHuEpo is structurally similar to endogenous erythropoietin, it was developed a direct urinary test that could differentiate isoforms in isoelectric profiles. This test became the first approved method for the direct detection of rHuEpo doping, even its limited detection of hours. The Athlete Biological Passport (ABP) was introduced in 2008 as a new tool to detect alteration of blood markers, which is explained in *6.4.1 Screening for blood parameters*. Nevertheless, ABP is not enough sensitive and some athletes are now using microdoses of rHuEpo in a range from 10 to 40 IU/kg body mass. These microdoses increase haemoglobin mass without large fluctuations according to ABP. To sum up, adapting to anti-doping methods by the athletes requires a constant developing and implementing new detection methods (5).

6.1.3 Legislation of gene doping

WADA is the international organization responsible for promoting, coordinating and monitoring the global fight against doping in sport in all its forms. In order to cope with doping, WADA has collected different substances and methods related to doping. Substances are classified as Non-approved substances, Anabolic agents, Beta-2 agonists, Hormone and metabolic modulators, Diuretics and masking agents and Peptide hormones, growth factors, related substances and mimetics. The classification of prohibited methods includes manipulation of blood and blood components, chemical or physical manipulation, and gene doping.

WADA has admitted that gene doping represents a threat to the integrity of sport and the health of athletes so it was forbidden in 2004. However, genetic modification is still not fully theorised in anti-doping policy and it is already causing problems in elite sportsmen. To handle with this unfair situation, WADA has organized workshops about gene doping since 2002. Moreover, in 2004, it was created an expert group on gene doping, which task is to study the latest advances in the field of gene therapy and the methods for detecting gene doping (6).

6.2 ERYTHROPOIETIN

Erythropoietin is a hormone involved in the proliferation and differentiation of erythrocyte and the maintenance of a physiological level of erythrocyte mass. It is a glycoprotein synthesized by the kidney in response to low blood oxygenation, among other factors. It has been widely studied since it was discovered his first therapeutic use with anemia. In fact, this molecule is an effective treatment for severe anemia associated with chronic kidney disease, acquired immune deficiency syndrome (AIDS) and chemotherapy of cancer. Nowadays, it is been studying in anemic patients with cardiac failure, like strokes, as a neuroprotective agent (7).

6.2.1 Erythropoiesis

Erythropoiesis is part of the haematopoiesis, which involves the production of mature cells in the blood and lymphoid organs. In periods of increased erythrocyte loss, due to haemolysis or haemorrhage, the production of erythrocytes increases. However, overproduction of erythrocytes does not occur, even in a severe loss of erythrocytes. Maturing erythroid progenitor cells expand in number and decrease in size. The first committed erythroid cell type forms characteristic colonies called a burst-forming unit-erythroid cell and further differentiate into colony-forming unit-erythroid cells. These cells begin synthesis of haemoglobin and differentiate into erythroblasts, which enucleate and form reticulocytes. After several days, mitochondria are degraded, reticulin declines, and the cells become mature red blood cells (RBC) (*Figure 1*). As erythrocytes lack DNA, they can neither divide nor alter gene expression in response to stimuli (7).

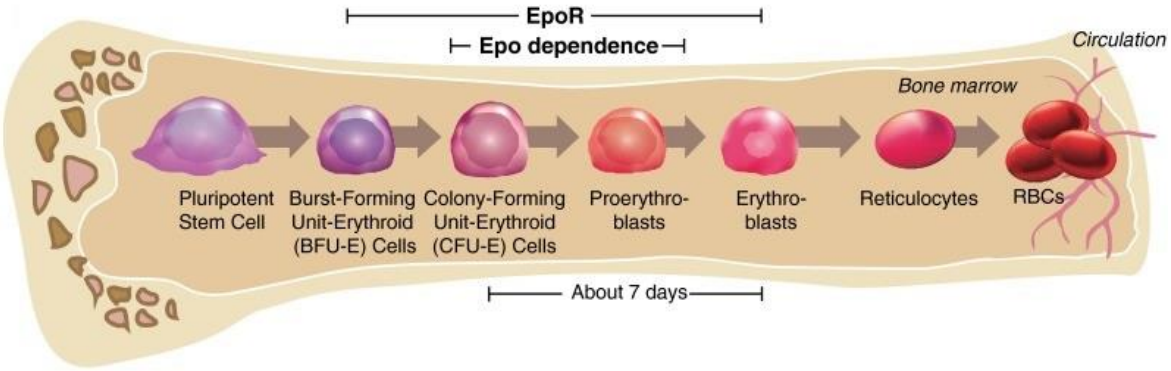

Figure 1: The process of erythropoiesis, adapted from (7)

Erythropoiesis occurs in specialized zones of bone marrow, surrounded with macrophages. In healthy humans, erythrocytes constitute 99% of circulating cells and approximately 45% of the blood volume. To sustain this level of RBC production, a 25% of the cells in a normal bone marrow are erythroid precursors. Although erythroid

precursors only represent a smaller proportion of 1%, its lifespan is 3–4 months under normal conditions, but it can be decreased in chronic kidney diseases (7).

6.2.2 Erythropoietin gene expression

The expression of the Epo gene is mainly in the liver encoded in chromosome position 7q22 and it is under the control of several transcription factors. GATA binding protein 2 (GATA-2) and nuclear factor kappa B (NF-κB) act and inhibit Epo gene expression on the 5' promoter. On the other hand, the main mechanism by which hypoxia stimulates the expression of the Epo gene is binding of hypoxia inducible transcription factor (HIF) *(Figure 2)*.

The hypoxia-inducible Epo enhancer, which is located on 3' of the Epo gene, contains two transcription factor binding sites. HIF binds the proximal site of the Epo enhancer downstream. HIF-α protein levels are controlled by HIF-prolyl hydroxylases (HIF-PH), enzimes that hydroxylate the α-subunit of HIF, targeting it for ubiquitination by the Von Hippel–Lindau protein and subsequent degradation by the proteasome. HIF-PH activity generally increases with high levels of oxygen, which led an augmented HIF protein levels and the rate of Epo production and, consequently, erythropoiesis also increases (8).

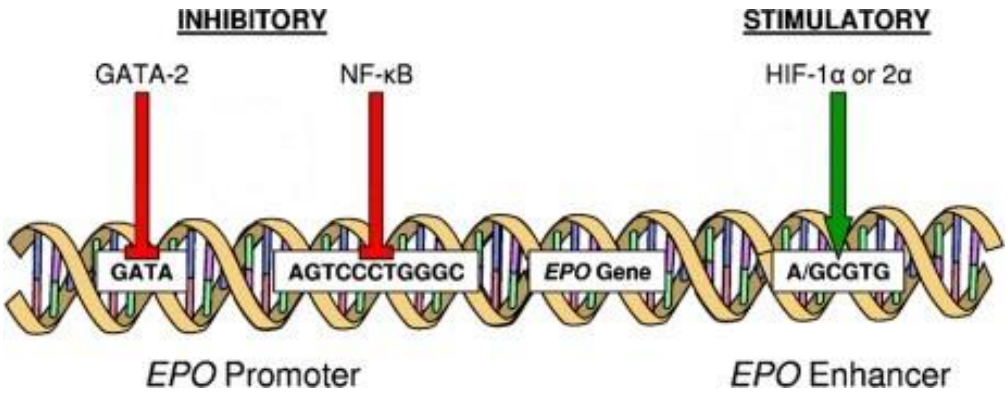

Figure 2: Transcriptional factors that stimulates or inhibates Epo gene expression (9)

Epo deficiency is the main cause of the anemia in chronic kidney disease and a contributing factor in the anemias induced for inflammation and cancer. There are some active compounds capable of stimulating endogenous Epo production in preclinical or clinical trials for treatment of anemia. These agents include stabilizers of the HIFs, which stimulate his expression through the Epo enhancer, and GATA inhibitors (8).

6.2.3 EpoR

Erythropoietin receptor (EpoR) is a type I transmembrane protein that belongs to the cytokine receptor superfamily and its principal function is regulation of erythropoiesis. Activation of EpoR is initiated by the direct binding of a single Epo molecule to two transmembrane EpoR proteins that form a homodimer on the surface of erythroid progenitor cells. The binding of Epo induces a conformational change in EpoR that makes the transmembrane and intracellular regions of the receptor get closer. Following binding, the Epo–EpoR complex is activated, internalized, and some is degraded in lysosomes, with the remainder recycled to the cell surface *(Figure 3).*

Moreover, EpoR requires a tyrosine kinase janus kinase 2 (JAK2), to induce the signaling cascade. JAK2, which interacts with EpoR at the juxtamembrane region, is transphosphorylation and consequently, activated. After JAK2 activation, JAK2 phosphorylates tyrosine residues in EpoR, which serve as docking sites for mediators of the signal transducer and activator of transcription 5 (STAT5) and phosphatidylinositol-3 (PI3) kinase/ protein kinase B (Akt) signaling pathways. Survival, proliferation and differentiation of erythroid progenitor cells are thereby stimulated (10).

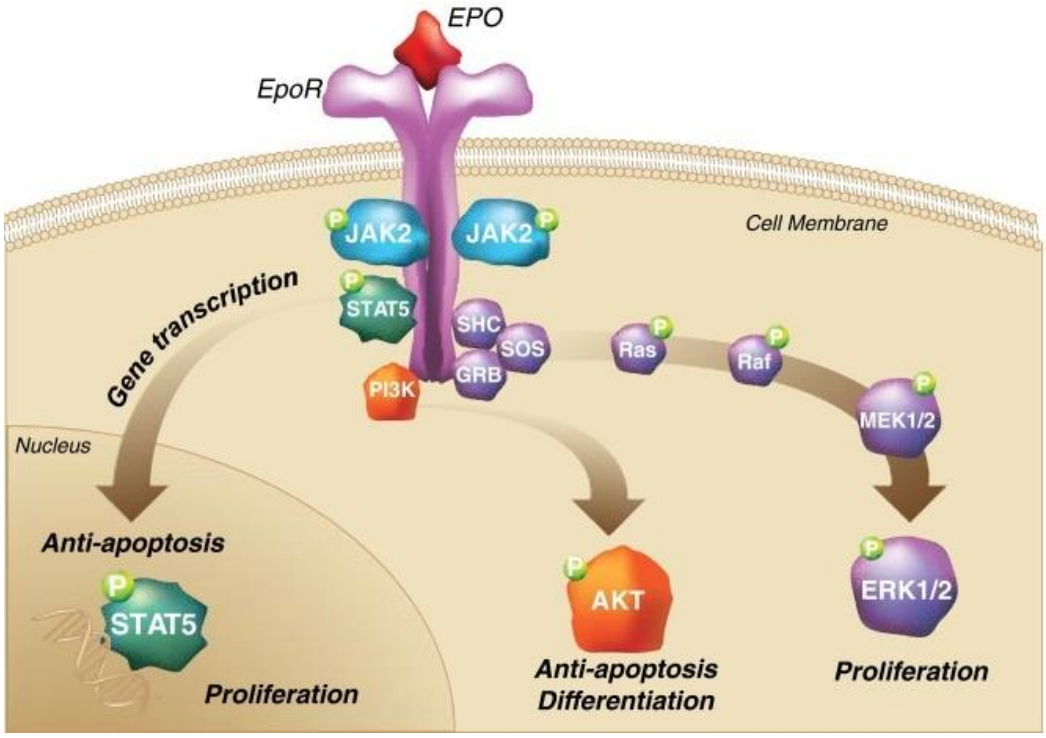

Figure 3: The signaling pathways stimulated by EpoR upon binding to Epo (10)

6.2.4 Adverse side effects of erythropoietin

Under normoxic physiological conditions, haemoglobin levels are regulated by the blood oxygen mainly via HIF. Epo and their derivatives increase the erythrocyte number and the transport capacity of oxygen, which increases blood viscosity and the probability of thromboembolic events. Besides increasing blood viscosity, long-term use of Epo can result in various side effects such as red cell aplasia and heart failure. In patients with an iron deficiency, Epo can elevate platelet counts and increase the risk of cardiovascular problems, including cardiac arrest, arrhythmia, hypertension, thrombosis, myocardial infarction and edema. Moreover, Epo is involved in angiogenesis, and his withdrawal may lead to lysis of young RBC called neocytolysis.

Otherwise, Epo has also been reported to have other effects, such as promotion of tumor cell growth or survival. One mechanism could involve the expression of functional EpoR in tumors or endothelial cells. Consequently, Epo directly stimulated tumor growth or antagonized tumor therapies (11).

6.3 EPO GENE DOPING

In contrast to gene therapy, the objective of gene doping is the overproduction of specific substances, hormones or enzymes which can enhance realisation of sport. The main method used is viral gene transfer. The gene is usually inserted with a carrier called vector, which are certain viruses that can deliver the new gene by infecting the cell. These viruses are modified in order not to cause any disease *6.3.1.1 Viral gene transfer*. The vector can be injected intramuscular or intravenously into a specific tissue in the body, where it is taken up by individual cells, where the vector injects new gene into nucleus and consequently the cell makes protein using this new gene (*Figure 4*). On the other hand, a sample of the patient's cells can be removed and exposed to the vector in a cell culture, after the cells containing the vector are returned to the patient; this is called *ex vivo* gene transfer (12).

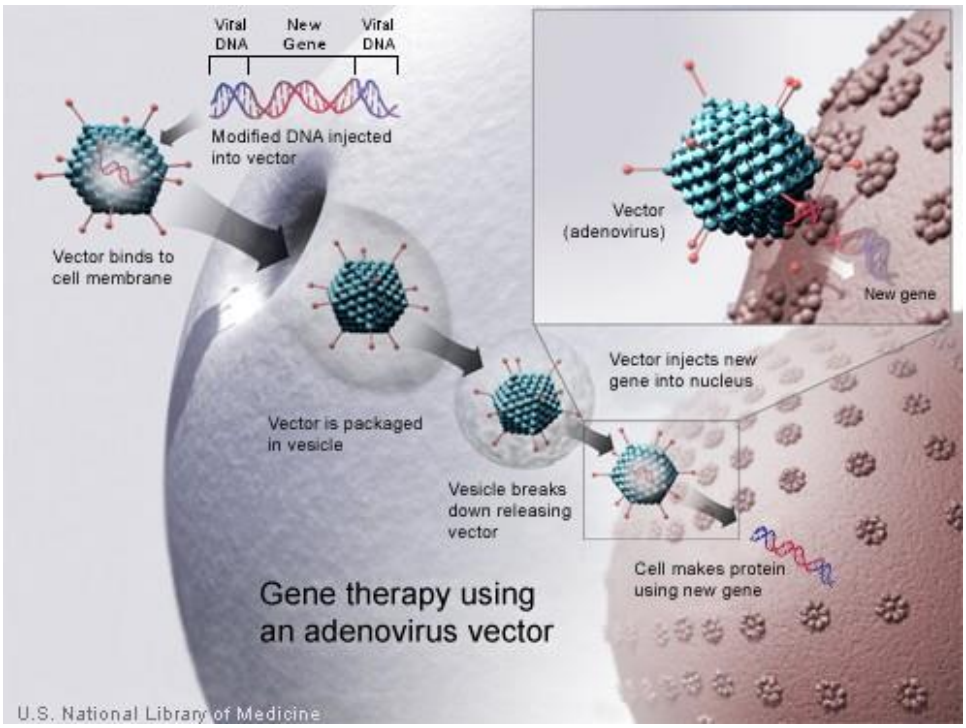

Figure 4: Scheme of gene therapy (12)

6.3.1 Types of Epo gene doping

6.3.1.1 Viral gene transfer

Viruses are vectors that get access to host cells and transfer their genetic material with high efficiency. Removing distinct viral genes allow viral replication to be hindered and genes can be inserted. The choice of virus depends on some features such as the packaging capacity of the recombinant vector or the chromosomal integration *(Table 2)*. The main drawback for transgene expression is the host's immune system, which can attack the viral vector system or the transgene product (13).

	Adenovirus	Adeno-associated virus	Retrovirus
Features of wild type virus	Non-enveloped double stranded DNA virus	Small non-enveloped single stranded DNA virus	Enveloped single stranded RNA virus (two copies)
Packaging capacity of the recombinant vector	30 kb	4,5 kb	8–10 kb
Chromosomal integration	Remains mainly episomal	Small part integrates into host cell genome (≈1%)	Integrates into the host cell genome of dividing cells
Risk of insertional oncogenesis	Small	Small	High risk
Long term gene expression	No	Yes	Yes
Advantages for Epo gene transfer	- 3rd generation shows less immunogenicity compared to further generations - Can be produced at high titres	- Low immunogenicity - High efficiency to transduce muscle cells	- Shows long term gene expression - Usually used *ex vivo*
Disadvantage for *in vivo* Epo gene transfer	Short term gene expression due to its remaining immunogenicity	Difficult to produce at high titres	Inability to infect non-dividing cells such as muscle cells.

Table 2: Features of viral vector systems, adapted from (12)

Following the initial success of viral Epo gene transfer in immune-deficient mice, more studies were conducted using adenovirus, adeno-associated virus, and retrovirus in rodents and non-human primates *(Table 3)*.

Due to the low immunogenicity of rAAV and its high efficiency in transducing muscle tissue, almost all studies employed rAAV vector systems for Epo gene delivery. It was shown that prolonged transgene expression and elevated haematocrit over the course of 600 days in non-human primates after a single i.m. injection of rAAV carrying the Rhesus macaques Epo gene. However, it was reported some immune responses targeting the endogenous Epo protein, leading to severe anemia and death. The complete mechanism is not already known, but it seems to be associated with high level production of Epo at ectopic sites (13).

6.3.1.2 Regulation of Epo transgene expression

A high number of regulatory systems have been developed to control or adjust transgene expression in order to avoid uncontrolled delivery of the Epo gene that leads to polycythemia with hemodynamic and rheologic problems. These systems include

antibiotics like doxycycline or immunosuppressives such as rapamycin, which only have found application in non-human primates studies. *(Table 3)* (13).

Using a Tetracycline On (Tet-On) regulatory system in non-human primates wherein administration of doxycycline led to high Epo secretion and increased haematocrit. Tet Response Element (TRE) is 7 repeats of a 19 nucleotide tetracycline operator sequence, and is recognized by the tetracycline repressor (tetR). Using a Tet-On system, if tetracycline is present, tetR will bind to tetracycline and not to the TRE, permitting transcription. However, it was found a humoral immune response directed against regulatory system proteins, in which was attacked genetically modified myofibers (14).

In contrast to the tetracycline dependent systems, the rapamycin regulated system contains exclusively human proteins, which should reduce the immune reaction. So, it was demonstrated long term expression of the Epo transgene in non-human primates for 6 years, which became the longest non-human primate study without immune response. However, it was described a persistent regulation of gene expression so, the degree of i.m. Epo gene transfer safety is still unclear and systematic administration is not possible yet (15).

Viral Epo gene transfer	Rodents	Non-human primates	Duration of the study
rAd	-	Cynomolgus macaques	84 days
rAAV	-	Rhesus macaques	>600 days
rAAV + Tet-regulated	-	Cynomolgus macaques	>5 years
rAAV + rapamycin regulated	-	Rhesus macaques	>6 years

Table 3: Selected Epo gene transfer studies using viral vector systems, adapted from (13)

6.3.1.3 Non-viral gene transfer

The most used vector system for non-viral gene delivery is plasmid DNA (pDNA). Compared to viral vectors, pDNA vectors are less immunogenic and they can be produced in large quantities easily at a reduced cost. However, the major limitation of *in vivo* use of pDNA is poor transfection efficiency and short duration of transgene expression (*Table 4*) (15).

Non-viral Epo gene transfer	Rodents	Non-human primates	Duration of the study
Naked pDNA	Balb/c mice	-	90 days
pDNA + HLV	-	Rhesus monkeys	>450 days
Modified mRNA	Balb/c mice	-	4 weeks

Table 4: Selected Epo gene transfer studies using non-viral gene transfer (13)

The best results in the field of non-viral Epo gene transfer were improved using the hydrodynamic limb vein (HLV). This method can deliver pDNA, viral and non-viral vectors to limb skeletal muscle. This procedure for pDNA delivery requires an injection of a large volume of DNA containing fluid. Although it was observed a high interindividual variability, animals reached normal physiological counts of RBC, which is indicative of a lack of autoimmune reaction against endogenous Epo.

In addition to DNA based gene delivery, it has also been reported the use of modified mRNA for Epo production. It was demonstrated that double modification of Epo mRNA significantly reduced binding to pattern recognition receptors and decreased activation of the immune system *in vivo* compared to unmodified RNA and prolongs transgene expression *in vivo*. mRNA interactions with various Toll-like receptors were reduced as well as activation of the innate immune system commonly associated with mRNA transcribed *in vitro* and *in vivo*. Moreover, the stability of mRNA increased *in vivo* and increase the haematocrit from 51% to 64% in 4 weeks. However, the efficiency of modified mRNA in non-human primates remains to be proven (13).

6.3.1.4 Ex vivo gene transfer

Ex vivo experiments have become other option in gene transfer, which is based on retrieving cells from a donor and genetically modifying these cells in culture. The first *ex vivo* gene therapies for Epo gene transfer was founded in the transplantation of myoblasts, smooth muscle cells, and fibroblasts. However, there were some limitations including immune response and poor survival of transduced cells *(Table 5)* (13).

On the other hand, 10 human patients with moderate chronic renal failure reimplanted subcutaneous 'Biopump-Epo' implants. These consist in virus-vector-treated with autologous subcutaneous tissue for continuous Epo production. Depending on the pre-determined *ex vivo* Epo secretion, it was released between 1 and 7 Biopumps. Following transplantation, Epo levels peaked on day 3 in most cases and declined until day 10 because it was reported an immune response without increasing the haematocrit. Moreover, the implant was removable and the pharmacological effect could be reversed.

It seemed that encapsulating cells in microcapsules would be the solution to avoid immune response. The membrane surrounding cells would allow diffusion of proteins, simultaneously preventing cell contact with immune competent host cells. Nevertheless, microcapsules have not achieved clinical significance although it has been tested different cell types and materials. The main drawbacks seemed to be related to an inflammatory response against the capsule material and the easily deterioration of the microcapsules (15).

Ex-vivo Epo gene transfer	Rodents	Primates	Duration of the study
Transplantation of myoblasts	Fischer 344 rats	-	80 days
Biopump-Epo implants	-	Homo sapiens	12 months
Myoblast cells secreting Epo within polymer microcapsules	Balb/c mice	-	100 days

Table 5: Ex-vivo Epo gene transfer

6.3.2 Safety concerns related to gene doping

Unfortunately, there are some problems related to gene doping *(Table 6)*. One drawback is the quality of the gene material, this needs to be purified in order to avoid contamination with wild-type viruses. There is a risk of possible mutagenesis, which consists in a definitive change in the cell's nucleic acid and the capacity to induce genetic modifications of chromosome structure, and consequently it could cause different diseases, like cancer or leukemia.

Moreover, it could be induced immunogenic reactions for the introduction of vectors, genes and the protein produced, which could slightly differ from the physiologically one. It could be an obstacle the modulation of gene expression, with the risk of excessive or insufficient production. This could originate a deregulation or an activation of an oncogene as there is a lack of control of transgene insertion.

Gene doping can affect the environment for the elimination of body fluids, which contain genetically modified viruses or their derivatives. As well as, it could be the integration of genetic material into germ cells, consequently it could be transmitted to following generations (16).

Safety concerns related to gene doping	
Quality of the gene material	Risk of mutagenesis
Induce immunogenic reactions	Change gene expression
Environmental risk	Integration into germ cells

Table 6: Summary of safety concerns related to gene doping

6.4 DETECTION OF EPO GENE DOPING IN BLOOD

Suitable gene doping detection methods are a major requirement for preventing athletes from using gene doping technologies. The detection method must be enough sensitive to detect any type of doping and samples must be easily accessible with minimally invasive techniques, since biopsies are not accepted. There are many potential strategies to detect the abuse of Epo gene transfer in relation to different levels of detection *(Table 7)* (17).

Level of detection	Type of detection	Problem	Solution
Doping effect: immune or blood parameters	Indirect	Specificity	Determine limit values
Transgenic protein	Direct	Homology to the natural protein	Detection of posttranslational differences
Transcriptome	Indirect	Specificity	mRNA reference levels
Transgenic DNA	Direct	Homology to the genomic DNA	Primers pairs for pre-amplification must only bind to the exon-junctions of cDNA

Table 7: Different types of Epo gene transfer detection.

6.4.1 Screening for blood parameters

Screening for blood parameters was the first detection of Epo gene doping. In 2009, WADA approved the ABP, which is based on monitoring athletes' biological variables over time to facilitate indirect detection of doping. They are evaluated blood parameters, such as the concentration of haemoglobin and reticulocytes and the subsequent enhancement of oxygen transport. It can be also used the diagram OFF Score to amplify changes observed in haemoglobin concentrations and percentage of reticulocytes. In *figure 5* it is shown an example of using Epo, ON phase, related to high percentage of reticulocytes before racing that compresses samples 3–7. Samples 8-10 show no use of this substance. Afterwards, there is a cessation of erythropoietic stimulation that leads to a prolonged suppression of reticulocytes and an elevated erythrocytes and a slightly increase in haemoglobin, this is OFF phase. In addition, variation in ABP haematological parameters due to training or hypoxia exposure can influence the interpretation of the ABP results. Nevertheless, sophisticated doping protocols enable athletes to continuously dope below the detectable threshold. Consequently, indirect detection methods should be replaced by direct detection methods that allow unequivocal identification of gene doping (18).

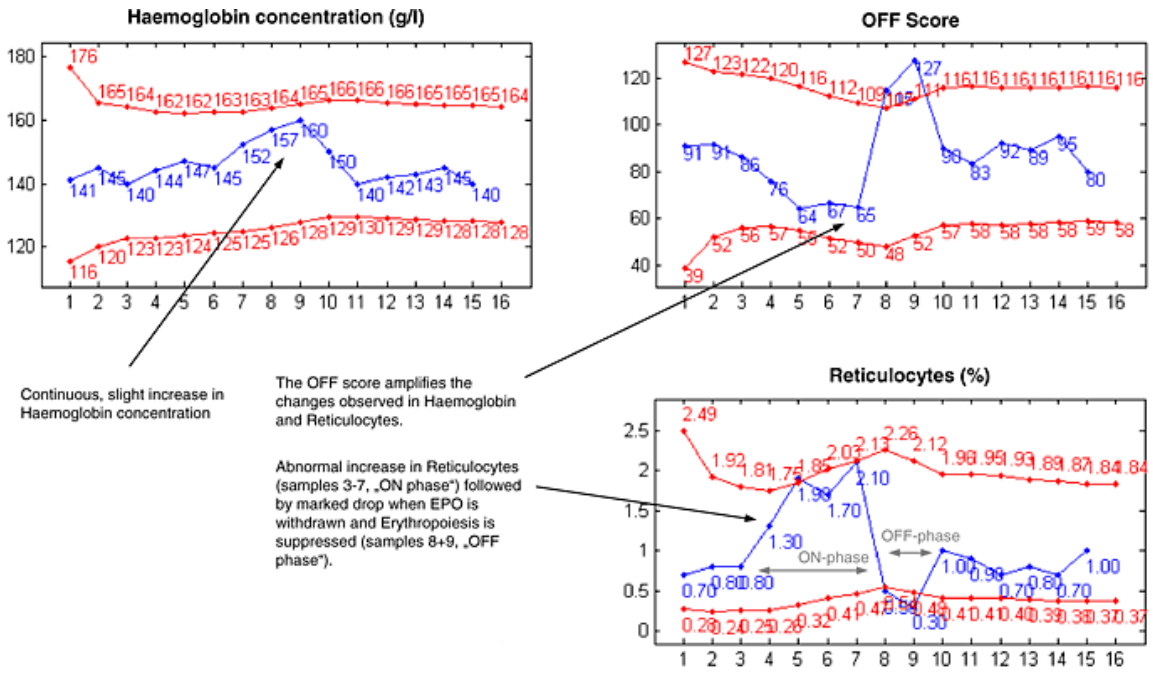

Figure 5: Haematological module of the ABP (18)

6.4.2 Detection of transgenic Epo protein

It was discovered that endogen, genomic DNA (gDNA) and proteins artificially encoded by transgenic DNA (tDNA) in muscle cells can be distinguished by a conventional Epo test consisting of double blotting and migration on isoelectric focusing. The difference in these Epo molecules is their glycosylation pattern due to different post-translational modifications in various tissues. However, that post transcriptional modifications may differ depending on the gene transfer protocol, the route of vector administration, the vector used, the target tissue and finally of course, the target species (13).

6.4.3 Detection of immune response

Other option to distinguish gene doping would be to identify specific immune responses to the vector system or the transgene protein. In fact, T-lymphocyte and antibodies against vector particles could be detected easily using Enzyme-linked immunosorbent assay (ELISA). However, adaptive cell-mediated and adaptive humoral immune responses seem drawbacks and it could be some false positive test in case of a natural viral infection. Unfortunately, viral vectors used have a high prevalence and incidence of natural infections. Furthermore, such detection procedures would also have limited use as non-viral mediated gene transfer is unlikely to produce any immune response (13).

6.4.4 Detection using transcriptomics

Abuse of gene doping can be detected screening the blood's transcriptome, as it can change gene expression patterns due to distinct influences such as diseases, exercise or the abuse of doping substances. Screening by microarrays allows defining specific biomarker or gene expression patterns. A potential advantage of the transcriptomic approach would be the ability to detect a wide range of Epo doping procedures, including all kinds of gene doping, as all of them share a common pathway following Epo-receptor activation. Furthermore, another approach would be to detect Epo mRNA expression at ectopic sites, which is indicative of gene transfer. However, the interindividual and intraindividual variations are drawbacks to validate potential biomarkers and to establish reference levels of mRNA because they are very similar to levels that would provide evidence for doping. However, it seems an alternative the use of micro RNA (miRNA), non-coding RNA molecules of approximately 22 nucleotides that modulate gene expression post transcriptionally, as a reliable biomarker. Nevertheless, knowledge about transcriptome and their variables is still limited and some athletes might carry an innate genetic feature or mutation, like an undiagnosed pathological condition, which could also alter their individual profile (13).

6.4.5 Detection of Epo cDNA

Other approach could be to target Epo cDNA using two quantitative nested qPCR (quantitative PCR) assays. It is combined a first round endpoint PCR of 25 cycles, with a second round of nested qPCR of 40 cycles. The product of the pre-amplification step is a linear molecule that is subsequently detected by qPCR. The nested qPCR assay is based on the strategy to pre-amplify five replicates per sample in the first round PCR. Afterwards, these samples are pooled and diluted before a second round qPCR. The establishment of a standard curve in the nested qPCR assay enables cDNA quantification.

The priming strategy of the two nested qPCR assays involves two assays using the same pre-amplification primers to generate a 437 bp linear amplicon. In the second round qPCR a 114 bp amplicon (Assay #1), and a 133 bp amplicon (Assay #2) are generated. The pre-amplification primers bind to the exon junction 1 and 2 and exon junction 4 and 5, respectively. Regarding nested qPCR, primers in Assay #2 targets the exon-exon boundary 2 and 3 and exon 4 and both of them uses a common probe ***(Figure 6)*** (19).

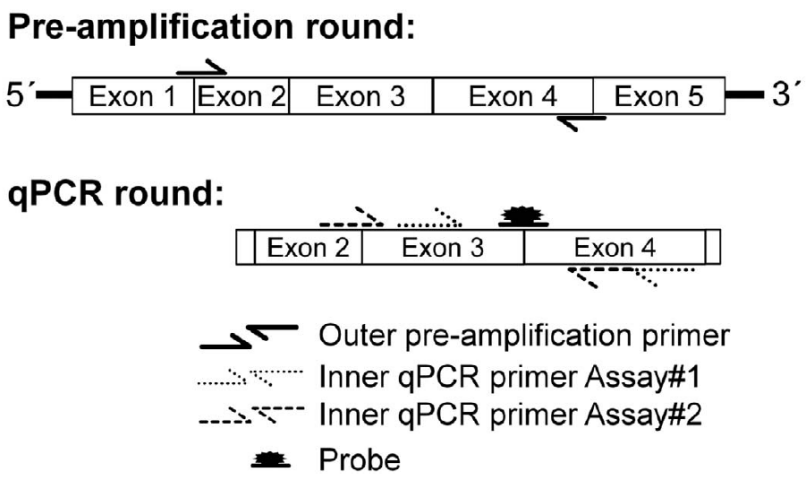

Figure 6: Pre-amplification round: Setup of the nested qPCR assay. 5 replicates of a sample undergo a pre-amplification round of 25 cycles. In the pre-amplification round both assays use the same primer pair, which binds exon junction 1 and 2 and exon junction 4 and 5, respectively. qPCR round: Priming strategy of the nested qPCR setup for the amplification of the human Epo cDNA sequence. In the qPCR round Assay#1 and Assay#2 use different primer pairs, whereas the same binding site for the probe is used (19).

The main requirement for pre-amplification round, was that both primers crossed an exon/exon junction. Its effectiveness may be compromised if Epo cDNA sequence is modified by insertion of small introns in a targeted exon junction or by site-directed mutagenesis of sequences for primers and/or probe annealing. Extensive modifications of Epo cDNA by changing all four exon/exon junctions, which would mask this transgene, it may also complicate its efficient transgene expression. The resultant from insertions of introns may be limited by viral vector packaging capacity. Also, as mRNA splicing is target tissue specific, the presence of introns may result in aberrant splicing when the transgene is expressed ectopically, as in the case of Epo expression in muscle rather than in its natural site of production, the kidney, potentially leading to a non-functional protein. It can also be hypothesized that the developed gene doping detection approach may be confounded by the presence of processed pseudogenes leading to a false-positive result. However, it has not been reported any pseudogenes for Epo human genome (20).

6.5 ADVANCES IN DETECTING EPO GENE DOPING

Analysis of vector genomes and transgene expression is typically performed by quantitative PCR (qPCR) using plasmid with transgenic sequences. Unfortunately, these methods differ between manufacturers, leading to inaccurate quantification or contaminations. To deal with these problems, in 2016, Baoutina et *al.* developed a method using synthetic certified DNA reference material (RM) to analyse human erythropoietin transgene. The authors elaborated a design strategy for synthetic RM with modified transgenic sequences to prevent false positives due to cross-contamination. When this RM was amplified in transgene-specific assays, the amplicons differed in size and sequence from transgene's amplicons. Afterwards, these differences could be established in post-PCR DNA fragment size analysis (DNA-FSA). In this study, it was used two vectors carrying the Epo transgene, non-viral, naked Epo pDNA and viral, Epo rAAV (21).

6.5.1 Design of the RM sequence

Achieving a unique synthetic RM suitable for vector-independent measurements of tDNA was particularly important in gene doping detection, as the nature of the vector used for gene transfer was unknown. There were compared three forms, a circular and a linear plasmid form and a shorter DNA fragment form, each one in viral and non-viral vectors.

The measurement system consisted in five validated qPCR assays targeting Epo cDNA. Epo RM incorporated synthetic reference sequence (RS) with five assays for reference sequence (ARS), one for each PCR assay. In each ARS design, there were sequences for binding the oligonucleotides of the assay. The sequences between these sites and the length of the amplicon from the ARS were different from those for the amplicon from the transgene. Afterwards, these differences could be established in DNA-FSA. Each ARS was designed by either removing several bases from the assay template in Epo cDNA, like in assays 1, 2 and 5, or inserting short sequences into the assay template, like in assays 3 and 4. The five ARS were *in silico* assembled into the RS together with three spacers *(Figure 7)* (21).

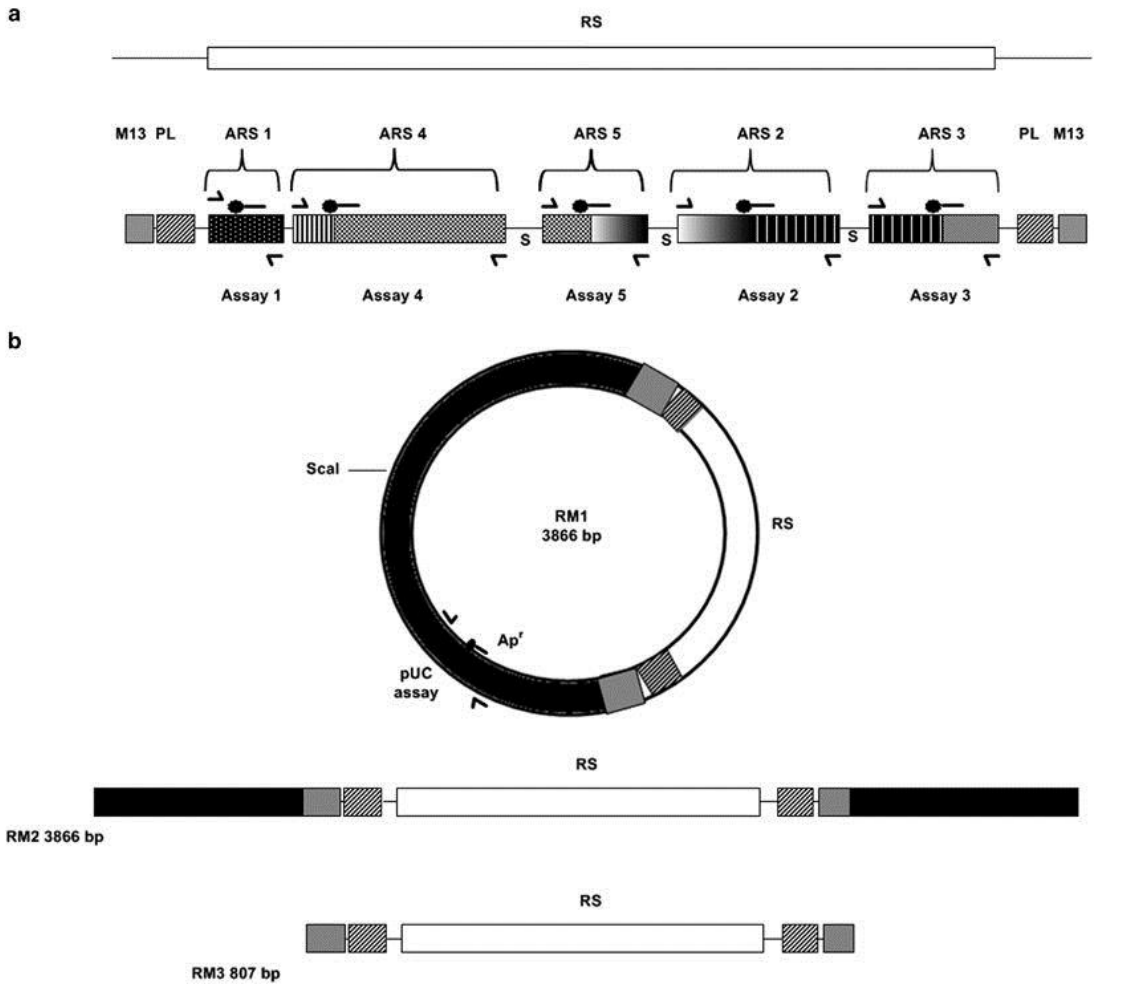

Figure 7: Diagram of the designed RM. (a) RS comprises five ARS for five Epo transgene-specific PCR assays and spacers (S), and is flanked by polylinkers (PL) and sequences for the M13 primers (M13). There is a similar pattern within sections of different ARS, which indicates complementarity to the same exon within Epo cDNA. (b) Three forms of the RM with approximate locations of the site for ScaI used to linearise the circular plasmid and of the plasmid pUC (pUC) assay. The oligonucleotides forming each assay are schematically shown as a one-sided arrow (primers) or a single line with a star head (probe). The bars representing different fragments, like ARS, S and PL (21).

6.5.2 Contributions of this study

This design strategy could serve as a prototype for development of measurement tools for other transgenes in order to achieve results comparability between laboratories. RM could facilitate implementation of a PCR-based analysis of genetic material, since gene doping until genetic disorders, as well as to determine dosage and monitor biodistribution. Moreover, it could be generated a RM with modified sequences from several transgenes, so that one RM could be used for analysis of multiple transgenes or vectors. Furthermore, this method could detect gene doping based on the analysis of transcriptomics biomarkers (21).

7. DISCUSSION

Potentiality of detection of Epo gene doping

While traditional banned doping substances or methods are easily detectable, detection of gene doping does not have an official method yet. Screening for blood parameters was the first detection of Epo gene doping implemented by WADA. Biological variables were monitored, such as the concentration of haemoglobin and reticulocytes and the subsequent enhancement of oxygen transport. The main drawback of this method was that these parameters can be influenced by training or hypoxia exposure.

The detection of vectors, even based on the immune response of the body to viral vectors, was often unable to discriminate between natural infection and artificial introduction of the virus. Moreover, it might not be possible to detect *ex vivo* gene transfer including Biopump or encapsulated cells because tDNA remains to transplanted cells and it is unlikely to spread to other cells or tissues. So, another approach would be to detect Epo mRNA expression at ectopic sites, which is indicative of gene transfer.

Identification in body fluids of the small molecules like antibiotics used as promoters of inducible gene activity provides indirect evidence of gene manipulation without medical treatment. However, some of these drugs are commonly used and they are not included in the WADA list of prohibited substances and methods. Direct detection of vectors or locally injected genes is only possible if the analysis is conducted early enough after administration, the local treatment site is known in the case of injection and the athlete accepts invasive procedures such as biopsy.

On the other hand, proteins encoded by gDNA and tDNA can distinguished by double blotting and migration on isoelectric focusing as they have different glycosylation pattern. This pattern is due to different post-translational modifications that may differ depending on the gene transfer protocol, the route of vector administration, the vector used, the target tissue and finally the target species.

Screening the blood's transcriptome allows to detect changes in mRNA levels compared with physiological levels. This quantification may be the main inconvenient as it would be require repeated measurements from gene expression patterns or specific biomarkers using microarrays. However, it would be an alternative the use of miRNA, nucleotides that modulate gene expression post transcriptionally, as a reliable biomarker. Nevertheless, knowledge of this field is still limited and some athletes might carry an innate genetic feature or mutation, like an undiagnosed pathological condition, which could also alter their individual profile.

Other approach could be to target Epo cDNA using two quantitative nested qPCR assays. It is combined a first round endpoint PCR and then a second round of nested qPCR of 40 cycles. The nested qPCR assay is based on the strategy to pre-amplify five replicates per sample in the first round PCR. Afterwards, these samples are pooled and diluted before a second round qPCR. The establishment of a standard curve in the nested qPCR assay enables cDNA quantification. The main requirement for pre-amplification round, was that both primers crossed an exon/exon junction. Its effectiveness may be compromised if Epo cDNA sequence is modified by insertion of small introns in a targeted exon junction or by site-directed mutagenesis of sequences for primers or probe annealing. It should be emphasized that extensive modifications of Epo cDNA by changing exon/exon junctions, would mask this transgene and also complicate its efficient expression.

Finally, a promising method of direct detection of Epo transgene performed by qPCR using synthetic certified DNA RM. When this RM was amplified in transgene-specific assays, the amplicons differed in size and sequence from transgene's amplicons. Afterwards, these differences could be established in post-PCR DNA-FSA. The main advantage to this design strategy was that it could serve as a prototype for development of measurement tools for other transgenes or transcripts in order to achieve results comparability between laboratories. RM could facilitate implementation of a PCR-based analysis of genetic material, from gene doping to genetic disorders. Moreover, it could be generated a RM with modified sequences from several transgenes, so that one RM could be used for analysis of multiple transgenes or vectors. Furthermore, this method could detect gene doping based on the analysis of transcriptomics biomarkers.

8. CONCLUSIONS

1. Nowadays, gene therapy is limited to some particular and serious diseases, while in the future it could be applied as a banned gene doping practice. In this situation, the sports medicine community will have to work closely with WADA in order to change and adjust legislation, particularly the genetic anti-doping rules.

2. Erythropoietin gene expression requires a signaling phosphorylation cascade, which stimulates pathways of anti-apoptosis, proliferation and differentiation of erythroid progenitor cells.

3. Adapting to anti-doping methods by the athletes requires a constant developing and implementing new detection methods. In order to ensure uniformity of results among laboratories, a method should be developed and standardised.

4. The chosen method used to detect erythropoietin gene doping is based on RM. This RM, used as an intern control, could facilitate implementation of a PCR-based routine test for Epo gene doping that could withstand legal scrutiny. Furthermore, the modified sequence design strategy can be easily adapted to generate synthetic nucleic acid RMs for analysis of any transgene.

9. BIBLIOGRAPHY

(1) Wells D. Gene doping: the hype and the reality. Br J Pharmacol. 2008; 154(3): 623-631.

(2) Miah A. Genetically modified athletes: biomedical ethics, gene doping and sport. 1st edition. London: Routledge Taylor & Francis Group; 2004.

(3) Carvalho M, Sepodes B, Martins A. Regulatory and scientific advancements in gene therapy: State-of-the-art of clinical applications and of the supporting european regulatory framework. Front Med. 2017; 4(182): 1-18.

(4) Azzazy H. Gene Doping. Handb Exp Pharmacol. 2010; 195: 485-512.

(5) Pitsiladis YP, Durussel J, Rabin O. An integrative 'Omics' solution to the detection of recombinant human erythropoietin and blood doping. Br J Sports Med. 2014; 48: 856-861.

(6) World anti-doping agency. What is prohibited [Internet]. Montreal: WADA; 2018 [updated 2018; access 10 november 2017]. Available in: https://www.wada-ama.org/en/content/what-is-prohibited/prohibited-at-all-times/gene-doping

(7) Elliott S, Sinclair A. The effect of erythropoietin on normal and neoplastic cells. Biologics. 2012; 6: 163–189.

(8) Jelkmann W. Novel erythropoietic agents: a threat to sportsmanship. Med Sportiva. 2007; 11(2): 32-42.

(9) Macdougall, I. New Anemia Therapies: Translating Novel Strategies From Bench to Bedside. Am J Kidney Dis. 2012; 59(3): 444-451.

(10) Elliott S, Sinclair A, Collins H, Rice L, Jelkmann W. Progress in detecting cell-surface protein receptors: The erythropoietin receptor example. Ann Hematol. 2014; 93(2): 181-192.

(11) Salamin O, Kuuranne T, Saugy M, Leuenberger N. Erythropoietin as a performance-enhancing drug: Its mechanistic basis, detection, and potential adverse effects. Mol Cell Endocrinol. 2017; 440: 1-13.

(12) U.S. National library of medicine. How does gene therapy work? [Internet]. Bethesda: National institutes of health; 2018 [updated 23 January 2018; access 17 november 2017]. Available in: https://ghr.nlm.nih.gov/primer/therapy/procedures

(13) Neuberger E, Jurkiewicz M, Moser D, Simon P. Detection of Epo gene doping in blood. Drug Test Anal. 2012; 4(11): 859-869.

(14) Addgene: the non-profit plasmid repository. Tetracycline Inducible Expression. [Internet]. Cambridge: Addgene; 2017 [updated 23 January 2017; access 15 desember 2017]. Available in: https://www.addgene.org/tetracycline/

(15) Baoutina A, Alexander I, Rasko J, Emsli K. Developing strategies for detection of gene doping. J Gene Med. 2007; 10(1): 3-20.

(16) Fischetto G, Bermon S. From gene engineering to gene modulation and manipulation: can we prevent or detect gene doping in sports? Sports Med. 2013; 43(10): 965-977.

(17) E Brzeziańska, D Domańska, A Jegier. Gene doping in sport – perspectives and risks. Biol Sport. 2014; 31(4): 251-259.

(18) Schumacher Y, Saugy M, Pottgiesser T, Robinson N. Detection of Epo doping and blood doping: the haematological module of the Athlete Biological Passport. Drug Test Anal. 2012; 4(11): 846-853.

(19) Neuberger E, Perez I, Le Guiner C, Moser D, Ehlert T, Allais M, et *al.* Establishment of two quantitative nested qPCR assays targeting the human Epo transgene. Gene Ther. 2016; 23: 330–339.

(20) Baoutina A, Coldham T, Bains G, Emslie K. Gene doping detection: evaluation of approach for direct detection of gene transfer using erythropoietin as a model system. Gene Ther. 2010; 17: 1022-1032.

(21) Baoutina A, Bhat S, Zheng M, Partis L, Dobeson M, Alexander IE, et *al*. Synthetic certified DNA reference material for analysis of human erythropoietin transgene and transcript in gene doping and gene therapy. Gene Ther. 2016; 23(10): 708–717.

10. LIST OF ABBREVIATIONS

ABP: Athlete biological passport

ADA: Adenosine deaminase

AIDS: Acquired immune deficiency syndrome

Akt: protein kinase B

ARS: Assays for reference sequence

DNA-FSA: DNA fragment size analysis

ELISA: Enzyme-linked immunosorbent assay

EMA: European Medicines Agency

Epo: Erythropoietin

EpoR: Erythropoietin receptor

FDA: Food and Drug Administration

GATA-2: GATA binding protein 2

gDNA: genomic DNA

HIF: Hypoxia inducible transcription factor

HIF-PH: HIF-prolyl hydroxylases

HLV: Hydrodynamic limb vein

IMIM: Institute of medical investigation

JAK2: Janus kinase 2

M13: M13 primers

miRNA: micro RNA

NF-κB: nuclear factor kappa B

pDNA: plasmid DNA

PI3: Phosphatidylinositol-3

PL: Polylinkers

pUC: Plasmid pUC

qPCR: quantitative PCR

rAAV: recombinant adeno-associated virus

rAd: recombinant adenovirus

RBC: Red blood cells

rHuEpo: recombinant human Epo

RM: Reference material

SCID: Severe combined immunodeficiency

RS: Reference sequence

STAT5: Signal transducer and activator of transcription 5

Tet: tetracycline

Tet-On: Tetracycline on

tetR: tetracycline repressor

TRE: Tet response element

tDNA: transgenic DNA

WADA: World anti-doping agency