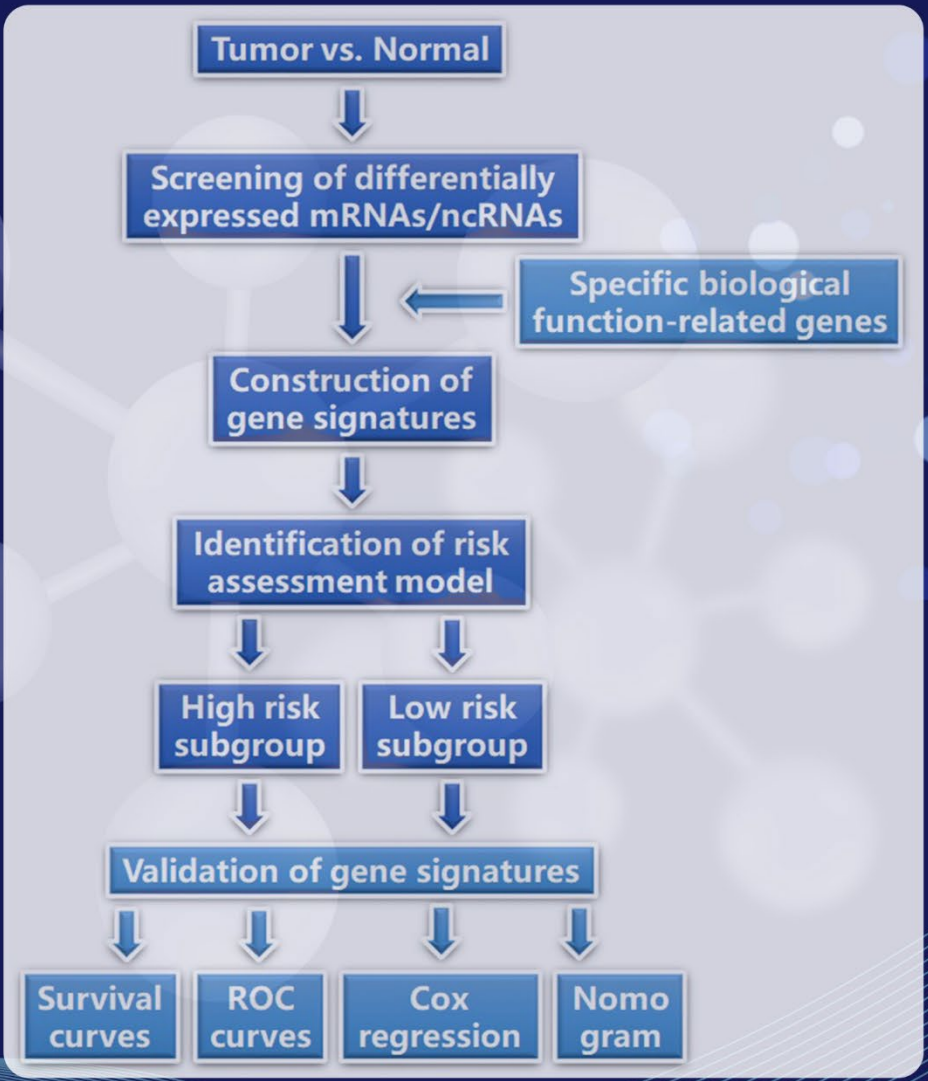
Gene & Protein in Disease

ACCSCIENCE PUBLISHING



Comprehensive Prognostic Signatures in Thyroid Cancer: A Summarized Review for Molecular Signatures Construction Strategies

ISSN: 2811-003X (Online)
Volume 2 · Issue 3
September 2023

Online ISSN: 2811-003X

Gene & Protein in Disease

Gene & Protein in Disease is an international journal for molecular and translational medicine. The journal primarily focuses on publishing investigations on the molecular bases and experimental therapeutics of human diseases.

Scan to access website:

Scan to submit papers:

About the Publisher

AccScience Publishing is a publishing company based in Singapore. We publish a range of high-quality, open-access, peer-reviewed journals and books from a broad spectrum of disciplines.

Contact Us

Managing Editor
gpd.office@accscience.sg

AccScience Publishing
8 Burn Road, #15-03 Trivex, Singapore 369977.

Volume 2 • Issue 3 • September 2023
ISSN 2811-003X (online)

GENE & PROTEIN IN DISEASE

ACCSCIENCE PUBLISHING

Access Science Without Barriers

GENE & PROTEIN IN DISEASE

ISSN: 2811-003X (online)

Editorial and Production Credits

Publisher: AccScience Publishing
Managing Editor: Juliana Meng
Production Editor: Ian Wong
Journal Development Editor: Felicia Wang
Special Issue Commissioning Editor: Felicia Wang
Article Layout and Typeset: Sinjore Technologies (India)
Cover Design: ProPub (China)

For all advertising queries, contact gpd.office@accscience.sg.

Supplementary file

Supplementary files of articles can be obtained at https://accscience.com/journal/GPD/2/3.

About the Cover

A graphic illustration of double-stranded DNA

Disclaimer

AccScience Publishing is not liable to the statements, perspectives, and opinions contained in the publications. The appearance of advertisements in the journal shall not be construed as a warranty, endorsement, or approval of the products or services advertised and/or the safety thereof. AccScience Publishing disclaims responsibility for any injury to persons or property resulting from any ideas or products referred to in the publications or advertisements. AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Gene & Protein in Disease

Editorial Board

ACCSCIENCE PUBLISHING

Volume 2 | Issue 3 | September 2023 |

Gene & Protein in Disease

https://accscience.com/journal/GPD

CONTENTS

ACCSCIENCE PUBLISHING

Gene & Protein in Disease

REVIEW ARTICLE

Control of mortality by human serum albumin

Andrew S. Johnson[1], and William Winlow[1,2]*

[1]Department of Biology, University of Napoli Federico II, 80138, Naples, Italy

[2]Institute of Ageing and Chronic Diseases, University of Liverpool, Liverpool, UK

Abstract

In a fit, resting human, efficient heart function depends on sufficient return of blood to the heart. At any stage of acclimatization, whether through changes in external environment or illness, pressures in the cardiovascular system must equilibrate to ensure adequate supply of nutrients to the deep capillaries, simultaneously providing adequate osmotic pressure to the cells and systemic pressure for venous return. About 80% of colloid osmotic pressure (COP) is controlled by levels of human serum albumin (HSA), which is the ultimate, homeostatic determinant of fluid volume (FV) in all compartments of the body and is controlled by the liver. The state of cardiovascular output is regulated by COP, which is determined by HSA colloidal pressure, which has clinical significance in many areas of medicine. In COVID-19 and other diseases like *Streptococcus*, where albumin binding has been suspected of being in deficit, we suggest that raising available HSA binding will alleviate vulnerabilities to disease and sepsis, unlike present forms of fluid therapy (FT). In this review, we describe how HSA binds and delivers nutrients as an intermediary transporter as it circulates, being especially important in the lymphatic system. Insufficient HSA nutrient binding leads to cell stress and albumin-binding deficiency (ABD) as both colloidal pressure and nutrients change lead to symptoms of sepsis. HSA binding can be reduced because of a lack of HSA or external ligands taking up binding sites on HSA. FV also affects long-term metabolic and cardiac diseases due to incorrect pressure and nutrients. The HSA lymphatic nutrient pump therefore utilizes HSA as an intermediate transporter of nutrient and waste ligands, and their circulation is determined by the lymph rather than the cardiac output. An elevation of the whole body colloidal pressure will increase the whole body FV and subsequently trigger production of all other nutrients in blood plasma and lymph: For example, body fluid increase will decrease blood oxygen to the kidneys promoting erythropoietin release culminating in the creation of new blood cells. We provide evidence that the hepatic portal vein (HPV) is essential in the precise control of HSA, glucose, and ketone metabolism and that the HPV is the appropriate target for FT and glucose control. We propose that in a clinical situation where this method is applicable, raising HSA will alleviate the symptoms of ABD and decrease the likelihood of serious illness and death by sepsis.

Keywords: Albumin; COVID-19; Diabetes; Fluid therapy; Sepsis

***Corresponding author:**
William Winlow
(bill.winlow@gmail.com)

Citation: Johnson AS, Winlow W, 2023, Control of mortality by human serum albumin. *Gene Protein Dis,* 2(3): 0328.
https://doi.org/10.36922/gpd.0328

Received: March 13, 2023
Accepted: August 10, 2023
Published Online: September 1, 2023

1. Introduction

During illness and old age, the systemic changes stemming from the loss of human serum albumin (HSA) binding (Figure 1) cause cumulative cellular changes[1-7]. These

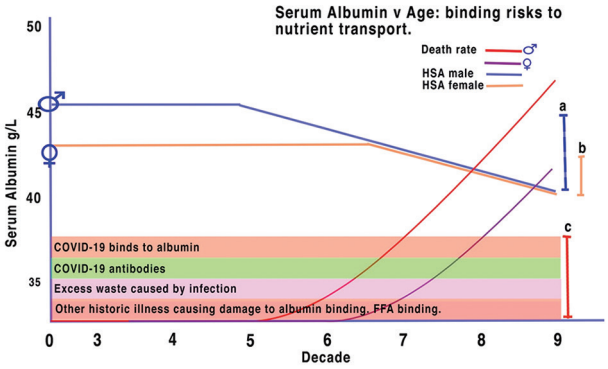


Figure 1. HSA levels are age-dependent and have a direct relationship to mortality due to COVID-19 and other illnesses[1,2,4,5]. HSA levels drop in males from the age of 50 and females at about 65, corresponding to an increased risk of illnesses; in this case, the risk of serious illnesses and death stems from COVID-19[1,2,4,5].

changes are reflected in the ability of the liver to maintain levels needed for systemic delivery of almost all nutrients, including gaseous exchange[2,3].

There are two main mechanisms for adjusting cellular health when illness occurs: Either remove the irritation or adjust the medium the cells are bathed in to maintain the best survival. In cases of COVID-19, influenza, and bacterial diseases, the conventional wisdom is to remove the offending virus or bacteria. This is usually done by a vaccine, which uses the body's own immune system to create antibodies to promote dissipation or antibiotics which destroy the bacteria *in situ*. The second method is to stabilize cells and the immune system by altering the medium in which they sit. This method must take into consideration a number of factors because the systemic system controlling individual organs and cellular structures is diverse, and during illness, normal vascular control is usually impeded. We suggest one method of increasing the immunity of the body is to improve nutrient supply and colloid oncotic pressure to the cells. In our previous papers, we covered evidence such as the age profile of has levels (Figure 1) and the possibility of raising HSA. This can be done by infusion of HSA to the circulation through the liver and increasing whole body fluids (WBF) to a level at which the cardiovascular/lymphatic system is most efficient. Colloid osmotic pressure (COP) has been found as a useful biomarker of fatal outcome in pregnant COVID-19 patients[8]. In our previous papers, we have already explained the relevance of HSA to pregnancy[1,2,4,5] that the pregnant mothers can become HSA binding-deficient leading to increased vulnerability because the fetus does not produce its own HSA.

In anesthesiology and surgery, maintaining correct WBF is essential for the efficiency of heart and lungs both in terms of pressure and nutrients[9]. Appropriate concentrations of cellular nutrients improve recovery from surgery and diabetes[10]. All systemic pharmaceuticals depend on not only the levels of WBF but also the binding constraints of HSA in determining the constitution of the blood, interstitial fluids, lymph, and other body fluids. A successful protocol to maintain HSA at set levels will alleviate diurnal and systemic variation, permitting control of glucose and ketone levels by allowing the precise maintenance of pancreatic and adrenal hormones, which also compete for transport by binding to HSA. As liver HSA production increases all other components of the blood change correspondingly, each component forms an equilibrium determined by its timing and distribution through the interstitial fluid and lymph. All other proteins and cellular distribution of blood components are maintained secondary to HSA including nutrients, red and white blood cell count, and the levels of hormones and waste. HSA through its control of colloidal pressure and WBF therefore underlies both the maintenance of blood pressure and the efficiency of the heart.

2. COVID-19 and albumin (HSA)

We have recently demonstrated the HSA lymphatic nutrient pump (HSALNP) and its critical importance to health (Figure 2)[1] and HSA (Figure 1)[11-13]. The HSALNP independently pumps nutrients by means of circulation of nutrient-bound HSA in the blood and lymph. HSA is the primary intermediate nutrient transporter. HSA production, circulation, and concentration therefore maintain the body's ability to adapt on acclimatization and injury. We defined the symptoms of sepsis as being caused by insufficient binding of HSA causing insufficient HSA-bound nutrients and colloidal pressure in the interstitial spaces flowing to the lymph[1,2]. Insufficient HSA binding can be caused by the lack of HSA or temporary occupation of binding sites on HSA (when COVID-19 virions and other immunoglobulins bind). Cellular damage is timing-dependent on essential nutrients, such as glucose being controlled and distributed by the liver to the deep cells of organs. HSA transport of proteins, for example, C-reactive protein (CRP), a known component of the endothelial glycocalyx, from the liver is common during infection[6,14,15], and endothelial glycocalyx degradation is a known factor in COVID-19 infection[15].

The HSALNP (Figure 2) is a circulatory system that should be considered separately from the conventional cardiovascular capillary circuit and is concerned with nutrition of the cells. The HSALNP circulates within the interstitial spaces to the lymph then to venous return during multiple-hour half-life. It is pressure- and hormone-dependent. HSA concentration is maintained

within narrow limits and any deviation from normal blood flow or ability to bind ligands will decrease nutrients in the HSALNP and cause albumin-binding deficiency (ABD) with corresponding changes in nutrients, colloidal pressure, and diastolic pressure. Nutrients travel around the body bound, with the most abundant binding protein being HSA. Sites of nutrient production are linked to targets so that sufficient nutrition is constantly maintained as in the regulation of glucose. Both the HSALNP and ABD are a logical explanation of the processes that must exist to

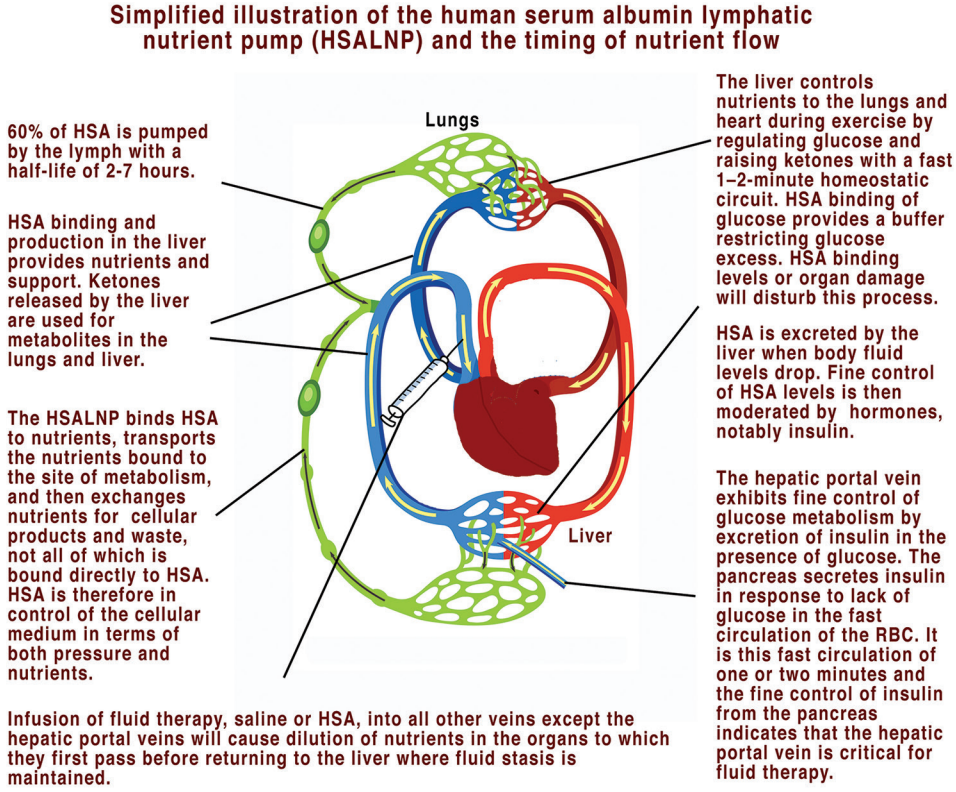


Figure 2. The human serum albumin lymphatic nutrient pump.

• Hypoalbuminemia is a known factor in sepsis patients[15-18] and COVID-19[18]. • Hypoalbuminemia is associated with inflammation[18-21] and is a marker for checkpoint blockade[22]. • Insulin production is a known determinant of HSA[23]. • HSA is implicated in recovery from post-cardiac arrest[24] and coronary heart disease[25]: • It is a useful predictor of ketoacidosis[26,27].
• Since early 2020, researchers have been reporting the relationship between COVID-19 and both intravascular HSA levels and hypoalbuminemia[28-38]. • Albumin oxidation in COVID-19 is known to occur[37] as is structural damage[38] and glycogenesis[39], in each case HSA binding is decreased. • Binding of ligands affects HSA levels: HSA binds to the COVID-19 spike 1 subunit and predicts in-hospital survival[40].
• The progress of HSA in sepsis and COVID-19 has been evaluated[38] and oxidative stress is known to occur during HSA deficiency[39]. • HSA levels may identify patients with SARS-CoV-2 infection in whom inflammatory processes are occurring and serve as a potentially useful marker of disease severity and prognosis[29-33].
• The existence of a mechanism, through the HSALNP that explains the kinetics of serious COVID-19 vulnerabilities, such as ABD, is described here.
• HSA is self-regulated through feedback from pressure within the hepatic portal vein – 80% of this pressure is from HSA. • HSA is therefore the primary determinant of body fluids as all other mechanisms of the blood regulate around this changing concentration. • In the 19th century, earnest starling investigated the effects of colloidal pressure on capillaries and understood that HSA as the primary colloid was the determinant of WBF[13,14] and noted the effect of colloids on nutrition. Starling understood the relationship between colloidal pressure and diastolic pressure.

Figure 3. Evidence that vulnerabilities and sepsis in COVID-19 are caused by human serum albumin (HSA) binding deficiency. Based on the evidence presented here, the regulation of HSA levels forms the basis of mean cardiovascular return and output when the body is at rest, with systolic/diastolic pressures being optimal. Any lowering of pressure causes production of HSA and an increase in plasma volume. Glucose and ketones are then corrected at this level as are other nutrient ligands, according to their own mechanisms.

explain nutrition of cells and failure during sepsis[16-18] and multiple organ failure as shown in Figure 3.

The HSALNP provides binding from the target (liver) to the cells, but ABD occurs when binding of systemic nutrients is impaired. Common manifestations of COVID-19 in children include non-dermatological age-related symptoms, such as multisensory inflammation syndrome. Age-related dermatological symptoms change from urticaria in children under 2 years old to chronic lymphocytic leukemia in the middle teens. Distribution of HSA is known to be extensive in the children's skin, which is a relatively larger organ than that in adults[1].

3. Raising HSA concentration to correct deficiencies that cause vulnerabilities in ABD

Gaseous exchange to cells takes place in the lungs, cells are supplied from the capillary network, with almost all gaseous distribution by direct diffusion. In contrast, metabolites are transferred through the cellular interstitial spaces determined by HSA binding and interstitial fluid and lymph flow. However, over 50% of HSA forms a secondary circulation, which has a half-life of many hours, from the leakage of HSA through the interstitial spaces into the lymph – this is the HSALNP. This secondary circuit is fed from the primary circuit which is maintained by HSA liver production and is dependent upon organ movement. In turn, metabolites such as glucose and ketones are controlled in the primary circuit by circulating hormones mainly from the adrenal glands (cortisol, aldosterone, adrenaline, and noradrenaline) and pancreas (insulin and glucagon). Long-term maintenance of glucose is therefore dependent on individual organ's use of glucose, organ usage reduces glucose levels which are then detected by the pancreas as the lymph returns to the plasma. Note that glucose levels are maintained in the fast primary cardiocapillary circuit and not the HSALNP with insulin flowing directly from the pancreas through the hepatic portal vein (HPV). The result is that liver control of glucose by insulin from the HPV and pancreas is maintained within this fast circulation every few minutes at most.

The ability to raise or decrease WBF has many clinical consequences in almost all areas of medicine. In ABD (Figure 4), an increase in albumin levels will gradually stabilize the HSA diurnal rhythm by settling variations in the cardiovascular capillary circulation and stabilizing cardiovascular output. Once HSA is raised to an appropriate concentration, ABD will be reversed as the chain of nutrients supplied by the liver and the complex equilibrium of nutrients in the plasma re-establish themselves. In the longer term, the 7-h half-life of the HSALNP, which supplies the interstitial spaces and deep cells with nutrients and maintains correct colloidal pressure, will resume providing interstitial spaces and cells as appropriate.

The HSALNP separates the cardiovascular system into different circulations according to the half-life of substances transported. Thus, conventionally gaseous respiration provides exchange in minutes to the cells from the lungs. The speed and distribution of all other components in the blood is controlled by the binding levels of albumin; 60% of which passes into the interstitial spaces and lymph, taking many hours to return to the liver. It is this second longer nutrient circulation, which has been overlooked, and the timing of which may be weeks. Concentrations of all other components of the blood respond to changes in HSA and normalize correspondingly.

4. Maintenance of physiologically normal HSA levels

In a healthy body, HSA production in the liver is regulated by pressure in the hepatocytes of the liver. Increased levels

- A decrease in HSA binding availability causes albumin-binding deficiency (ABD), whereby insufficient binding for nutrient affects oncotic pressure and cellular nutrients.
- ABDs are defined symptoms caused by lack of correct binding to HSA, resulting in illness caused by resulting nutrient deficiency and adverse oncotic pressure changes.
- Sepsis may well be an ABD similar to metabolic diseases and many cardiac symptoms[2,4,17,18,40].

- ABDs cause endothelial collapse, and the endothelial glycocalyx plays an important role in vascular homeostasis, regulating vascular permeability and cell adhesion. It has antithrombotic and anti-inflammatory functions and is bound within the glycocalyx, thus contributing to stability of the layer.
- Proteins that support the glycocalyx like C-reactive protein which are created in the liver are transported by HSA and become integral to the endothelial glycocalyx[6,15].
- HSA also transports sphingosine-1-phosphate, which has protective endothelial effects[5].

- The human serum albumin lymphatic nutrient pump (HSALNP) distributes nutrients selectively to organs, simultaneously maintaining oncotic pressure[1,2,4,5].

- Present regimens for saline and colloids do not consider physiological reality[4,5]. Infusion of the periphery can produce cellular damage.

- An appropriate procedure for fluid therapy is by direct infusion to the liver, and the most convenient route is by the hepatic portal vein (HPV) by ultrasound-guided core needle. Note the HPV is the only site of entry for both insulin and HSA where it is possible to maintain appropriate concentrations. Dilution of insulin or HSA by infusion at the periphery is predicted to cause insulin resistance as it is primarily fed to the periphery.

Figure 4. Albumin-binding deficiency and an appropriate procedure to overcome it.

of HSA produce an increase in ketone bodies from the liver, a preferential metabolite for the lungs, heart, and brain. During ketogenesis, any ketones not metabolized during illness continue to circulate in excess. Glucagon plays a limited role in ketogenesis during fasting or in response to SGLT2 inhibition[7]. The lungs are the primary site of plasma and lymph flow from the liver. The delivery of ketones to the lungs is immediate; an excess of ketones may produce ketoacidosis[10,44-47], while a deficit causes inefficiency in cellular function. A chain of events maintains the normal supply of ketones from the lungs and depends on albumin levels. Both ketones and glucose levels are also controlled by levels of insulin and glucagon secreted from the pancreas directly into the HPV. Pancreatic plasma is sampled for glucose levels from cardiac output and is primarily reactive to changes in glucose concentrations within one or 2 min from detection of plasma restricted to capillaries. However, plasma that flows into the lymph may take many hours in the interstitial lymph circulation. Glucose tolerance is thus one of fast (HPV) and slow correction (HSALNP) with present procedures for glucose stabilization and diabetes only referencing these mechanisms in combination.

The safety of raising HSA concentration relies on the timing of distributed nutrients through the interstitial cells and the constitution of the resulting lymph. Although the timing to infuse HSA to the cardiovascular capillary circulation takes a few minutes, the resulting flow through the interstitial-lymph circuit takes many hours, as mentioned above. Many nutrients such as glucose are only partly bound to HSA and their timing through the interstitial/lymph does not follow that of HSA. This means glucose variations follow discrete timings as lymph returns to the venous system slowly. Glucose levels are maintained mainly by insulin and glucagon excretion by the pancreas as well as adrenaline in times of stress. Insulin, bound to HSA, is maintained by concentrations in the HPV as it passes to the liver. Levels of glucose are therefore dependent on only the insulin levels in the HPV and not the rest of the circulatory system. This is reflected in the rapidity of glucose homeostasis. Therefore, in illness where the lymph flow is irregular, insulin measurements should ideally be taken from the HPV or arterial blood where a more direct relationship exists. In the liver, insulin promotes glycogenesis, converting glucose to glycogen for storage, thus leading to a reduction of blood sugar. This feedback loop that regulates glucose is entirely dependent on the cardiovascular capillary circulation, as the pancreas is primarily sampling glucose through this loop with changes in lymph occurring over the longer term depending on lymph flow. The interdependency of insulin, HSA, glucose metabolism, and levels of insulin can also be seen in the age variations in Type 2 diabetes in terms of both obesity and the age profile. In diabetes, hepatic production of serum albumin decreases, and it has been long established that insulin positively controls albumin gene expression[21]. Serious illness in COVID-19 occurs during the secondary phase when IgG3 cells are at their highest[37]. In addition, glycosylation of HSA decreases available HSA binding.

A chain of nutrient exchange therefore exists between the liver–lungs–heart and the rest of the body, such that release of proteins and nutrients of the liver are exchanged in the lungs before entering the capillary circulation of the periphery. The lungs and heart are both ketone metabolizers and therefore have a different dependency on HSA binding than the periphery. Maintaining this chain is critical for the health of the heart and lungs, especially during stress or illness. As blood flows from one organ to the next in series, binding by albumin and corresponding nutrient binders of nutrients change to reflect the correct medium for cell growth and cardiovascular efficiency. It is very important that this chain of binding and concentrations of metabolites is maintained so that individual organs operate within their correct nutrient medium and pressures.

Ketones, (released concurrently with HSA production), are preferentially metabolized by the lungs, heart, and brain during prolonged exercise or when glucose levels drop during illness[9,43,44,46]. Increased synthesis and use of ketone bodies as ancillary fuel during periods of deficient food supply, and low insulin levels causes oxidative stress in the mitochondria, which initiates a protective response allowing cells to cope with decreased energy availability[9,41,46]. In a normal healthy individual who has a sedentary lifestyle, ketone levels are balanced by the release of HSA. Exercise produces a relative drop in pressure in the liver, producing both nutrient-bound has and ketone bodies. In a healthy individual, any excess acetone from ketone metabolism is largely excreted by the lungs, thus avoiding ketoacidosis. In a respiratory-compromised individual, whose the lungs are infected by COVID-19, any increase in HSA and ketones will result in ketoacidosis and cellular death. An inverse relationship exists between HSA levels and ketosis for Type 2 diabetes[46]. As the release of ketones is dependent on natural release of HSA, any infused HSA will therefore lead to a drop in ketones as plasma volume increases. In acute respiratory distress syndrome, this may have a beneficial effect on increasing antioxidants and decreasing risk of ketoacidosis. If HSA is administered rapidly, both ketones and glucose may require adaptation, either by hormonal influence or direct infusion. The ketone-body metabolism is maintained by the anabolic hormone insulin and the primarily catabolic hormones, glucagon, cortisol, catecholamines, and growth

hormone[44-47]. The level of ketones should be assessed, ideally in the hepatic vein, arterial blood, and venous blood, to monitor their usage and production.

Albumin and magnesium concentrations are linearly related at high and low albumin concentrations[48,49]. Both magnesium and calcium are bound to HSA, with up to 80% of the calcium bound to HSA. The relationship between this competitive binding is well known, with the metabolism of magnesium and calcium, which are interdependent, closely related to the intestinal absorption and the renal excretion[48,50]. Plasma calcium concentration is maintained within a narrow range by the coordinated action of parathyroid hormone (PTH), $1,25(OH)_2D_3$, calcitonin, and ionized calcium (iCa^{2+})[46]. HSA is an intermediary that transport both calcium and magnesium, which compete for binding and dissolution in surrounding fluid. Transport of both calcium and magnesium is dependent on HSA levels for binding and appropriate oncotic pressures to adequately supply cells. Modified calcium and magnesium ratios are implicated in COVID-19 vulnerabilities[51,52].

Appropriate pressure regulation of cardiac function is dependent on the concentrations of body fluids and their distribution. Levels of overall pressure in the cells supplying nutrients are then maintained by repetitive heartbeats and colloidal pressure, which depends on HSA and maintains fluid volume to the capillary endothelia and interstitial spaces. For a healthy heart, sufficient blood must be returned to the heart with every filling of the atria so that the ventricles can operate within their limits as defined by the Frank-Starling effect. It is the liver and HSA production that defines WBF, blood volume, and content providing both a stable WBF and cardiovascular output within the limits of each individual physiology, by sampling pressure at its lowest in the HPV.

Serum albumin concentration is an important predictor of both baseline hemoglobin and erythropoietin sensitivity[53]. Erythropoietin released during sustained hypoxia leads to changes in hemoglobin following 6 weeks of acclimatization. Increasing WBF and blood volume will have an immediate effect on increasing blood flow, which may alleviate some sensitivity to HSA infusion. However, acute-phase reactants such as highly sensitive CRP[42,43], lactate dehydrogenase (LDH), ferritin, procalcitonin, interleukin 6, tumor necrosis factor-alpha, and interleukin 1 are elevated in most cases of sepsis, and early elevation is associated with high mortality if left untreated.

5. Discussion

Successful increase in HSA level depends on different pathologies of individuals and the technology to cannulate the HPV, evaluate, and correct nutrients and hormones. Core needle biopsy under ultrasound is now a regular procedure and has been used for the liver[54], and this technique can be used for safe cannulation of the HPV. The importance of the HPV has been shown during insulin clamp[55].

5.1. Lungs and COVID-19

The lungs are almost always the first point of infection in COVID-19. Any damage to the lungs affects the primary circuit of the HSALNP, affecting all metabolites in the blood. Metabolites controlled by the liver, such as glucose and ketones, are selectively metabolized according to the activity of the lungs. In damaged tissue, ketosis is a danger. Reduced respiration will reduce the level of acetone removal and promote ketoacidosis. In cases of lung damage, infusion of has to the liver should be to a well-ventilated supine subject, with at least ketone/glucose levels monitored and adjusted. Levels of adrenaline should be at least stable.

5.2. Albumin-binding deficiency

Here, we provide evidence that ABD is the primary cause of sepsis. In ABD where lung inflammation is not problematic, direct infusion to the HPV should adequately raise WBF and reduce vulnerabilities and sepsis. An HSA rise will affect all other components of the blood accordingly. In ill patients, timing is still critical and care should be taken as the effluent from the lymph mixes with that of the plasma. It is the various timings of ligand-HSA complexes passing through the interstitial spaces that result in lymph formation. The lymph flows back into the venous circuit and changes the components of the plasma. This process may take many hours through the interstitial spaces to the lymph, depending on the specific ligand and flow rate. The effect of all organs releasing lymph contributes to variations in concentrations, such as diurnal variations. With insufficient HSA or a depletion of binding potential caused by blocked HSA as in glycation, or competitive binding with immunoglobulins, nutrients may be exchanged before reaching the deep interstitial cells. This produces both colloidal and nutrient deficiency at the level of both the cells and the cardiovascular circulation. An increase of HSA concentration increases blood volume, and maintenance of that concentration should permit deep cells to be optimally perfused by HSA.

5.3. Suggested protocol for administration of HSA

We suggest monitoring glucose, ketones and corresponding hormones, insulin, and glucagon with the aim of maintaining a balance, where necessary, as close as the physiologically normal levels in the HPV. Both pancreatic and adrenal hormones are important in maintaining primary ketone/glucose balance during stress and must be

maintained at or near optimal as HSA rises in their point of action, the liver. HSA should be infused directly into the liver where it is formed naturally and where binding to metabolites and homeostasis originates. It is essential that HSA is infused to the liver and not into the periphery where unbound HSA causes stress through colloidal pressure change and nutrient loss. To ensure accuracy of timing for infusion, the HPV should be used, because any variations in concentration within the HPV are critical for those of both glucose and HSA. Ketones and glucose are the secondary variants in the cardiovascular circuit after HSA: Their concentrations are determined by the homeostasis of pancreatic and adrenal hormones, which can be modified through the HPV to ensure continuity.

The liver has highly complex mechanisms for moderating the content of the plasma's diverse ligands in the blood. Fundamental to this is the level of COP maintained by HSA. The liver is adaptable to huge variation in its production of HSA and metabolites. There is no evidence to suggest that healthy liver cannot produce, and recharge metabolites linked to infused HSA at any physiological levels. In addition, almost all systemic pharmaceuticals form ligands with albumin or other related protein binders like prealbumin are therefore affected by the HSA concentration. Changes in WBF will also change the output of ligand transport to cells by increasing cardiac efficiency.

1. HSA '∝' K
2. $In^{(G)}$ = ↑HSA ↑K ↓G ↓Fa
3. $gl^{(-in)}$ = ↑G ↑K
4. ep = ↑K ↑G ↓Fa
5. $co^{(-in)}$ = ↑K ↑G ↓Fa

K = ketones, G = glucose, Fa = fatty acids, in = insulin, gl = glucagon, ep = adrenaline, co = cortisol.

Figure 5. Hormones maintaining HSA, glucose, and ketone levels. These represent measured results taken from the references and they are non-linear reactions. (-in) indicates that insulin deficiency is required; for example, glucagon can increase ketogenesis acutely when insulin secretion is inhibited[14]. Ketone production and release of HSA and ketones are linked. (1) Ketone production increases proportionally to HSA levels. (2) Insulin increases HSA and ketone production in the presence of glucose, which is decreased with fatty acids. (3) Glucagon in the absence of insulin increases the levels of both glucose and ketones. (4) Adrenaline increases ketones and glucose with loss of fatty acids. (5) Cortisol in low insulin increases ketones and glucose with loss of fatty acids. In addition, as the HSA levels rise, the glucose concentrations fall as a consequence of the dilution of plasma fluid from colloidal pressure change and from the additional binding of glucose to HSA. A logical feedback mechanism that exists between HSA, glucose, and insulin can be manipulated to regulate the liver and by implication, the health of the body. The implication is that HSA can be raised appropriately when glucose and insulin are present. Rapid increase of HSA may precipitate adverse effects in pulmonary and heart disease due to the release of ketone bodies. To maintain or increase HSA levels to a new equilibrium, glucose should be present and insulin added to the HPV.

Increase in WBF caused by HSA infusion leads to changes in systemic hormones, which lead to concomitant changes in both glucose and ketone metabolism. Ketones are metabolized in the fast cardiovascular capillary circuit, which is important during exercise when the lungs and heart metabolize ketones under stress and on the release of adrenaline. An insulin surge in the HPV is followed by a decrease in glucose and increasing levels of HSA, ketones, and fatty acids. Glucagon can also act to increase glucose, and when insulin is deficient, ketone levels also increase. Adrenaline raises concentrations of ketones, glucose, and fatty acids. Cortisol like glucagon raises concentrations of ketones and glucose but decreases fatty acids when insulin is deficient. Note that the level of glucagon is controlled by the pancreas, and the cortisol level is controlled by a separate blood supply to the adrenal gland. Blood is therefore sampled from two diverse sites.

Deconstructing the actions of these hormones in respect to their sampling of the plasma demonstrates that HSA levels are also partially maintained from the glucose levels when insulin is present (Figure 5). HSA is primarily controlled by overall pressure in the HPV and this further attribute may have practical value for increasing HSA during illness. It may be possible to increase HSA naturally by fine-tuning levels of these hormones accurately and maintaining ketone and glucose levels.

5.4. Importance of HSA infusion into the HPV

The HPV is vital to the timely control of glucose, HSA, and ketones, all of which are maintained by sampling and correction of plasma within seconds or minutes. HSA and insulin should be administered so that the homeostatic process controlling HSA, glucose, and ketones is uninterrupted. We have already shown that the infusion through the periphery leads to incorrect colloidal pressures and nutrient deficits in the deep cellular structures. Infusion of insulin or HSA to the periphery reduces effective concentrations of these hormones in the liver and interferes with their respective binding in the plasma. Using the HPV, HSA equilibrium can theoretically be maintained. Maintaining HSA concentration within the physiological range should produce immediate benefits in patients with heart and lung diseases. In a healthy individual, the levels of ketone and glucose rise rapidly during exercise and as the lungs and heart metabolize them, the homeostasis of ketone and glucose can be maintained. It is this equilibrium and the balance between ketones, glucose, and the rest of the nutrients that maintain body health.

6. Conclusion

HSA concentrations affect every cell in the body and a reduction in HSA levels is directly implicated in many areas

of health. Evidence shows that raising HSA concentration by successful infusion into the HPV will have the beneficial consequences of stabilizing insulin and glucose levels, thereby reducing vulnerability to diabetes. There is ample evidence that a healthy body should respond to an increase of WBF by replacing all nutrients and blood cells in due time, optimizing cellular nutrients and oncotic pressure as well as delivery of nutrients into cells. Such optimization of the colloidal pressure of plasma will also optimize WBF, bringing diastolic and systolic pressures to optimal and ensuring optimal distribution of blood flow throughout the body. Clinical maintenance of WBF by HSA infusion to the liver will stabilize HSA variability and diurnal changes, permitting the fine control of metabolites, including glucose and ketones.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare no conflicts of interest.

Author contributions

Conceptualization: Andrew S. Johnson
Writing – original draft: Andrew S. Johnson
Writing – review & editing: All authors

All authors have read and agreed to the published version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Not applicable.

References

1. Johnson AS, Winlow W, 2022, The HSA lymphatic nutrient pump (HSALNP) and its critical and central importance to health. *EC Pharmacol Toxicol*, 10(12): 32–34.

2. Johnson AS, Winlow W, 2021, COVID-19 vulnerabilities are intensified by declining human serum albumin levels. *Exp Physiol*, 107: 674–682.

   https://doi.org/10.1113/EP089703

3. Levitt DG, Levitt MD, 2016, Human serum albumin homeostasis: A new look at the roles of synthesis, catabolism, renal and gastrointestinal excretion, and the clinical value of serum albumin measurements. *Int J Gen Med*, 9: 229–255.

   https://doi.org/10.2147/IJGM.S102819

4. Johnson AS, Fatemi R, Winlow W, 2020, SARS-CoV-2 bound human serum albumin and systemic septic shock. *Front Cardiovasc Med*, 7: 153.

   https://doi.org/10.3389/fcvm.2020.00153

5. Johnson AS, Polese G, Johnson M, *et al*., 2022, Appropriate human serum albumin fluid therapy and the alleviation of COVID-19 vulnerabilities: An explanation of the HSA lymphatic nutrient pump. *COVID*, 2: 1379–1395.

   https://doi.org/10.3390/covid2100099

6. Aldecoa C, Llau JV, Nuvials X, *et al*., 2020, Role of albumin in the preservation of endothelial glycocalyx integrity and the microcirculation: A review. *Ann Intensive Care*, 10: 85.

   https://doi.org/10.1186/s13613-020-00697-1

7. Capozzi ME, Coch RW, Koech J, *et al*., 2020, The limited role of glucagon for ketogenesis during fasting or in response to SGLT2 inhibition. *Diabetes*, 69(5): 882–892.

   https://doi.org/10.2337/db19-1216

8. Zavala-Barrios B, Cérbulo-Vázquez A, García-Espinosa M, *et al*., 2023, The Colloido-Osmotic Pressure as a useful biomarker of fatal outcome in pregnant COVID-19 patients. *medRxiv*, 2023.06.29.23292050.

   https://doi.org/10.1101/2023.06.29.23292050

9. Casey JD, Brown RM, Semler MW, 2018, Resuscitation fluids. *Curr Opin Crit Care*, 24(6): 512–518.

   https://doi.org/10.1097/MCC.0000000000000551

10. Laffel L, 1999, Ketone bodies: A review of physiology, pathophysiology and application of monitoring to diabetes. *Diabetes Metab Res Rev*, 15: 412–426.

    https://doi.org/10.1002/(sici)1520-7560(199911/12)15:6<412:aid-dmrr72>3.0.co;2-8

11. Weaving G, Batstone GF, Jones RG, 2016, Age and sex variation in serum albumin concentration: An observational study. *Ann Clin Biochem*, 53: 106–111.

    https://doi.org/10.1177/0004563215593561

12. Gom I, Fukushima H, Shiraki M, *et al*., 2007, Relationship between serum albumin level and aging in community-dwelling self-supported elderly population. *J Nutr Sci Vitaminol* (*Tokyo*), 53(1): 37–42.

    https://doi.org/10.3177/jnsv.53.37

13. Starling EH, 1896, On the absorption of fluids from the connective tissue spaces. *J Physiol*, 19: 312–326.

    https://doi.org/10.1113/jphysiol.1896.sp000596

14. Woodcock TE, Woodcock TM, 2012, Revised Starling equation and the glycocalyx model of transvascular fluid exchange: An improved paradigm for prescribing intravenous fluid therapy. *Br J Anaesth*, 108(3): 384–394.

    https://doi.org/10.1093/bja/aer515

15. Veraldi N, Vivès R, Blanchard-Rohner G, *et al.*, 2022, Endothelial glycocalyx degradation in multisystem inflammatory syndrome in children related to COVID-19. J *Mol Med (Berl*), 100: 735–746.

    https://doi.org/10.1007/s00109-022-02190-7

16. Sun JK, Sun F, Wang X, *et al.*, 2015, Risk factors and prognosis of hypoalbuminemia in surgical septic patients. *PeerJ*, 3: e1267.

    https://doi.org/10.7717/peerj.1267

17. Gupta L, James BS, 2012, 727: Hypoalbuminemia as a prognostic factor in sepsis, severe sepsis and septic shock. *Crit Care Med*, 40(12): 1–328.

    https://doi.org/10.1097/01.ccm.0000424942.33592.72

18. Takegawa R, Kabata D, Shimizu K, *et al.*, 2019, Serum albumin as a risk factor for death in patients with prolonged sepsis: An observational study. *J Crit Care*, 51: 139–144.

    https://doi.org/10.1016/j.jcrc.2019.02.004

19. Soeters PB, Wolfe RR, Shenkin A, 2019, Hypoalbuminemia: Pathogenesis and clinical significance. *JPEN J Parenter Enteral Nutr*, 43(2): 181–193.

    https://doi.org/10.1002/jpen.1451

20. Mantovani A, Garlanda C, 2023, Humoral innate immunity and acute-phase proteins. *N Engl J Med*, 388(5): 439–452.

    https://doi.org/10.1056/nejmra2206346

21. Seccombe A, Sapey E, 2018, What is the evidence base for fluid resuscitation in acute medicine? *Clin Med* (*Lond*), 18(3): 225–230.

    https://doi.org/10.7861/clinmedicine.18-3-225

22. Zheng M, 2022, Serum albumin: A pharmacokinetic marker for optimizing treatment outcome of immune checkpoint blockade. *J Immunother Cancer*, 10: e005670.

    https://doi.org/10.1136/jitc-2022-005670

23. Chen Q, Lu M, Monks BR, *et al.*, 2016, Insulin is required to maintain albumin expression by inhibiting forkhead box O1 protein. *J Biol Chem*, 291(5): 2371–2378.

    https://doi.org/10.1074/jbc.M115.677351

24. Lee H, Lee J, Shin H, *et al.*, 2022, Association between early phase serum albumin levels and outcomes of post-cardiac arrest patients: A systematic review and meta-analysis. *J Pers Med*, 12: 1787.

    https://doi.org/10.3390/jpm12111787

25. Kuller LH, Eichner JE, Orchard TJ, *et al.*, 1991, The relation between serum albumin levels and risk of coronary heart disease in the Multiple Risk Factor Intervention Trial. *Am J Epidemiol*, 134(11): 1266–1277.

    https://doi.org/10.1093/oxfordjournals.aje.a116030

26. Karthikeyan J, Rajaragupathy S, 2018, Diagnostic usefulness of serum albumin as a predictor of diabetic ketoacidosis. *Indian J Crit Care Med*, 22(10): 733–736.

    https://doi.org/10.4103/ijccm.ijccm_91_18

27. Kheir M, Saleem F, Wang C, *et al.*, 2021, Higher albumin levels on admission predict better prognosis in patients with confirmed COVID-19. *PLoS One*, 16(3): e0248358.

    https://doi.org/10.1371/journal.pone.0248358

28. Ramadori G, 2020, Hypoalbuminemia: An underestimated, vital characteristic of hospitalized COVID-19 positive patients? *Hepatoma Res*, 6: 28.

    https://doi.org/10.20517/2394-5079.2020.43

29. Islam N, Shkolnikov VM, Acosta RJ, *et al.*, 2021, Excess deaths associated with Covid-19 pandemic in 2020: Age and sex disaggregated time series analysis in 29 high income countries. *BMJ*, 373: n1137.

    https://doi.org/10.1136/bmj.n1137

30. Zekri-Nechar K, Zamorano-León JJ, Segura-Fragoso A, *et al.*, 2022, Albumin binds COVID-19 spike 1 subunit and predicts in-hospital survival of infected patients-possible alteration by glucose. *J Clin Med*, 11(3): 587.

    https://doi.org/10.3390/jcm11030587

31. Turcato G, Zaboli A, Kostic I, *et al.*, 2022, Severity of SARS-CoV-2 infection and albumin levels recorded at the first emergency department evaluation: A multicentre retrospective observational study. *Emerg Med J*, 39: 63–69.

    https://doi.org/10.1136/emermed-2020-210081

32. Aziz M, Fatima R, Lee-Smith W, *et al.*, 2020, The association of low serum albumin level with severe COVID-19: A systematic review and meta-analysis. *Crit Care*, 24: 255.

    https://doi.org/10.1186/s13054-020-02995-3

33. Violi F, Cangemi R, Romiti GF, *et al.*, 2021, Is albumin predictor of mortality in COVID-19? *Antioxid Redox Signal*, 35: 139–142.

    https://doi.org/10.1089/ars.2020.8142

34. Xie C, Wang S, Zhou J, *et al.*, 2022, Albumin level as an independent predictive factor for adverse outcomes in COVID-19 patients: A retrospective cohort study. *J Infect Dev Ctries*, 16: 1696–1702.

    https://doi.org/10.3855/jidc.16880

35. Abdeen Y, Kaako A, Amin ZA, *et al.*, 2021, The prognostic effect of serum albumin level on outcomes of hospitalized COVID-19 patients. *Crit Care Res Pract*, 2021: 9963274.

    https://doi.org/10.1155/2021/9963274

36. Prudente R, Franco E, Machado L, *et al*., 2022, Urea/albumin ratio associated with mortality in hospitalized COVID-19 patients: A cohort study. *Eur Respir J*, 60(Suppl 66): 1460.

https://doi.org/10.1183/13993003.congress-2022.1460

37. Wybranowski T, Napiórkowska M, Bosek M, *et al*., 2022, Study of albumin oxidation in COVID-19 pneumonia patients: Possible mechanisms and consequences. *Int J Mol Sci*, 23(17): 10103.

https://doi.org/10.3390/ijms231710103

38. Georgieva E, Karamalakova Y, Arabadzhiev G, *et al*., 2022, Site-directed spin labeling EPR spectroscopy for determination of albumin structural damage and hypoalbuminemia in critical COVID-19. *Antioxidants* (*Basel*), 11: 2311.

https://doi.org/10.3390/antiox11122311

39. Iles RK, 2023, The COVID-19 antibody responses, isotypes and glycosylation: Why SARS-CoV-2 spike protein complex binding of IgG3 is potentiated in some and immuno-pathologies manifest. *medRxiv*, 2023.01.13.23284524

https://doi.org/10.1101/2023.01.13.23284524

40. Siddiqui SS, Chakraborty N, Muzaffar SN, *et al*., 2022, Albumin kinetics in sepsis and COVID-19. *Crit Care Explor*, 4(3): e0651.

https://doi.org/10.1097/CCE.0000000000000651

41. Badawy M, Yasseen BA, El-Messiery RM, *et al*., 2021, Neutrophil-mediated oxidative stress and albumin structural damage predict COVID-19-associated mortality. *Elife*, 10: e69417.

https://doi.org/10.7554/eLife.69417

42. Sproston NR, Ashworth JJ, 2018, Role of C-reactive protein at sites of inflammation and infection. *Front Immunol*, 9: 754.

https://doi.org/10.3389/fimmu.2018.00754

43. Akkececi NS, Cetin GY, Gogebakan H, *et al*., 2019, The C-reactive protein/albumin ratio and complete blood count parameters as indicators of disease activity in patients with Takayasu arteritis. *Med Sci Monit*, 25: 1401–1409.

https://doi.org/10.12659/msm.912495

44. Mey JT, Hari A, Axelrod CL, *et al*., 2020, Lipids and ketones dominate metabolism at the expense of glucose control in pulmonary arterial hypertension: A hyperglycaemic clamp and metabolomics study. *Eur Respir J*, 55(4): 1901700.

https://doi.org/10.1183/13993003.01700-2019

45. Kolb H, Kempf K, Röhling M, *et al*., 2021, Ketone bodies: From enemy to friend and guardian angel. *BMC Med*, 19: 313.

https://doi.org/10.1186/s12916-021-02185-0

46. Cheng PC, Hsu SR, Cheng YC, 2016, Association between serum albumin concentration and ketosis risk in hospitalized individuals with Type 2 diabetes mellitus. *J Diabetes Res*, 2016: 1269706.

https://doi.org/10.1155/2016/1269706

47. Alberti KG, Johnston DG, Gill A, *et al*., 1978, Hormonal regulation of ketone-body metabolism in man. *Biochem Soc Symp*, 43: 163–182.

48. Kroll MH, Elin RJ, 1985, Relationships between magnesium and protein concentrations in serum. *Clin Chem*, 31(2): 244–246.

https://doi.org/10.1093/clinchem/31.2.244

49. Paunier L, 1992, Effect of magnesium on phosphorus and calcium metabolism. *Monatsschr Kinderheilkd*, 140(9 Suppl 1): S17–S20.

50. Jeon US, 2008, Kidney and calcium homeostasis. *Electrolyte Blood Press*, 6(2): 68–76.

https://doi.org/10.5049/ebp.2008.6.2.68

51. Guerrero-Romero F, Mercado M, Rodriguez-Moran M, *et al*., 2022, Magnesium-to-calcium ratio and mortality from COVID-19. *Nutrients*, 14(9): 1686.

https://doi.org/10.3390/nu14091686

52. Coman AE, Ceasovschih A, Petroaie AD, *et al*., 2023, The significance of low magnesium levels in COVID-19 patients. *Medicina* (*Kaunas*), 59: 279.

https://doi.org/10.3390/medicina59020279

53. Agarwal R, Davis JL, Smith L, 2008, Serum albumin is strongly associated with erythropoietin sensitivity in hemodialysis patients. *Clin J Am Soc Nephrol*, 3(1): 98–104.

https://doi.org/10.2215/CJN.03330807

54. Johnson KD, Laoveeravat P, Yee EU, *et al*., 2020, Endoscopic ultrasound guided liver biopsy: Recent evidence. *World J Gastrointest Endosc*, 12(3): 83–97.

https://doi.org/10.4253/wjge.v12.i3.83

55. Farmer TD, Jenkins EC, O'Brien TP, *et al*., 2015, Comparison of the physiological relevance of systemic vs. portal insulin delivery to evaluate whole body glucose flux during an insulin clamp. *Am J Physiol Endocrinol Metab*, 308(3): E206–E222.

https://doi.org/10.1152/ajpendo.00406.2014

ACCSCIENCE PUBLISHING

Gene & Protein in Disease

REVIEW ARTICLE

Comprehensive prognostic signatures in thyroid cancer: A summarized review for molecular signatures construction strategies

Xiaoyan Lu[1,2†], Yuanyuan Zhang[1,2†], Pei Yang[1,2], Minjun Yi[1,2], Luyao Wang[1,2], Jing Chen[1,2], Han Wang[1,2], Mengke Li[1,2], Yufei Jiang[1,2], Bingbing Guo[1,2], Wenyuan Lu[1,2], Shijia Li[1,2], Jiahao Chen[1,2], Yingying Lian[1,2], Xinyu Li[1,2], Binbin Zhao[1,2], Xiaoqing Wang[1,2*], and Yang An[1,2*]

[1]Department of Biochemistry and Molecular Biology, Cell Signal Transduction Laboratory, School of Basic Medicine, Bioinformatics Center, Henan University, Kaifeng, 475004, China

[2]Henan Provincial Engineering Center for Tumor Molecular Medicine, Kaifeng Key laboratory of cell signal transduction, Kaifeng, 475004, China

Abstract

Thyroid carcinoma (TC) is one of the most common endocrine carcinomas with an increasing rate of morbidity in recent decades. With a high risk of relapse and metastasis occurring in TC patients, it is essential to identify potential prognostic signatures for TC patients. Here, through a comprehensive review, we summarized 45 prognostic signatures for TC patients and concluded three main strategies for signature establishment after an extensive investigation. In particular, these signatures were classified according to different construction strategies, and the verification methods were summarized. Besides, we found that 18 key genes were overrepresented in reported signatures. This review provides a comprehensive understanding, systematic summary, and integrated analysis of current prognostic signatures of TC, which may help researchers to further understand cancer progression, construct prognostic signatures of TC, and guide future clinical treatment.

Keywords: Thyroid cancer; Prognostic signatures; Survival outcome

[†]These authors contributed equally to this work.

***Corresponding authors:**
Xiaoqing Wang
(wangxiaoqing@henu.edu.cn)
Yang An
(anyang@henu.edu.cn)

Citation: Lu X, Zhang Y, Yang P, *et al*., 2023, Comprehensive prognostic signatures in thyroid cancer: A summarized review for molecular signatures construction strategies. *Gene Protein Dis*, 2(3): 1138. https://doi.org/10.36922/gpd.1138

Received: June 26, 2023
Accepted: August 16, 2023
Published Online: September 20, 2023

1. Introduction

TC is considered one of the most common endocrine carcinomas with a considerably high morbidity and mortality[1-3]. The average annual incidence of TC is more than 6% and has appreciably increased in recent decades[4]. According to distinct origins and differentiation degree of TC cells, TC could be mainly divided into five pathological types: Papillary TC (PTC), follicular TC (FTC), poor-differentiated TC (PDTC), anaplastic TC (ATC), and medullary TC (MTC)[5-8]. Among these, PTC and FTC were thought to be well-differentiated TC, the long-term survival rates of which are higher than 90%, and PTC is considered the most common type of thyroid cancer[9,10]. In addition to the main five types of TC, a rare subtype of TC also exists, namely, spindle cell thyroid cancer (SCTC)[11].

Although most TC patients exhibited good prognosis, some patients still underwent high risks of cancer recurrence and distant metastasis[12-14]. At present, clinical treatments for

TC patients are surgical resection, radioiodine, and systemic therapy according to different pathological types[9,15,16]. Therefore, if risk factors of TC progression could be precisely predicted, early intervention and targeted therapy might prevent the deterioration of TC, which would improve the prognosis of TC patients[17,18]. Recently, researchers have identified several signatures of prognostic values for TC patients and developed risk prediction models of TC progression to predict survival outcomes, classify patients by risk stratification, and guide treatments for TC[19-21].

This retrospective review discussed the previously reported signatures of TC, concluded three main strategies for signature construction related to TC prognosis, and summarized the methods of verifying signatures after integrated analysis.

2. Methods

2.1. Data collection and selection

To estimate previously reported prognostic signatures for TC, a total of 150 studies were obtained by searching the PubMed database with the keywords "prognostic signature AND thyroid cancer." Prognostic signatures derived from these studies included four categories: mRNA signatures, non-coding RNAs (ncRNAs) signatures, genomic signatures, and signatures related to biological functions. Selection of studies was then conducted according to the consistent exclusion criteria listed below: (1) review; (2) not focused on TC; (3) not in English; (4) only one molecule mentioned; and (5) no prognostic signature constructed/no survival analysis mentioned. As a result, a total of 42 studies correlated with TC prognosis were included in the study (Figure 1).

2.2. Evaluation of signatures through meta-analysis

To evaluate the prognostic abilities of different signatures for TC, survival data from the training set were collected and summarized, including (1) survival curves and risk scores, (2) receiver operating characteristic (ROC) plots, (3) univariate and multivariate Cox regression analysis, and (4) nomogram analysis. To visually estimate the abilities of risk stratification of signatures for TC patients, values of hazard ratios (HR) and 95% confidence intervals (CI) were extracted from multivariate regression analysis data of these studies and described as a forest plot (Figure 2) using GraphPad Prism 8.0.2. As described, 28 prognostic signatures estimated by meta-analysis presented great prediction performance that most high-risk TC patients had poorer survival rates than low-risk patients (Figure 2).

3. Results

Comprehensive signature information for TC prognosis was described in Tables 1-4 and Figure 2. By screening, a total of 42 studies were collected (Figure 1), 45 prognostic signatures for TC were summarized, and data types include mRNA, miRNA, and lncRNA. These prognostic signatures were associated with overall survival (OS), disease-free survival (DFS), recurrence-free survival (RFS), or progression-free interval (PFI) of TC patients. After analyzing and summarizing these signatures, we proceeded with the following works: (1) sorting out data sources and applications; (2) categorizing these prognostic signatures; (3) summarizing three main strategies for constructing signatures; (4) summing up the verification methods of signatures; (5) summarizing nomograms of these prognostic signatures; and (6) screening overrepresented prognostic genes.

3.1. Summary of the sources of data for signature construction

In general, for the establishment of prognostic signatures, the most pervasive source of data acquisition to download expression profiling and patients' clinical information was from the TCGA or GEO database. Besides, some studies obtained data from some specific databases, for instance, the Human Autophagy Database (HADb)[22,23], the Immunology Database, and the Analysis Portal (ImmPort) database[24-26]. These databases were typically rich in genes and molecules associated with specific biological functions. In addition, some researchers obtained data from array or sequencing results of a certain number of qualified thyroid cancer and normal tissue samples derived from TC patients[21,27-33]. For instance, in addition to obtaining 165 transcriptome data and 125 PTC patients' clinical data from the Nucleotide Archive database, Teng *et al.* also included 11 patients who had undergone total thyroidectomy in Beijing Cancer Hospital[21]. The data obtained from different approaches or platforms were not only the first step to constructing molecular signatures but also laid the foundation for researchers to carry out subsequent works.

3.2. Classification of prognostic signatures for TC

Derived from different data types, these signatures were divided into four categories. Signatures of 19 studies were identified by analyzing mRNA data[19,20,34-40] and ncRNA data involving long ncRNAs (lncRNAs)[33,41-43] and microRNAs (miRNAs)[29,44]. Signatures of 11 studies were associated with specific biological functions, including immune-related genes (IRGs)[24-26], autophagy-related genes (ARGs)[24-26], epithelial-mesenchymal transition (EMT)-related genes[45], RBPs-associated genes[30], and ferroptosis-related genes[27,31,46]. In addition, signatures of eight studies were related to mutation or methylation, including methylation-driven genes[47] and regulators[48,49].

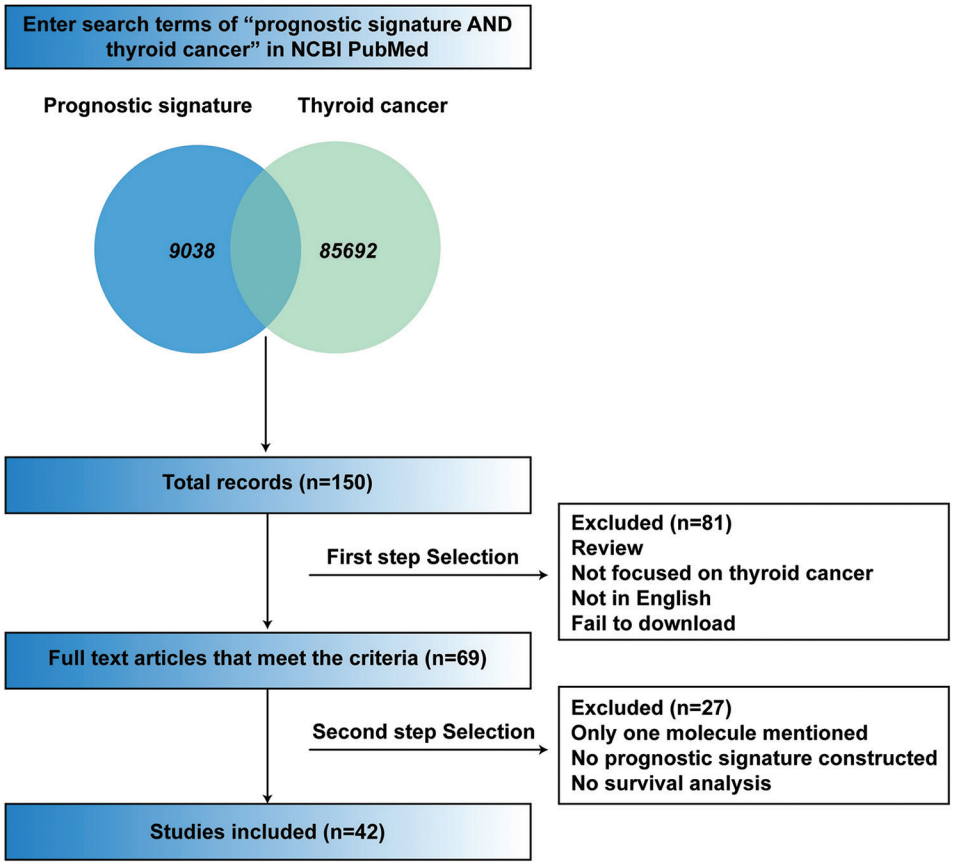


Figure 1. The flowchart of articles screening.

In recent years, many signatures have been established according to different research objectives and methods, including glucose metabolism[50], splicing events[51,52], gene mutations[53], and chromosomal alteration[54].

Occasionally, some studies contained two or more types of signatures. For instance, Li *et al.* constructed two prognostic signatures (Lnc2mi1m2 and Lnc5m4) which were capable of efficiently predicting the long-term OS of TC patients[55]. Wang *et al.* established three different types of signatures including a 6-mRNA-based classifier, a 5-lncRNA-based classifier, and a 4-miRNA-based classifier[56], to develop a comprehensive and reliable model for predicting the prognosis of TC patients. The predictive ability and accuracy of signatures might be improved by employing multiple types of prognosis signatures as well as constituting a network.

3.3. The strategies for signature construction

For diverse research objectives, based on different data sources and platforms, researchers usually preferred different research and analysis methods to construct prognostic labels. Through comprehensive analysis, three main strategies for constructing prognostic signatures for TC were summarized.

3.3.1. Strategy 1: Signatures based on differentially expressed genes

In strategy 1, signatures were mainly constructed through the following steps (Figure 3): (1) The selection and classification of research objects. Mostly, samples were selected as tumor tissues derived from TC patients and normal tissues derived from paracancerous tissues or other healthy people[29,40,44]. Ruiz *et al.* defined differentially expressed genes (DEGs) between N0 and N1 (N0 means no lymph node metastasis; N1 means primary lymph node metastasis) samples[20]. You *et al.* divided the data from TCGA into three groups: (i) Tumor and normal samples; (ii) PTC samples with or without lymph node metastasis; and (iii) PTC samples with stages 1 – 2 and stages 3 – 4[33]; (2) identification of differentially expressed mRNAs/lncRNAs/miRNAs. By comparing the different sample groups mentioned above, most authors applied edgeR to identify differentially expressed mRNAs/lncRNAs/miRNAs. Teng *et al.* used the Bioconductor package edgeR

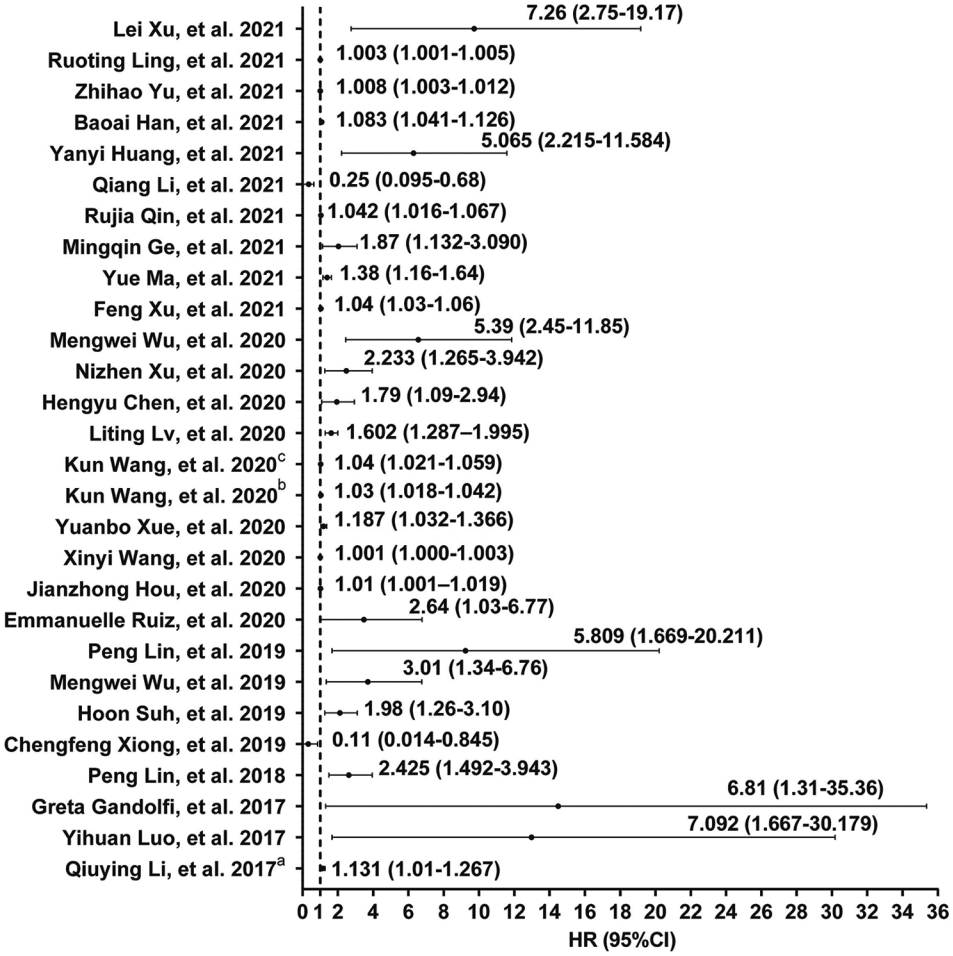


Figure 2. Evaluation of 28 prognostic signatures through meta-analysis. In the Forest plot, Qiuying Li, *et al.* 2017[a] means, the data of the multivariate Cox regression analysis were derived from the training set; Kun Wang, *et al.* 2020[b] means, HR (95% CI) was extracted from multivariate regression analysis of a 6-gene biomarker; and Kun Wang, *et al.* 2020[c] means, HR (95% CI) was extracted from multivariate analysis of a 5-lncRNA biomarker.

to identify differentially expressed genes and lncRNAs in two independent cohorts[21]. Moreover, Wu *et al.* applied the LIMMA package[19]; (3) signature construction. As differentially expressed mRNAs/lncRNAs/miRNAs were identified, a series of analyses were applied to construct signatures, including lease absolute shrinkage and selection operator (lasso) regression, Cox regression, and functional enrichment analysis. Ruiz *et al.* employed a machine learning model with linear discriminant analysis (LDA) to construct a 25-gene signature of great predictive ability[20]. Wu *et al.* applied lasso regression analysis to identify DEGs significantly associated with PFI[19]. Li *et al.* performed univariate Cox regression analysis to obtain lncRNAs related to the survival of TC patients[41]; (4) signatures validation. Most researchers applied ROC curves to reflect the predictive capabilities of signatures[19,20,42]. Moreover, univariate and multivariate logistic regression analyses were also performed to validate the association between signatures and other clinical parameters[33,41,44]. Teng *et al.* used Kaplan-Meier survival analysis as well as univariate and multivariate regression analyses to assess the prognostic power of the gene signature and whether it was considered an independent prognostic factor[21]. The flow chart of strategy 1 is shown in Figure 3 and the summarized relevant information of signatures based on strategy 1 is presented in Table 1. For instance, Wu *et al.* constructed a gene signature using the method of DEGs screening mentioned in strategy 1[19]. First, by comparing the differences in expression profiling between 510 tumors and 58 normal samples, 295 DEGs including 137 downregulated and 158 upregulated genes were found by the LIMMA package. Second, the authors applied univariate Cox regression analysis to identify 50 DEGs associated with PFI. The lasso regression model was then applied to develop a 5-gene prognostic signature (FXYD6, PLP2, FABP4, LYVE1, and TGFBR3). Finally, it is

Table 1. The prognostic signatures of TC patients are summarized in strategy 1

Author	Signature type	Signature	Survival event	Survival curve *P*-value	AUC of the ROC curve	Signature outcome	Univariate Cox analysis		Multivariate Cox analysis		References
							HR (95% CI)	*P*-value	HR (95% CI)	*P*-value	
Lei Xu, *et al.*, 2021	mRNA	ENTPD1, PAPSS2, LRRC75A, KIAA1211L, ADRA1B, PCOLCE2, HSPA6, PDLIM3, MPL, ANKRD37, LOXL2, ADAM22, SMIM10L2B, CTXND1, NECAB1, LBX2, ITPKA, F2RL2, MAST1, SIGLEC11, and EFNB3	OS	*P*=4e-04	AUC at 1-, 3-, and 5-years OS: 0.994, 0.987, and 1	Unfavorable	7.27 (3.4-15.5)	*P*=3.16e-07	7.26 (2.75 – 19.17)	*P*=6.37e-05	Xu *et al.*, 2021[39]
Yanyi Huang, *et al.*, 2021	mRNA	MMP9, AKR1C1, PLA2G2E, CARTPT, and SLC5A1	DFS	*P*=4.613e-06	AUC at 1-, 3-, and 5-year DFS: 0.804, 0.772, and 0.72	Unfavorable	5.015 (2.330 – 10.793)	*P*<0.001	5.065 (2.215 – 11.584)	*P*<0.001	Huang *et al.*, 2021[34]
Ruchong Pan, *et al.*, 2021	mRNA	TENM1, FN1, APOD, F12, and BTNL8	DFS	*P*=7.427e-14	AUC at 1-, 3-, and 5-year DFS: 0.716, 0.705, and 0.681	Unfavorable					Ruchong *et al.*, 2021[36]
He Ren, *et al.*, 2021	mRNA	CDH3, CTGF, CYR61, OGN, FGF13, and CHRDL1	OS	*P*=0.00087	AUCs for 1-, 3-, and 5-year OS: 0.81, 0.67, and 0.72	Unfavorable					Ren *et al.*, 2021[35]
Emmanuelle Ruiz, *et al.*, 2020	mRNA	a 25-gene signature	DFS	*P*=0.001	0.86	Unfavorable	2.81	*P*=0.002	2.64 (1.03 – 6.77)		Ruiz *et al.*, 2020[20]
Zhihao Yu, *et al.*, 2021	mRNA	FTO, RBM15, and KIAA1429	OS	*P*=0.01798	0.729	Unfavorable	1.007 (1.003 – 1.011)	*P*<0.001	1.008 (1.003 – 1.012)	*P*=0.001	Yu *et al.*, 2021
Linkun Zhong, *et al.*, 2020	mRNA	ADRA1B, RIPPLY3, PCOLCE, TEKT1, and SALL3	OS	*P*=0.0000259	AUC of 3- and 5-year survival: 0.937 and0.925	Unfavorable					Zhong *et al.*, 2020[40]
Mengwei Wu, *et al.*, 2020	mRNA	CLCNKB, FXBO27, FXYD6, RIMS2, SPC24, and CDKN2A	PFI	*P*<0.0001		Unfavorable	6.9 (3.45 – 13.79)	*P*<0.0001	5.39 (2.45 – 11.85)	*P*<0.0001	Wu *et al.*, 2020[38]
Mengwei Wu, *et al.*, 2019	mRNA	PLP2, LYVE1, FABP4, TGFBR3, and FXYD6	PFI	*P*=0.0001	AUC of 1-, 2-, 3-, 4-, and 5-year PFI: 0.6623, 0.6740, 0.6791, 0.6656, and 0.6484	Unfavorable	4.15 (1.88 – 9.16)	*P*=0.0004	3.01 (1.34 – 6.76)	*P*=0.0077	Wu *et al.*, 2019[19]

(Cont'd...)

https://doi.org/10.36922/gpd.1138

Table 1. ***(Continued)***

Author	Signature type	Signature	Survival event	Survival curve *P*-value	AUC of the ROC curve	Signature outcome	Univariate Cox analysis		Multivariate Cox analysis		References
							HR (95% CI)	*P*-value	HR (95% CI)	*P*-value	
Yun Zhang, *et al.*, 2019	lncRNA	DOCK9-DT, LINC00900, C8orf34-AS1, LINC01736, ATP2C2-AS1, FAM111A-DT, LINC02550, and LINC01896	OS	*P*=0.0001627	0.862	Unfavorable					Zhang *et al.*, 2019[43]
Qiuying Li, *et al.*, 2017[a]	lncRNA	RP11-536N17.1, RP11-508M8.1, AC026150.8, and CTD-2139B15.2	OS	*P*=0.016	0.833	Unfavorable	1.094 (1.028 – 1.163)	*P*=0.004	1.131 (1.01 – 1.267)	*P*=0.034	Li *et al.*, 2017[41]
Yihuan Luo, *et al.*, 2017	lncRNA	AC079630.2, CRNDE, and CTD-2171N6.1	OS	*P*<0.001		Unfavorable	8.592 (2.878 – 25.651)	*P*<0.001	7.092 (1.667 – 30.179)	*P*=0.008	Luo *et al.*, 2017[42]
Xin You, *et al.*, 2018	lncRNA	PRSS3P2, KRTAP5-AS1, and PWAR5	OS	*P*=0.022	0.739	Unfavorable					You *et al.*, 2018[33]
Tong Liu, *et al.*, 2018	miRNA	Has-miR-181a-2-3p and has-miR-138-1-3p	OS	*P*=0.003	0.784	Unfavorable					Liu *et al.*, 2018[11]
Chengfeng Xiong, *et al.*, 2019	miRNA	Has-mir-6843, hsa-mir-6730, hsa-mir-196a-2, and hsa-mir-206	OS	*P*=0.00006	0.886	Unfavorable	19.091 (2.519 – 144.674)	*P*=0.004	0.11 (0.014 – 0.845)	*P*=0.034	Chengfeng *et al.*, 2019[44]
Huajing Teng, *et al.*, 2018	mRNA/ lncRNA	HIGD1B, SDPR, PACERR, and CYP4A22-AS1	DFS								Teng *et al.*, 2018[21]
Mihai Saftencu, *et al.*, 2019	mRNA/ miRNA	ITGA10 and MSLN	DFS								Saftencu *et al.*, 2019[37]
Xiaoping Li, *et al.*, 2020	mRNA/ miRNA/ lncRNA	Lnc2mi1m2 (DIO3OS, MLIP-AS1, has-mir-96, CADM2, TPPP)	OS	*P*=0.012	AUC values for predicting 5-year and 10-year survival: 0.894 and 0.957	Unfavorable	4.39 (1.24 – 15.57)	*P*<0.05			Zhang *et al.*, 2020[55]
	mRNA/ lncRNA	Lnc5m4 (AC006487.1, AC022148.1, DGUOK-AS1, TLR8-AS1, AC024563.1, ALPL, GALNT7, ARMCX6, IL11)	OS	*P*=0.00244	AUC values for predicting 5-year and 10-year survival: 0.843 and 0.89	Unfavorable	11.81 (1.55 – 89.93)	*P*=0.0172	9.28 (1.19 – 72.58)	*P*=0.0338	Zhang *et al.*, 2020[55]

(Cont'd...)

Table 1. ***(Continued)***

Author	Signature type	Signature	Survival event	Survival curve *P*-value	AUC of the ROC curve	Signature outcome	Univariate Cox analysis		Multivariate Cox analysis		References
							HR (95% CI)	*P*-value	HR (95% CI)	*P*-value	
Kun Wang, *et al.*, 2020[b]	mRNA	PIMREG, PAQR4, RHBDL1, JAKMIP3, GAS2L2, and PLEKHG4	PFI	*P*=0.000002462	AUC of the first, third, and fifth years in the training set were 0.83, 0.82, and 0.82	Unfavorable	1.032 (1.022, 1.042)	*P*<0.001	1.030 (1.018, 1.042)	*P*<0.001	Wang *et al.*, 2020[66]
	lncRNA	AC136475.1, LINC02154, AC010969.1, and AC064805.2	PFI	*P*=0.00005696	AUC of the first, third, and fifth years in the training set were 0.75, 0.84, and 0.85	Unfavorable	1.047 (1.030, 1.065)	*P*<0.001	1.040 (1.021, 1.059)	*P*<0.001	Wang *et al.*, 2020[66]

Abbreviations: OS: Overall survival; DFS: Disease-free survival; PFI: Progression-free interval. Li *et al.* 2017[a] means the data were derived from the training set in this study; Wang *et al.* 2020[b] means this study identified three prognostic signatures in the training set, including a 6-mRNA-based signature, a 5-lncRNA-based signature, and a 4-miRNA-based signature. Unfavorable means patients with a high risk of prognostic signatures are associated with a worse prognosis.

essential to verify the prognostic effect of this signature. By calculating the risk score and determining the cutoff value, patients were divided into high-risk and low-risk groups. The survival curve showed a better prognosis in the low-risk group than in the high-risk group. Furthermore, the area under the receiver operating characteristic (ROC) curve (AUC) and Harrell's concordance index (C-index; 0.7600) also demonstrated that this signature has a remarkable prognostic ability.

3.3.2. Strategy 2: Signatures related to specific biological functions

In strategy 2, these signatures were built based on biological functions. The process of constructing signatures was as follows: (1) Identification of differential expression genes/lncRNAs. Most authors downloaded available transcriptome data from various databases[22,23,30] or hospitals[27,32]; (2) identification of genes associated with specific biological functions, mainly containing immune-related and autophagy-related genes, which were derived from the immunology database[24-26] and the Human Autophagy Database (HADb)[22,23], respectively. Moreover, gene functions associated with RNA-binding proteins[30], ferroptosis[27,31,46], EMT[45], and glycolysis[32] were all involved. The signatures related to different biological functions are illustrated in detail below; (3) integrate and analyze the molecules to obtain differentially expressed biological function-related genes or lncRNAs, followed by functional enrichment analyses being performed; and (4) establishment of prognostic signatures. To construct a prognostic gene signatures, studies based on strategy 2 adopted the Cox proportion hazards regression model and introduced the risk score formula to stratify TC patients[22-27,30-32,45]; and (5) validation of signatures. Methods of verification included Kaplan-Meier survival curves, ROC curves, and univariate and multivariate Cox regression analyses[24-27,30-32,46]. The flow chart of strategy 2 is shown in Figure 4 and the summarized relevant information of signatures based on strategy 2 is presented in Table 2.

According to strategy 2, several functional signatures associated with IRGs and ARGs were established[22-26]. Xue *et al.* developed a 5-immune-associated genes (IAGs) signature to predict TC progression for patients[26]. The construction strategy was described as follows: on the one hand, 509 TC and 58 normal samples were obtained from the TCGA database; on the other hand, a list of 2498 IAGs was acquired from the Immunology Database (ImmPort). Eighty-two differentially expressed IAGs were then screened by an integrated analysis. After univariate, multivariate, and lasso regression analyses, a reliable 5-IAGs signature was constructed, including CXCL5,

Table 2. The prognostic signatures of TC are summarized in strategy 2

Authors	Signature type	Signature	Survival event	Survival curve *P*-value	Risk score	Signature outcome	AUC of the ROC curve	Univariate Cox analysis		Multivariate Cox analysis		References
								HR (95% CI)	*P*-value	HR (95% CI)	*P*-value	
Rujia Qin, *et al.*, 2021	IRGs	ULBP2, S100A5, LTF, PLXNA4, FAM3B, GIPR, RORB, and TGFBR3	PFI	$P = 0.0000026$	Risk score = (−0.876 * ULBP2) + (0.301 * S100A5) + (−0.582 * LTF) + (−1.003 * PLXNA4) + (−0.975 * FAM3B) + (1.901 * GIPR) + (1.606 * RORB) + (−0.907 * TGFBR3)	Unfavorable	AUCs for 1-, 3-, and 5-year PFI: 0.802, 0.729, and 0.703	1.046 (1.027 – 1.065)	$P < 0.001$	1.042 (1.016 – 1.067)	$P = 0.001$	Qin *et al.*, 2021[25]
Yuanbo Xue, *et al.*, 2020	IRGs	CXCL5, AZU1, NOD1, TNFRSF11B, and VGF	OS	$P = 0.001524$	Risk score = (0.6846 * expression value of CXCL5) + (1.1556 * expression value of AZU1) + (−0.3156 * expression value of NOD1) + (0.0542 * expression value of TNFRSF11B) + (0.0952 * expression value of VGF)	Unfavorable	0.888	1.018 (1.004 – 1.033)	$P = 0.011$	1.187 (1.032 – 1.366)	$P = 0.016$	Xue *et al.*, 2020[26]
Peng Lin, *et al.*, 2019	IRGs	AGTR1, CTGF, FAM3B, IL11, IL17C, PTH2R, and SPAG11A	PFI	$P < 0.001$	Risk score = [Expression level of AGTR1 * (−0.1212)] + [Expression level of CTGF * (−0.3284)] + [Expression level of FAM3B * (−0.1675)] + [Expression level of IL11 * 0.3089] + [Expression level of IL17C * 0.2368] + [Expression level of PTH2R * (−0.2823) + [Expression level of SPAG11A * (-0.9550)]	Unfavorable	0.792	2.711 (2.046 – 3.592)	$P < 0.001$	2.425 (1.492 – 3.943)	$P < 0.001$	Lin *et al.*, 2019[24]
Gang Hu, *et al.*, 2020	ARGs	ITPR1, CCL2, and CDKN2A	OS	$P = 0.00042$	Risk score = (0.123337693147 * expression value of ITPR1) + (0.028524 * expression value of CCL2) + (0.192202 * expression value of CDKN2A)	Unfavorable	AUCs of the signature predicting the 5- and 10-year OS rates: 0.839 and 0.886					Hu *et al.*, 2020[22]
Peng Lin, *et al.*, 2018	ARGs	ATG9B, BID, and DNAJB1	OS	$P = 0.0023$	Risk score = 0.469 * ATG9B expression – 0.796 * BID expression + 0.782 * DNAJB1 expression	Unfavorable						Lin *et al.*, 2018[23]

(Cont'd...)

Table 2. *(Continued)*

Authors	Signature type	Signature	Survival event	Survival curve *P*-value	Risk score	Signature outcome	AUC of the ROC curve	Univariate Cox analysis		Multivariate Cox analysis		References
								HR (95% CI)	*P*-value	HR (95% CI)	*P*-value	
Yue Ma *et al.*, 2021	RBPs	AZGP1, IGF2BP2, MEX3A, NUDT16, NUP153, and USB1	OS	$P = 0.009$	Risk score = (1.021 * Exp AZGP1) + (−0.09 * Exp IGF2BP2) + (0.241 * Exp MEX3A) + (0.127 * Exp NUDT16) + (0.388 * Exp NUP153) + (−0.236 * Exp USB1)	Unfavorable	AUCs for 3-, and 5-year OS: 0.766 and 0.784	1.35 (1.15 – 1.58)	$P < 0.001$	1.38 (1.16 – 1.64)	$P < 0.001$	Ma *et al.*, 2021[30]
Mingqin Ge, *et al.*, 2021	Ferroptosis-related genes	DPP4, GPX4, GSS, AKR1C1, HMGCR, TFRC, SQLE, and PGD	OS	$P = 0.001186$	Risk score = (0.340143834733433) * AKR1C1 expression + (−0.319305763663027) * DPP4 expression + (1.3413362890122) * GPX4 + (−2.69880844893117) * GSS + (0.171053637326744) * HMGCR + (1.0923479137719) * TFRC + (0.0892997114590679) * SQLE + (0.499540011822522) * PGD	Unfavorable	AUCs for 1-, 2- and 3-year OS: 0.887, 0.890 and 0.842	2.528 (1.716 – 3.725)	$P < 0.001$	1.87 (1.132 – 3.090)	$P = 0.015$	Ge *et al.*, 2021[27]
Xiaoyu Qian, *et al.*, 2021	Ferroptosis-related genes	HMGCR, GSS, TFRC, DPP4, and PGD	OS	$P = 0.0049$	Risk score = (1.051 * the level of expression of HMGCR + 0.913 * the level of expression of GSS + 1.098 * the level of expression of TFRC + 0.977 * the level of expression of DPP4 + 1.008 * the level of expression of PGD)	Unfavorable	AUCs for 1-, 2- and 3-year OS: 0.621, 0.728, and 0.875	10.697 (1.328 – 86.173)	$P = 0.026$	11.682 (1.454 – 93.878)	$P = 0.021$	Qian *et al.*, 2021[31]
Ruoting Lin, *et al.*, 2021	Ferroptosis-related genes	Ac008063.2, Apoe, Bcl3, Acap3, Alo × 5ap, Atxn2l, and B2m		$P = 6.329e-4$		Unfavorable	0.748	1.006 (1.004 – 1.008)	$P < 0.0001$	1.003 (1.001 – 1.005)	$P = 0.007$	Lin *et al.*, 2021[46]
Qiang Li, *et al.*, 2021	ERGs	FN1, ITGA2, and KIT	DFS	$P = 0.0017$		Unfavorable	AUCs for 1-, 2- and 3-year DFS: 0.659,0.637, 0.611					Li *et al.*, 2021[45]

(Cont'd...)

https://doi.org/10.36922/gpd.1138

Table 2. *(Continued)*

Authors	Signature type	Signature	Survival event	Survival curve *P*-value	Risk score	Signature outcome	AUC of the ROC curve	Univariate Cox analysis		Multivariate Cox analysis		References
								HR (95% CI)	*P*-value	HR (95% CI)	*P*-value	
Feng Xu, *et al.*, 2021	Glycolysis-based genes	CHST6, POM121C, PPFIA4, STC1, TGFBI, and FBP2	OS	$P = 0.0007$	Risk score = (0.0149 * TGFBI expression level) + (0.0517 * STC1 expression level) + (1.866 * PPFIA4 expression level) + (0.345 * POM121C expression level) + (0.542 * CHST6 expression level) + (2.672 * FBP2 expression level)	Unfavorable	0.929	1.05 (1.04 – 1.07)	$P < 0.001$	1.04 (1.03 – 1.06)	$P < 0.001$	Xu *et al.*, 2021[39]

Abbreviation: PFI: Progression-free interval; OS: Overall survival; IRGs: Immune-related genes; ARGs: Autophagy-related genes; RBPs: RNA binding proteins; ERGs: EMT-related genes. Unfavorable means patients with a high risk of prognostic signatures are associated with a worse prognosis.

AZU1, NOD1, TNFRSF11B, and VGF. To verify the prognostic effect of this signature, samples were divided into high- and low-risk groups by risk score. The survival curve showed that the high-risk group presented a poor prognosis. The ROC curve also indicated the predictive power of this signature. Other immune- or autophagy-related studies also adopted similar signature construction methods.

In recent years, some studies manifested that ferroptosis could be linked to various types of cancers[57-60]. Thus, some researchers have explored the relationship between ferroptosis and thyroid cancer and constructed ferroptosis-related signatures with high prognostic values[27,31,46]. By an integrated analysis of DEGs and ferroptosis-related genes, a 5-gene prognostic signature was discovered, which consisted of DPP4, HMGCR, TFRC, PGD, and GSS[31]. Similarly, an 8-gene signature containing GPX4, AKR1C1, SQLE, and 5 aforementioned genes was constructed. Moreover, analysis of immune microenvironment and nomogram verification was performed to validate an 8-gene signature[27]. These two signatures (5-gene vs. 8-gene) exhibited different AUC values of ROC curves, indicating the difference in their predictive abilities. In addition, Lin *et al.* also constructed a ferroptosis-related 7-gene signature (*Apoe*, *Acap3*, *Bcl3*, *Ac008063.2*, *Alox5ap*, *B2m*, and *Atxn2l*) by comparatively analyzing two different PTC subtypes[46].

EMT has been regarded as an indicator in many types of tumors[61]. Li *et al.* screened 3 genes (ITGA2, FN1, and KIT) by analyzing the differentially expressed EMT-related genes, thus a 3-gene signature associated with EMT was established for the prognosis of TC patients[45], as the survival curve illustrated that the high-risk group has a worse prognosis than the low-risk group. It has been reported that RNA-binding proteins (RBPs) play a critical role in the post-transcriptional regulation of RNA and thus may affect the occurrence and development of cancer[30]. By analyzing the differently expressed RBPs, combined with pathway and functional enrichment analyses, Ma *et al.* constructed a 6-gene prognostic signature (NUDT16, NUP153, IGF2BP2, MEX3A, USB1, and AZGP1), which may provide a novel strategy for the risk prediction of TC[30]. Using the similar methods described above, Xu *et al.* selected 6 glycolysis-related genes (POM121C, PPFIA4, FBP2, CHST6, STC1, and TGFBI) to set up a 6-gene prognostic signature[32]. The AUC value of this signature is 0.929, indicating its reliable prognostic ability.

3.3.3. Strategy 3: Signatures associated with methylation

In recent years, epigenetic regulation has been found to be crucial for gene expression[62,63]. The importance

Table 3. The prognostic signatures of TC are summarized in strategy 3

Authors	Pathological type	Signatures type	Signatures	Signatures outcome	Survival event	Survival curve		Risk score	AUC of the ROC curve	Univariate Cox regression analysis			Multivariate Cox regression analysis			References
						Cut-off	*P*-value			HR	95% CI	*P*-value	HR	95% CI	*P*-value	
Xinyi Wang, *et al.*, 2020	PTC	m6A RNA methylation regulators	IGF2BP2, STT3A, MTHFD1, and GSTM4		PFS	The median risk score	*P* = 0.00008166	Risk score = (0.390 * MTHFD1) + (0.167 * IGF2BP2) + (0.152 * GSTM4) + (−0.133 * STT3A)	0.817	1.001	1.000 – 1.003	*P* = 0.047	1.001	1.000 – 1.003	*P* = 0.04	Wang *et al.*, 2020[66]
Liting Lv, *et al.*, 2020	TC	Methylation -driven genes	RDH5, TREM1, BIRC7, and SLC26A7	Unfavorable	OS	The median of risk scores	*P* < 0.001	Risk score = (0.331) * Expression value of RDH5 + (0.165) * Expression value of TREM1 + (0.017) * Expression value of BIRC7 + (0.016) * Expression value of SLC26A7	0.836	1.344	1.142 – 1.582	*P* < 0.001	1.602	1.287 – 1.995	*P* < 0.001	Lv *et al.*, 2020[47]
Zhiwei Chen, *et al.*, 2021	TC	Methylation -driven genes	ALDOC, C14orf62, DVL1 and PTPRC	Unfavorable	OS	The median of risk scores	*P* = 0.016	Risk score = (3.23) * β value (ALDOC) + (2.98) * β value (C14orf62)+ (−8.96) * β value (DVL1) + (18.23) * β value (PTPRC)	1-, 3- and 5-year AUCs: 0.73, 0.71 and 0.76							Chen *et al.*, 2021[79]
Jianzhong Hou, *et al.*, 2020	PTC	m6A RNA methylation regulators	RBM15 KIAA1429 FTO	Unfavorable	OS	The median risk score	*P* = 0.01541	Risk score = (1.139 * RBM15) + (0.194 * KIAA1429) + (0.225 * FTO)	0.716	1.013	1.005 – 1.021	*P* < 0.001	1.01	1.001 – 1.019	*P* = 0.022	Hou *et al.*, 2020[49]
Nizhen Xu, *et al.*, 2020	DTC	m6A RNA methylation regulators	WTAP, ZC3H13, HNRNPC and ALKBH5	Unfavorable	PFS	The median risk score	*P* = 0.0004741	Risk score = (0.0341 * expression value of HNRNPC) + (−0.0184 * expression value of ZC3H13) + (−0.0128 * expression value of ALKBH5) + (-0.0877 * expression value of WTAP)	0.747	5.058	3.212 – 7.963	*P* < 0.001	2.233	1.265 – 3.942	*P* = 0.006	Xu *et al.*, 2020[48]
Mariana Reis, *et al.*, 2017	WDTC	DNA methylation probes	21 DNA methylation probes**	Unfavorable	DFS											

(Cont'd...)

https://doi.org/10.36922/gpd.1138

Table 3. ***(Continued)***

Authors	Pathological type	Signatures type	Signatures	Signatures outcome	Survival event	Survival curve		Risk score	AUC of the ROC curve	Univariate Cox regression analysis			Multivariate Cox regression analysis			References
						Cut-off	*P*-value			HR	95% CI	*P*-value	HR	95% CI	*P*-value	
Hengyu Chen, *et al.*, 2020	PTC	DNA methylation sites	A 6-DNA methylation signature	Unfavorable	RFS	The median risk score	$P = 0.0001$	Risk score= −5.367*cg17749033 + 1.619*cg24221648 + 2.334*cg01664864 + 1.873 *cg09578568 − 3.486*cg24051057 + 5.693*cg05972352	1-, 3- and 5-year AUCs: 0.731,0.746 and 0.766	2.13	1.52 – 2.99	$P =$ 0.00001	1.79	1.09 – 2.94	$P = 0.019$	Chen *et al.*, 2020[78]
Qiuying Li, *et al.*, 2020	TC	methylation -driven lncRNA	A methylation-driven 5-lncRNA -based signature	Unfavorable	OS	The optimal risk cutoff point	$P < 0.001$	EpiLncPM = 13.1 * AP006248.2 + 2.53 * AC068580.3 + 33.2 * AC016396.2 + 3.12 * LINC01140 + 1.19 * LINC01135	AUC of 3,5 and 10-year survival rates: 0.948,0.965 and 0.949	50.097*	10.231 – 245.312	$P < 0.001$	29.05*	5.94 – 142.074	$P < 0.001$	Li *et al.*, 2019[28]

Abbreviations: TC: Thyroid cancer; PTC: Papillary thyroid cancer; DTC: Differentiated thyroid cancer; WDTC: Well-differentiated thyroid cancer; OS: Overall survival; RFS: Recurrence-free survival; PFS: Progression-free survival; DFS: Disease-free survival; EpiLncRNAs: Epigenetically regulated lncRNA-based prognostic scoring model. Unfavorable means patients with a high risk of prognostic signatures are associated with a worse prognosis.

of methylation in the clinical prognosis of TC has been widely recognized[64,65]. In strategy 3, these studies included not only methylation-driven genes[47] but also regulators of methylation[48,49,66]. In particular, regulators of m6A RNA methylation were considered important risk predictors[48,49,66]. Through comprehensive analysis, the steps to build these signatures were as follows: (1) transcriptome, methylation data, and relevant clinical information were downloaded from databases, and samples contained both tumor and normal tissues; (2) the key methylation-driven genes or regulators of methylation were screened through various analyses; (3) signatures were constructed by function enrichment analyses[47,66]; and (4) Cox regression analysis, ROC, and survival curve were applied to verify the prognostic capabilities of these signatures. The flow chart of strategy 3 is shown in Figure 5 and the summarized relevant information of strategy 3 is presented in Table 3.

In detail, based on methylation-driven genes, Lv *et al.* constructed a 4-gene signature composed of RDH5, BIRC7, TREM1, and SLC26A7[47]. Transcriptome data from 567 samples and DNA methylation data from 570 samples were collected to screen 51 methylation-driven genes (46 hypermethylation and 5 hypomethylation) using "MethylMix" and "limma" R package. In addition, after Cox regression analysis, the methylation-driven gene signature was established and validated by the Kaplan-Meier survival curve, certifying that patients in the high-risk group presented a worse prognosis[47]. Besides, Xu *et al.*, Hou *et al.*, and Wang *et al.* also discovered that the m6A RNA methylation regulators have a high-risk evaluation potency. Xu *et al.* and Wang *et al.* constructed two signatures of 4 RNA methylation regulators, respectively[48,66], whereas Hou *et al.* selected RBM15, FTO, and KIAA1429 to establish a 3-gene signature[49]. By comparing the differentially expressed m6A RNA methylation regulators, they selected 13 RNA methylation regulators. Cox and lasso regression analyses were applied to assess the relationship between OS and these methylation regulators. Thus, this 3-gene signature was constructed and validated.

3.3.4. Other strategies

In addition to the three main aforementioned strategies, few studies adopted other strategies to identify signatures[50-54], and the summarized relevant information of these strategies is presented in Table 4.

3.3.4.1. Signatures related to glucose metabolism

Increasing evidence has demonstrated that glucose metabolism and glucose transporters (GLUTs) play essential roles in TC progression[67-70]. Suh *et al.* have studied glucose metabolism by constructing GLUT signature and

Table 4. The prognostic signatures of TC are based on other strategies

Authors	Pathological type	Signatures type	Signature	Signatures outcome	Survival event	Survival curve		ROC curve AUC	Univariate Cox regression analysis			Multivariate Cox regression analysis			References
						Cut-off	*P*-value		HR	95% CI	*P*-value	HR	95% CI	*P*-value	
Suh HY, *et al.*, 2019	PTC, PDTC, ATC	Related to glucose metabolism	GLUT signature and glycolysis signature	Unfavorable	RFS				1.5	1.03 – 2.17	$p = 0.033$	1.98	1.26 – 3.10	$P = 0.003$	Suh *et al.*, 2019[50]
Lin P *et al.*, 2019	PTC	Related to AS events	20 survival-associated AS events	Unfavorable	PFIs	The median risk score	$P < 0.0001$	0.843				5.809	1.669 – 20.211	$P < 0.001$	Lin *et al.*, 2019[52]
Han B *et al.*, 2021	TC	Related to AS events	A final AS events signature	Unfavorable	OS		$P =$ 2.002e-05	AUC at 3, 5, and 8 years: 0.991, 0.984, and 0.976	1.04	1.026 – 1.054	$p < 0.001$	1.083	1.041 – 1.126	$P < 0.001$	Han *et al.*, 2021[51]
Gandolfi G *et al.*, 2018	PTC	Related to genetic alterations	THYT1: duplication of Chr1q, duplication of Chr5p harboring the TERT genomic locus and mutations of the TERT promoter	Unfavorable	OS		$P = 0.001$		7.29	1.63– 32.57	$p = 0.001$	6.81	1.31 – 35.36	$P = 0.022$	Gandolfi *et al.*, 2018[54]
Han Y *et al.*, 2021	PTC	Related to genetic mutation	FN1, MET, SLC34A2, NGEF, TBC1D2, PLCD3, PROS1, and NECTIN4												Han *et al.*, 2021[53]

Abbreviations: TC: Thyroid cancer; PTC: Papillary thyroid cancer; DTC: Differentiated thyroid cancer; WDTC: Well-differentiated thyroid cancer; OS: Overall survival RFS: Recurrence-free survival; PFS: Progression-free survival; DFS: Disease-free survival. Unfavorable means patients with a high risk of prognostic biomarkers are associated with a worse prognosis

https://doi.org/10.36922/gpd.1138

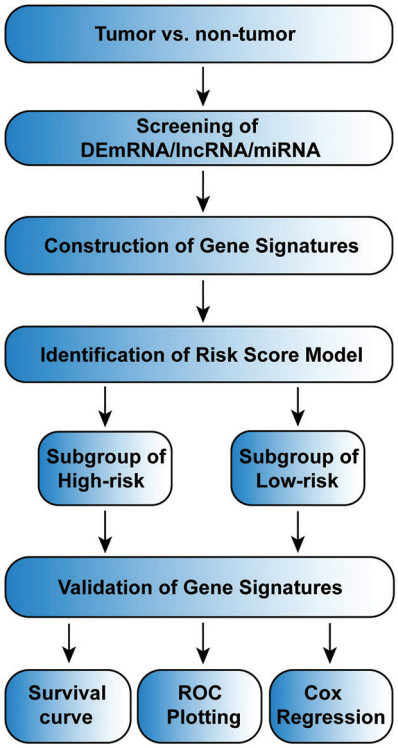


Figure 3. The flow chart of strategy 1: Signatures based on differentially expressed mRNAs/lncRNAs/miRNAs

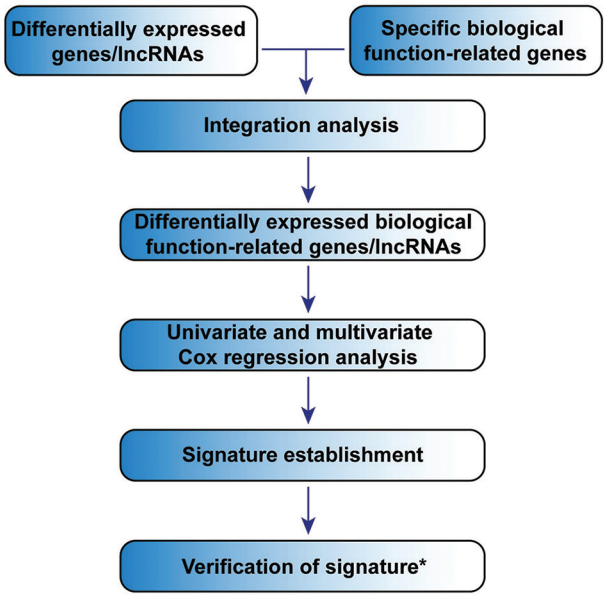


Figure 4. The flow chart of strategy 2: Signatures based on DEGs with specific biological functions. Verification of biomarker* means the validation methods include survival curve, ROC curve, and Cox regression analysis.

glycolysis signature[50] using microarray data of PDTC and ATC patients. The glycolysis and GLUT scores were then respectively estimated to define each signature, taking BRAFV600E mutation into account. Thus, K-M survival analysis showed that patients with high glycolysis or GLUT scores possessed a worse prognosis in PDTC/ATC[50].

3.3.4.2. Signatures related to alternative splicing events

As reported, splicing events significantly influence the occurrence of cancer[71-74]. Lin *et al.* obtained data from 496 patients with AS profiles derived from TCGA SpliceSeq and clinical data. Lasso regression analysis was performed to develop seven types (alternate acceptor site [AA], alternate promoter [AP], alternate terminator [AT], alternate donor site [AD], exon skip [ES], retained intron [RI], and mutually exclusive exons [ME]) of splicing events. Moreover, 20 alternative splicing (AS) events were screened to recognize the most significant prognostic signature. This signature was then verified as an independent predictor by the ROC curve (AUC value is 0.843) and multivariate analysis[52]. Furthermore, Han *et al.* have also concerned with the pivotal role of the alternative splicing events in TC prognosis and built a prediction model including AA, ES, AD, ME, and AT events with accurate prognostic efficacy[51], indicating that the prognosis effect of alternative splicing events on TC was a promising research subject.

3.3.4.3. Signatures correlated with genetic alterations

Recently, prognostic biomarkers with genetic alterations in various tumors have become a research focus[75]. Gandolfi *et al.* constructed a genetic alteration-related signature that consisted of three genetic variables[54]. The first step of their study was the acquisition of clinical data on PTC patients. A series of analyses of the genetic profiles were then conducted in distant metastasis (DM) PTCs and control samples to obtain the differential alterations between DMs and controls. As a result, three genetic variables including duplication of *Chr1q*, duplication of *Chr5p* harboring TERT locus, and mutations in TERT promoter displayed strong relevance with distant metastasis. Hence, a distinctive signature integrated with three genetic features was identified as Thyroid TERT *Chr1q* (THYT1). K-M survival analysis was then performed to assess the association of the THYT1 signature with the progression of distant metastasis. Furthermore, through the univariate and multivariate Cox models, the THYT1 signature was demonstrated to be an independent risk factor that was capable of predicting the aggressiveness in PTCs.

3.3.4.4. Signatures constructed with genetic mutation

Genetic mutation is a quite common phenomenon in TC[76,77]. Han *et al.* downloaded mutation data from 487 samples, and expression profiling data from 502 PTC and 58 normal samples. These samples were then

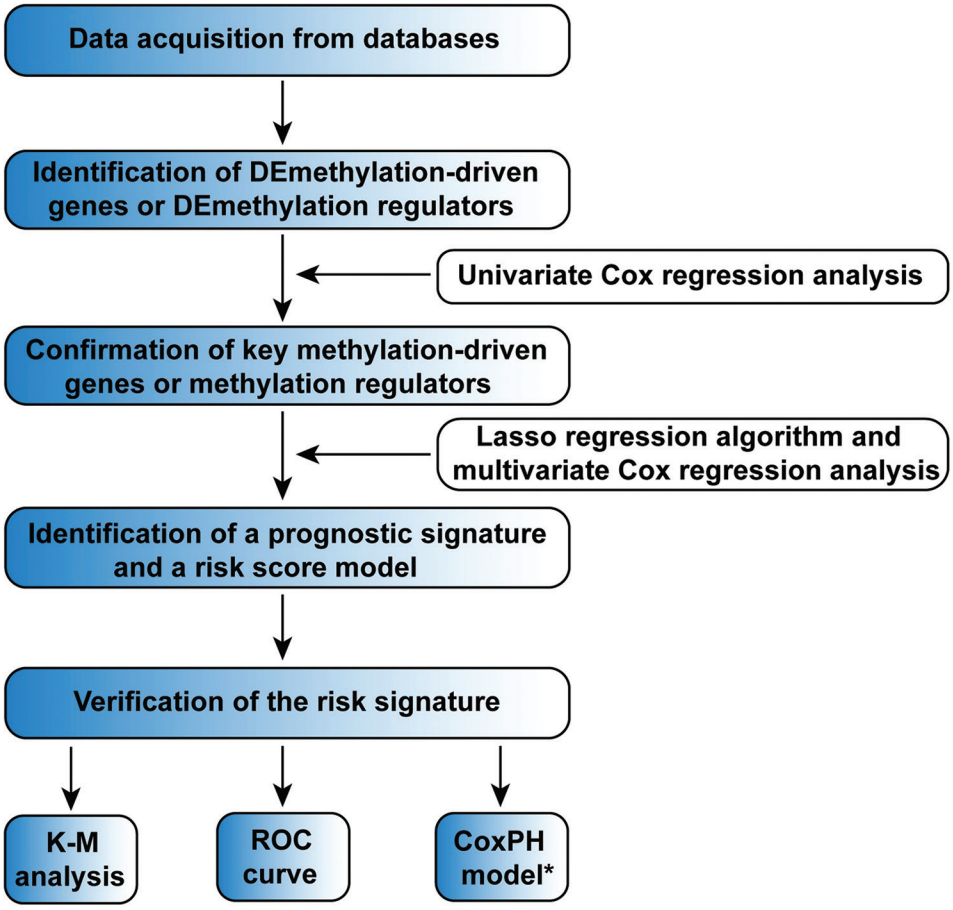


Figure 5. The flow chart of strategy 3: Signatures based on methylation-related molecules. CoxPH model* means the univariate and multivariate Cox proportional hazards regression analysis. DEmethylation, differentially expressed methylation.

divided into the BRAFWT group and the BRAFV600E group according to the mutation information to identify DEGs. Furthermore, weighted correlation network analysis (WGCNA) function or pathway enrichment, and correlation analysis was used to construct a co-expression network. In general, 8 BRAFV600E-associated prognostic biomarkers (FN1, MET, SLC34A2, NGEF, TBC1D2, PLCD3, PROS1, and NECTIN4) were discovered, which could be considered as targets for TC treatment[53].

3.4. The verification methods of signatures

In this review, the methods of constructing prognostic signatures were summarized. It is noteworthy that the validation effects of the constructed signatures could reflect the potential of prognostic signatures for clinical application. Through retrospective investigation and integrated analysis, we found that the verification methods described in this review were often closely related to the application of data acquisition and the process of signature construction. Therefore, according to the application of data, we divided the main validation methods into two categories: internal validation and external validation.

Data used to construct a signature were classified into three main types: the whole dataset, the training set, and the testing set. According to the research purpose, the verification methods of signatures include the following three categories: (1) the validation methods of prognostic signatures were performed in the whole dataset without distinguishing training and testing set[34,42,47,78]; (2) the process of validation was conducted in the training set and testing set, respectively[26,28,41]; and (3) distinguish among the entire dataset, training set, and testing set, and validate signatures in all three datasets[44,66]. In terms of specific verification methods, most researchers adopted common aforementioned validation approaches, including Kaplan-Meier survival, ROC curve, and univariate and multivariate Cox regression analyses. The main methods of internal validation were summarized above.

A common approach to conducting external validation was to employ hospital data sources as the validation

cohort. For instance, the internal validation was performed using transcriptome and clinical data of PTC patients derived from the Nucleotide Archive database, and the external validation was employed by inducing 11 patients who had undergone total thyroidectomy in Beijing Cancer Hospital[21]. In addition, Liu *et al.* and Lv *et al.* acquired the patients' data from the hospitals and performed further validation in the regulation of each molecule of the signature[29,47]. In addition to statistical validation methods, some studies also used experimental validation methods such as immunohistochemistry (IHC) and quantitative real-time polymerase chain reaction (qRT-PCR). As effective external validation methods, IHC and qRT-PCR were used to verify the differential expression of the biomarkers during the construction and validation of signatures[25,29,31,33,45,79]. As common methods of experimental verification, both of them could further verify the expression level of molecules in prognostic tags, making the signatures more convincing.

3.5. Nomogram based on prognostic signatures for TC

To demonstrate the prognostic abilities of signatures, some studies predicted the risk of various outcome variables of TC by constructing the risk assessment models: nomograms[19,27,36,38,56,78].

Nowadays, nomogram can be widely used for combining diagnosis with multiple indexes and also predict the recurrence and progression of many carcinomas, including PFI[19,38,56], RFS[78], DFS[36], and OS[27] for DTC, PTC, or TC patients. Based on the prognostic signature established by the multivariate regression hazard model and some relative clinical parameters, the nomogram could be developed. The assessment indexes including the C-index, AUC of ROC, and the calibration plots could evaluate the performance of the nomogram[19,36,38,78]. For instance, Pan *et al.* screened 5 genes to establish a prognostic signature and then validated its prognostic value for clinical applications. In addition, a nomogram of predicting DFS of PTC patients was constructed with a C-index of 0.797 (95% CI, 0.730–0.864), and AUCs for 1-, 3-, and 5-year DFS were 0.763, 0.777, and 0.755, respectively[36]. Chen *et al.* identified a DNA-methylated signature that has a good performance for RFS of PTC and then based on the risk score and multivariate regression model, a nomogram utilized in the clinic was constructed[78]. Wang

Table 5. Nomograms constructed for prognostic signatures of TC patients

Author	Signatures type	Pathological type	Survival event	Nomogram validation methods				References
				AUCs of the ROC curve (95% CI)	Harrell's concordance index (95% CI)	*P*-value of K-M analysis	Calibration plot	
Ruchong Pan, *et al.*, 2021	mRNA	DTC	DFS	AUCs for 1-, 3-, and 5-year DFS: 0.763, 0.777, and 0.755	0.797 (95% CI, 0.730 – 0.864)		Yes	Ruchong *et al.*, 2021[36]
Zhiwei Chen, *et al.*, 2021	mRNA	TC	OS				Yes	Chen *et al.*, 2021[79]
Mingqin Ge, *et al.*, 2021	mRNA	TC	OS				Yes	Ge *et al.*, 2021[27]
Mengwei Wu, *et al.*, 2020	mRNA	PTC	PFI	The AUCs for the 1-, 3-, and 5-year PFI: 0.855 (95% CI, 0.779 – 0.932), 0.799 (95% CI, 0.722 – 0.877), and 0.812 (95% CI, 0.718 – 0.907)	0.790 (95% CI, 0.652 – 0.927)	*P*<0.0001	Yes	Wu *et al.*, 2020[38]
Mengwei Wu, *et al.*, 2019	mRNA	PTC	PFI	The AUCs for the 1-, 3-, and 5-year PFI: 0.7480, 0.7550, and 0.7627	0.7600 (95% CI, 0.6759 – 0.8440)	*P*<0.0001	Yes	Wu *et al.*, 2019[19]
Kun Wang, *et al.*, 2020	mRNA/ lncRNA	PTC	PFI		0.792 (95% CI, 0.716 – 0.867)		Yes	Wang *et al.*, 2020[56]
Hengyu Chen, *et al.*, 2020	DNA methylation	PTC	RFS	The AUCs for 1-, 3-, and 5-year RFS: 0.850, 0.783, and 0.800	0.796 (95% CI, 0.704 – 0.888)		Yes	Chen *et al.*, 2020[78]

Abbreviations: PFI: Progression-free intervals; DFS: Disease-free survival; RFS: Recurrence-free survival; OS: Overall survival

Table 6. Repeatedly appeared prognostic genes of signatures were summarized

Name	Full Name	Frequency	Biological function	References
CDKN2A	Cyclin Dependent Kinase Inhibitor 2A	2	Cell circle regulation	Hu *et al.*, 2020[22] Wu *et al.*, 2020[38]
FXYD6	FXYD Domain Containing Ion Transport Regulator 6	2	Ion channel regulation and transmembrane transport	Wu *et al.*, 2019[19] Wu *et al.*, 2020[38]
CTGF	Connective Tissue Growth Factor	2	Promotion of proliferation and differentiation	Lin *et al.*, 2019[24] Ren *et al.*, 2021[35]
IGF2BP2	Recombinant Insulin Like Growth Factor 2 mRNA Binding Protein 2	2	RNA binding and translation regulation	Wang *et al.*, 2020[66] Ma *et al.*, 2021[30]
FTO	Alpha-ketoglutarate-dependent dioxygenase FTO	2	Mediation of different RNA oxidative demethylation	Hou *et al.*, 2020[49] Lin *et al.*, 2021[46]
RBM15	RNA-Binding Protein 15	2	Mediation of m6A RNA methylation	Hou *et al.*, 2020[49] Lin *et al.*, 2021[46]
KIAA1429	Vir Like M6A Methyltransferase Associated	2	Regulation of m6A methylation and modification of RNA processing	Hou *et al.*, 2020[49] Lin *et al.*, 2021[46]
AKR1C1	Aldo-Keto Reductase Family 1 Member C1	2	Regulation of hormone metabolism and catalytic reduction	Ge *et al.*, 2021[27] Huang *et al.*, 2021[34]
FN1	Fibronectin 1	2	Cell adhesion and the binding of compounds	Li *et al.*, 2021[27] Ruchong *et al.*, 2021[36]
FAM3B	FAM3 Metabolism Regulating Signaling Molecule B	2	Cell apoptosis induction	Lin *et al.*, 2019[24] Qin *et al.*, 2021[25]
TGFBR3	Transforming Growth Factor Beta Receptor 3	2	TGF-beta bind and capture	Wu *et al.*, 2019[19] Qin *et al.*, 2021[25]
HMGCR	3-Hydroxy-3-Methylglutaryl-CoA Reductase	2	Cellular cholesterol homeostasis maintaining	Ge *et al.*, 2021[27] Qian *et al.*, 2021[31]
GSS	Glutathione Synthetase	2	Glutathione biosynthesis	Ge *et al.*, 2021[27] Qian *et al.*, 2021[31]
TFRC	Transferrin Receptor	2	Mediation of the cellular uptake of iron	Ge *et al.*, 2021[27] Qian *et al.*, 2021[31]
DPP4	Dipeptidyl Peptidase 4	2	T-cell coactivation regulation	Ge *et al.*, 2021[27] Qian *et al.*, 2021[31]
PGD	Phosphogluconate Dehydrogenase	2	Catalysis of 6-phosphogluconate oxidative decarboxylation	Ge *et al.*, 2021[27] Qian *et al.*, 2021[31]
ADRA1B	Adrenoceptor Alpha 1B	2	Association with G proteins	Zhong *et al.*, 2020[40] Xu *et al.*, 2021[39]
PCOLCE2	Procollagen C-Endopeptidase Enhancer 2	2	C-terminal propeptide bind	Zhong *et al.*, 2020[40] Xu *et al.*, 2021[39]

Abbreviations: lncRNA: Long non-coding RNA; m6A: N6-methyladenosine

et al. built a nomogram based on the verified prognostic signatures and multivariate Cox proportional model. The C-indexes for measuring discrimination and the calibration plots for assessing accuracy were then used for verification of the nomogram model[56]. The elaborated summary of nomograms is shown in Table 5. Some studies also developed nomograms for prognostic assessment and clinical applications for TC patients[25,30,34,66].

3.6. Genes overrepresented in prognostic signatures

Through collation and summary, we found 18 genes that appeared repeatedly in previously reported prognostic signatures. These repeatedly appeared genes were involved in a wide range of biological functions, including immune-related genes[24,25], autophagy-related genes[22], m6A RNA methylation regulators[48,49,66], EMT-related genes[45], and ferroptosis-related genes[27,31], playing pivotal roles in mediating the elaborate process and mechanism of TC progression. The repetitions indicated that these genes would be closely related to the occurrence and development of TC, and also revealed that they might have more robust prognostic abilities of TC patients. The detailed information and data derived from studies are presented in Table 6.

4. Discussion

In the present review, we conducted a systematic review and an integrated analysis of previously reported prognostic signatures for TC. By comprehensive analysis, we summed up the construction processes of prognostic signatures for TC, classified the reported prognostic signatures, and summarized three main strategies for signature construction.

In particular, strategy 1 mainly focused on DEGs; strategy 2 confined the scope of signatures to specific biological functions, such as immune-related and ferroptosis-related genes[22,25,26,80], making these signatures more targeted to TC treatment; and strategy 3 mainly focused on epigenetic methylation, providing a new research direction for the prognosis of TC patients[47-49]. Among these, strategy 1 is the most extensive and common method of prognostic signature establishment. In addition to common types of prognostic signatures, plenty of signatures in the forms of ncRNAs, carbohydrate metabolism, and alternative splicing were explored as indicators of risk assessment[64,81].

To summarize, this review provides a comprehensive understanding, systematic summary, and integrated analysis of prognostic signatures for TC, which may help researchers further understand cancer progression and predict the recurrence risk of TC to improve the life quality of TC patients.

Acknowledgment

None.

Funding

This work was supported by the Program for Science and Technology Development in Henan Province (No.212102310616); Innovation Project for College Students of Henan University (No.20237003001; 20237003002; 20237003003; 202310475091; 20231022022); the Natural Science Foundation of Henan Province (No.232300421301); and the Program for Science and Technology Development in Kaifeng City (No.2203009).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Yang An
Investigation: Wenyuan Lu, Shijia Li, Jiahao Chen, Yingying Lian, Xinyu Li, Binbin Zhao
Methodology: Xiaoyan Lu, Yuanyuan Zhang, Pei Yang, Minjun Yi, Luyao Wang
Formal analysis: Xiaoyan Lu, Yuanyuan Zhang, Pei Yang, Minjun Yi, Luyao Wang, Yufei Jiang, Bingbing Guo
Data collection and curation: Jing Chen, Han Wang, Mengke Li
Software: Yingying Lian, Xinyu Li, Binbin Zhao
Project administration: Xiaoqing Wang
Funding acquisition: Xiaoqing Wang, Yang An
Validation: Yufei Jiang, Bingbing Guo, Wenyuan Lu, Shijia Li, Jiahao Chen
Writing – original draft: Xiaoyan Lu, Yuanyuan Zhang,
Writing – review & editing: Yang An

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Data availability

The data supporting that this systematic review was derived from previously reported studies, which were collected from the PubMed database and had been cited. The processed data are available from the corresponding author on request.

References

1. Antonelli A, La Motta C, 2017, Novel therapeutic clues in thyroid carcinomas: The role of targeting cancer stem cells. *Med Res Rev*, 37(6): 1299–1317.

   https://doi.org/10.1002/med.21448

2. Tofe S, Arguelles I, Forteza A, *et al.*, 2023, Age-standardized incidence, mortality rate, and trend changes of thyroid cancer in the Balearic Islands during the 2000–2020 period: A population-based study. *Eur Thyroid J*, 12(3): e220183.

   https://doi.org/10.1530/ETJ-22-0183

3. Gonzalez-Clavijo AM, Cuellar AA, Triana-Urrego J, *et al.*, 2023, Metastatic differentiated thyroid cancer: Worst prognosis in patients with metachronous metastases. *Endocrine*, 81(1): 90–97.

   https://doi.org/10.1007/s12020-023-03302-0

4. La Vecchia C, Malvezzi M, Bosetti C, *et al.*, 2015, Thyroid cancer mortality and incidence: A global overview. *Int J Cancer*, 36(9): 2187–2195.

   https://doi.org/10.1002/ijc.29251

5. Araque KA, Gubbi S, Klubo-Gwiezdzinska J, 2020, Updates on the management of thyroid cancer. *Horm Metab Res*, 52(8): 562–577.

   https://doi.org/10.1055/a-1089-7870

6. Chmielik E, Rusinek D, Oczko-Wojciechowska M, *et al.*, 2018, Heterogeneity of thyroid cancer. *Pathobiology*, 85(1–2): 117–129.

https://doi.org/10.1159/000486422

7. Lupoli GA, Fonderico F, Colarusso S, *et al.*, 2005, Current management of differentiated thyroid carcinoma. *Med Sci Monit*, 11(12): RA368–RA373.

8. Baloch ZW, Asa SL, Barletta JA, *et al.*, 2022, Overview of the 2022 WHO classification of thyroid neoplasms. *Endocr Pathol*, 33(1): 27–63.

https://doi.org/10.1007/s12022-022-09707-3

9. Cabanillas ME, McFadden DG, Durante C, 2016, Thyroid cancer. *Lancet*, 388(10061): 2783–2795.

https://doi.org/10.1016/S0140-6736(16)30172-6

10. Metovic J, Cabutti F, Osella-Abate S, *et al.*, 2023, Clinical and pathological features and gene expression profiles of clinically aggressive papillary thyroid carcinomas. *Endocr Pathol*.

https://doi.org/10.1007/s12022-023-09769-x

11. Liu Z, Zhao Q, Zeng W, *et al.*, 2018, Prognosis of a rare subtype of thyroid cancer: Spindle cell thyroid carcinoma. *Medicine* (*Baltimore*), 97(45): e13053.

https://doi.org/10.1097/MD.0000000000013053

12. Soares P, Celestino R, Melo M, *et al.*, 2014, Prognostic biomarkers in thyroid cancer. *Virchows Arch*, 464(3): 333–346.

https://doi.org/10.1007/s00428-013-1521-2

13. Yapa S, Mulla O, Green V, *et al.*, 2017, The role of chemokines in thyroid carcinoma. *Thyroid*, 27(11): 1347–1359.

https://doi.org/10.1089/thy.2016.0660

14. Wolff L, Steindl A, Popov P, *et al.*, 2023, Clinical characteristics, treatment, and long-term outcome of patients with brain metastases from thyroid cancer. *Clin Exp Metastasis*, 40(3): 217–226.

https://doi.org/10.1007/s10585-023-10208-8

15. Kang YJ, Stybayeva G, Hwang SH, 2023, Surgical completeness and safety of minimally invasive thyroidectomy in patients with thyroid cancer: A network meta-analysis. *Surgery*, 173(6): 1381–1390.

https://doi.org/10.1016/j.surg.2023.02.021

16. Sherman SI, 2003, Thyroid carcinoma. *Lancet*, 361(9356): 501–511.

https://doi.org/10.1016/s0140-6736(03)12488-9

17. Gulfidan G, Soylu M, Demirel D, *et al.*, 2022, Systems biomarkers for papillary thyroid cancer prognosis and treatment through multi-omics networks. *Arch Biochem Biophys*, 715: 109085.

https://doi.org/10.1016/j.abb.2021.109085

18. Patel J, Klopper J, Cottrill EE, 2023, Molecular diagnostics in the evaluation of thyroid nodules: Current use and prospective opportunities. *Front Endocrinol* (*Lausanne*), 14: 1101410.

https://doi.org/10.3389/fendo.2023.1101410

19. Wu M, Yuan H, Li X, *et al.*, 2019, Identification of a five-gene signature and establishment of a prognostic nomogram to predict progression-free interval of papillary thyroid carcinoma. *Front Endocrinol* (*Lausanne*), 10: 790.

https://doi.org/10.3389/fendo.2019.00790

20. Ruiz EM, Niu T, Zerfaoui M, *et al.*, 2020, A novel gene panel for prediction of lymph-node metastasis and recurrence in patients with thyroid cancer. *Surgery*, 167(1): 73–79.

https://doi.org/10.1016/j.surg.2019.06.058

21. Teng H, Mao F, Liang J, *et al.*, 2018, Transcriptomic signature associated with carcinogenesis and aggressiveness of papillary thyroid carcinoma. *Theranostics*, 8(16): 4345–4358.

https://doi.org/10.7150/thno.26862

22. Hu G, Feng HF, Zhan H, 2020, Identification of an autophagy-related signature predicting overall survival for papillary thyroid carcinoma. *Dose Response,* 18(1).

https://doi.org/10.1177/1559325819899265

23. Lin P, He Y, Wen DY, *et al.*, 2018, Comprehensive analysis of the clinical significance and prospective molecular mechanisms of differentially expressed autophagy-related genes in thyroid cancer. *Int J Oncol*, 53(2): 603–619.

https://doi.org/10.3892/ijo.2018.4404

24. Lin P, Guo YN, Shi L, *et al.*, 2019, Development of a prognostic index based on an immunogenomic landscape analysis of papillary thyroid cancer. *Aging* (*Albany NY*), 11(2): 480–500.

https://doi.org/10.18632/aging.101754

25. Qin R, Li C, Wang X, *et al.*, 2021, Identification and validation of an immune-related prognostic signature and key gene in papillary thyroid carcinoma. *Cancer Cell Int*, 21(1): 378.

https://doi.org/10.1186/s12935-021-02066-9

26. Xue Y, Li J, Lu X, 2020, A novel immune-related prognostic signature for thyroid carcinoma. *Technol Cancer Res Treat,* 19.

https://doi.org/10.1177/1533033820935860

27. Ge M, Niu J, Hu P, *et al.*, 2021, A ferroptosis-related signature robustly predicts clinical outcomes and associates with immune microenvironment for thyroid cancer. *Front Med* (*Lausanne*), 8: 637743.

https://doi.org/10.3389/fmed.2021.637743

28. Li Q, Wang P, Sun C, *et al.*, 2019, Integrative analysis of methylation and transcriptome identified epigenetically regulated lncRNAs with prognostic relevance for thyroid cancer. *Front Bioeng Biotechnol*, 7: 439.

https://doi.org/10.3389/fbioe.2019.00439

29. Liu T, You X, Sui J, *et al.*, 2019, Prognostic value of a two-microRNA signature for papillary thyroid cancer and a bioinformatic analysis of their possible functions. *J Cell Biochem*, 120: 7185–7198.

https://doi.org/10.1002/jcb.27993

30. Ma Y, Yin S, Liu XF, *et al.*, 2021, Comprehensive analysis of the functions and prognostic value of RNA-binding proteins in thyroid cancer. *Front Oncol*, 11: 625007.

https://doi.org/10.3389/fonc.2021.625007

31. Qian X, Tang J, Li L, *et al.*, 2021, A new ferroptosis-related gene model for prognostic prediction of papillary thyroid carcinoma. *Bioengineered*, 12(1): 2341–2351.

https://doi.org/10.1080/21655979.2021.1935400

32. Xu F, Xu H, Li Z, *et al.*, 2021, Glycolysis-based genes are potential biomarkers in thyroid cancer. *Front Oncol*, 11: 534838.

https://doi.org/10.3389/fonc.2021.534838

33. You X, Yang S, Sui J, *et al.*, 2018, Molecular characterization of papillary thyroid carcinoma: A potential three-lncRNA prognostic signature. *Cancer Manag Res*, 10: 4297–4310.

https://doi.org/10.2147/CMAR.S174874

34. Huang Y, Yi T, Liu Y, *et al.*, 2021, The landscape of tumors-infiltrate immune cells in papillary thyroid carcinoma and its prognostic value. *PeerJ*, 9: e11494.

https://doi.org/10.7717/peerj.11494

35. Ren H, Liu X, Li F, *et al.*, 2021, Identification of a six gene prognosis signature for papillary thyroid cancer using multi-omics methods and bioinformatics analysis. *Front Oncol*, 11: 624421.

https://doi.org/10.3389/fonc.2021.624421

36. Ruchong P, Haiping T, Xiang W, 2021, A five-gene prognostic nomogram predicting disease-free survival of differentiated thyroid cancer. *Dis Markers*, 2021: 5510780.

https://doi.org/10.1155/2021/5510780

37. Saftencu M, Braicu C, Cojocneanu R, *et al.*, 2019, Gene expression patterns unveil new insights in papillary thyroid cancer. *Medicina* (*Kaunas*), 55(8): 500.

https://doi.org/10.3390/medicina55080500

38. Wu M, Li S, Han J, *et al.*, 2020, Progression risk assessment of post-surgical papillary thyroid carcinoma based on circular RNA-associated competing endogenous RNA mechanisms. *Front Cell Dev Biol*, 8: 606327.

https://doi.org/10.3389/fcell.2020.606327

39. Xu L, Liu F, Li H, *et al.*, 2021, Comprehensive characterization of pathological stage-related genes of papillary thyroid cancer along with survival prediction. *J Cell Mol Med*, 25(17): 8390–8404.

https://doi.org/10.1111/jcmm.16799

40. Zhong LK, Gan XX, Deng XY, *et al.*, 2020, Potential five-mRNA signature model for the prediction of prognosis in patients with papillary thyroid carcinoma. *Oncol Lett*, 20(3): 2302–2310.

https://doi.org/10.3892/ol.2020.11781

41. Li Q, Li H, Zhang L, *et al.*, 2017, Identification of novel long non-coding RNA biomarkers for prognosis prediction of papillary thyroid cancer. *Oncotarget*, 8(28): 46136–46144.

https://doi.org/10.18632/oncotarget.17556

42. Luo YH, Liang L, He RQ, *et al.*, 2017, RNA-sequencing investigation identifies an effective risk score generated by three novel lncRNAs for the survival of papillary thyroid cancer patients. *Oncotarget*, 8(43): 74139–74158.

https://doi.org/10.18632/oncotarget.18274

43. Zhang Y, Jin T, Shen H, *et al.*, 2019, Identification of long non-coding RNA expression profiles and co-expression genes in thyroid carcinoma based on the cancer genome atlas (TCGA) database. *Med Sci Monit*, 25: 9752–9769.

https://doi.org/10.12659/MSM.917845

44. Chengfeng X, Gengming C, Junjia Z, *et al.*, 2019, MicroRNA signature predicts survival in papillary thyroid carcinoma. *J Cell Biochem*, 120(10): 17050–17058.

https://doi.org/10.1002/jcb.28966

45. Li Q, Jiang S, Feng T, *et al.*, 2021, Identification of the EMT-related genes signature for predicting occurrence and progression in thyroid cancer. *Onco Targets Ther*, 14: 3119–3131.

https://doi.org/10.2147/OTT.S301127

46. Lin R, Fogarty CE, Ma B, *et al.*, 2021, Identification of ferroptosis genes in immune infiltration and prognosis in thyroid papillary carcinoma using network analysis. *BMC Genomics*, 22(1): 576.

https://doi.org/10.1186/s12864-021-07895-6

47. Lv L, Cao L, Hu G, *et al.*, 2020, Methylation-driven genes identified as novel prognostic indicators for thyroid carcinoma. *Front Genet*, 11: 294.

https://doi.org/10.3389/fgene.2020.00294

48. Xu N, Chen J, He G, *et al.*, 2020, Prognostic values of m6A RNA methylation regulators in differentiated thyroid carcinoma. *J Cancer*, 11(17): 5187–5197.

https://doi.org/10.7150/jca.41193

49. Hou J, Shan H, Zhang Y, *et al.*, 2020, m^6A RNA methylation regulators have prognostic value in papillary thyroid carcinoma. *Am J Otolaryngol*, 41(4): 102547.

https://doi.org/10.1016/j.amjoto.2020.102547

50. Suh HY, Choi H, Paeng JC, *et al.*, 2019, Comprehensive gene expression analysis for exploring the association between glucose metabolism and differentiation of thyroid cancer.

BMC Cancer, 19(1): 1260.

https://doi.org/10.1186/s12885-019-6482-7

51. Han B, Yang M, Yang X, *et al.*, 2021, Systematic analysis of survival-associated alternative splicing signatures in thyroid carcinoma. *Front Oncol*, 11: 561457.

https://doi.org/10.3389/fonc.2021.561457

52. Lin P, He RQ, Huang ZG, *et al.*, 2019, Role of global aberrant alternative splicing events in papillary thyroid cancer prognosis. *Aging* (*Albany NY*), 11(7): 2082–2097.

https://doi.org/10.18632/aging.101902

53. Han Y, Yu X, Yin Y, *et al.*, 2021, Identification of potential BRAF inhibitor joint therapy targets in PTC based on WGCAN and DCGA. *J Cancer*, 12(6): 1779–1791.

https://doi.org/10.7150/jca.51551

54. Gandolfi G, Ragazzi M, de Biase D, *et al.*, 2018, Genome-wide profiling identifies the THYT1 signature as a distinctive feature of widely metastatic Papillary Thyroid Carcinomas. *Oncotarget*, 9(2): 1813–1825.

https://doi.org/10.18632/oncotarget.22805

55. Zhang Y, Zhang R, Liang F, 2020, Identification of metabolism-associated prostate cancer subtypes and construction of a prognostic risk model. *Front Oncol*, 10: 598801.

https://doi.org/10.3389/fonc.2020.598801

56. Wang K, Xu J, Zhao L, *et al.*, 2020, Prognostic lncRNA, miRNA, and mRNA signatures in papillary thyroid carcinoma. *Front Genet*, 11: 805.

https://doi.org/10.3389/fgene.2020.00805

57. Yu H, Guo P, Xie X, *et al.*, 2017, Ferroptosis, a new form of cell death, and its relationships with tumourous diseases. *J Cell Mol Med*, 21(4): 648–657.

https://doi.org/10.1111/jcmm.13008

58. Murakami H, Hayashi M, Terada S, *et al.*, 2023, Medroxyprogesterone acetate-resistant endometrial cancer cells are susceptible to ferroptosis inducers. *Life Sci*, 325: 121753.

https://doi.org/10.1016/j.lfs.2023.121753

59. Cao JY, Dixon SJ, 2016, Mechanisms of ferroptosis. *Cell Mol Life Sci*, 73(11–12): 2195–2209.

https://doi.org/10.1007/s00018-016-2194-1

60. Kim SE, Zhang L, Ma K, *et al.*, 2016, Ultrasmall nanoparticles induce ferroptosis in nutrient-deprived cancer cells and suppress tumour growth. *Nat Nanotechnol*, 11(11): 977–985.

https://doi.org/10.1038/nnano.2016.164

61. Dongre A, Weinberg RA, 2019, New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat Rev Mol Cell Biol*, 20(2): 69–84.

https://doi.org/10.1038/s41580-018-0080-4

62. Dawson MA, Kouzarides T, 2012, Cancer epigenetics: From mechanism to therapy. *Cell*, 150(1): 12–27.

https://doi.org/10.1016/j.cell.2012.06.013

63. Ilango S, Paital B, Jayachandran P, *et al.*, 2020, Epigenetic alterations in cancer. *Front Biosci*, (*Landmark Ed*), 25(6): 1058–1109.

https://doi.org/10.2741/4847

64. Zafon C, Gil J, Perez-Gonzalez B, *et al.*, 2019, DNA methylation in thyroid cancer. *Endocr Relat Cancer*, 26(7): R415–R439.

https://doi.org/10.1530/ERC-19-0093

65. Russo D, Damante G, Puxeddu E, *et al.*, 2011, Epigenetics of thyroid cancer and novel therapeutic targets. *J Mol Endocrinol*, 46(3): R73–R81.

https://doi.org/10.1530/JME-10-0150

66. Wang X, Fu X, Zhang J, *et al.*, 2020, Identification and validation of m^{6}A RNA methylation regulators with clinical prognostic value in Papillary thyroid cancer. *Cancer Cell Int*, 20: 203.

https://doi.org/10.1186/s12935-020-01283-y

67. Kushchayeva Y, Kushchayev S, Jensen K, *et al.*, 2022, Impaired glucose metabolism, anti-diabetes medications, and risk of thyroid cancer. *Cancers* (*Basel*), 14(3): 555.

https://doi.org/10.3390/cancers14030555

68. Heydarzadeh S, Moshtaghie AA, Daneshpoor M, *et al.*, 2020, Regulators of glucose uptake in thyroid cancer cell lines. *Cell Commun Signal*, 18(1): 83.

https://doi.org/10.1186/s12964-020-00586-x

69. Davidson CD, Tomczak JA, Amiel E, *et al.*, 2022, Inhibition of glycogen metabolism induces reactive oxygen species-dependent cytotoxicity in anaplastic thyroid cancer in female mice. *Endocrinology*, 163(12): bqac169.

https://doi.org/10.1210/endocr/bqac169

70. Matsuzu K, Segade F, Matsuzu U, *et al.*, 2004, Differential expression of glucose transporters in normal and pathologic thyroid tissue. *Thyroid*, 14(10): 806–812.

https://doi.org/10.1089/thy.2004.14.806

71. Marima R, Francies FZ, Hull R, *et al.*, 2021, MicroRNA and alternative mRNA splicing events in cancer drug response/resistance: Potent therapeutic targets. *Biomedicines*, 9(12): 1818.

https://doi.org/10.3390/biomedicines9121818

72. Climente-Gonzalez H, Porta-Pardo E, Godzik A, *et al.*, 2017, The functional impact of alternative splicing in cancer. *Cell Rep*, 20(9): 2215–2226.

https://doi.org/10.1016/j.celrep.2017.08.012

73. Park J, Kim D, Lee JO, *et al.*, 2022, Dissection of molecular and histological subtypes of papillary thyroid cancer using alternative splicing profiles. *Exp Mol Med*, 54(3): 263–272.

https://doi.org/10.1038/s12276-022-00740-0

74. Oltean S, Bates DO, 2014, Hallmarks of alternative splicing in cancer. *Oncogene*, 33(46): 5311–5318.

https://doi.org/10.1038/onc.2013.533

75. Abdullah MI, Junit SM, Ng KL, *et al.*, 2019, Papillary thyroid cancer: Genetic alterations and molecular biomarker investigations. *Int J Med Sci*, 16(3): 450–460.

https://doi.org/10.7150/ijms.29935

76. Nikiforov YE, Nikiforova MN, 2011, Molecular genetics and diagnosis of thyroid cancer. *Nat Rev Endocrinol*, 7(10): 569–580.

https://doi.org/10.1038/nrendo.2011.142

77. Zantut-Wittmann DE, Laus AC, Moreno DA, *et al.*, 2023, Extremely aggressive course in a poorly differentiated thyroid carcinoma presenting a double mutation of the TERT promoter. *Am J Med Sci*, 365(6): 532–537.

https://doi.org/10.1016/j.amjms.2023.03.019

78. Chen H, Ma X, Yang M, 2020, A methylomics-associated nomogram predicts recurrence-free survival of thyroid papillary carcinoma. *Cancer Med*, 9(19): 7183–7193.

https://doi.org/10.1002/cam4.3388

79. Chen Z, Liu X, Liu F, *et al.*, 2021, Identification of 4-methylation driven genes based prognostic signature in thyroid cancer: An integrative analysis based on the methylmix algorithm. *Aging* (*Albany NY*), 13(16): 20164–20178.

https://doi.org/10.18632/aging.203338

80. Mou Y, Wang J, Wu J, *et al.*, 2019, Ferroptosis, a new form of cell death: Opportunities and challenges in cancer. *J Hematol Oncol*, 12(1): 34.

https://doi.org/10.1186/s13045-019-0720-y

81. Zhao H, De Souza C, Kumar VE, *et al.*, 2021, Long non-coding RNA signatures as predictors of prognosis in thyroid cancer: A narrative review. *Ann Transl Med*, 9(4): 359.

https://doi.org/10.21037/atm-20-8191

ACCSCIENCE PUBLISHING

Gene & Protein in Disease

ORIGINAL RESEARCH ARTICLE

Analysis of multi-disease targeting effect of phytochemicals by AMPK stimulation–diabetes: A computational approach

Richa Goyal[1]*, Manoj Kumar[2], and Muhammad Anwar Mallick[3]

[1]Department of Biotechnology, Vinoba Bhave University, Hazaribagh, Jharkhand, India

[2]Department of Zoology, Vinoba Bhave University, Hazaribagh, Jharkhand, India

[3]Department of Botany/ Biotechnology, Vinoba Bhave University, Hazaribagh, Jharkhand, India

Abstract

Diabetes is a silent killer and a metabolic syndrome characterized by hyperglycemia that has been exponentially increasing in recent years. There is a need to develop therapeutic agents to control hyperglycemia and its secondary complications as well as protect and revive beta cells in diabetic patients. The target for first-line diabetes treatment is the adenosine monophosphate protein kinase (AMPK), which participates in cellular energy metabolism through phosphorylation of metabolic enzymes and transcription regulators. This study examined the drug-related properties as well as lead preparation of *Catharanthus roseus* alkaloids and testing molecular interaction at the AMPK targets to confirm their anti-diabetic effect. A control drug metformin and a library of 85 molecules of *C. roseus* alkaloids were crossed with the ADMET test, followed by the investigation of molecular interaction tested on AMPK1 and AMPK2 targets through an *in silico* docking process. Vindolinine (CID: 24148538), vindoline (CID: 425978), (+)-vindorosine (CID: 261578), Cr-1 (CID: 5315746), and Cr-2 (CID: 59908094) had passed the ADMET test. Molecular interaction of the tested *C. roseus* alkaloids on AMPK1 and AMPK2 targets had potential energy that varied from −7.4 to −5.3 kcal/mol, whereas binding energies of −4.0 kcal/mol for AMPK1-metformin interaction and −4.2 kcal/mol for AMPK2-metformin interaction were observed. The tested *C. roseus* alkaloids were shown to be more potent activators of AMPK than the control drug. All five biomolecules of *C. roseus* acted as modulators that have the potential to stimulate AMPK, reduce glucose production, and increase glucose utilization in hepatocytes. In addition, they diminished insulin resistance and secondary complications of diabetes by inhibiting acetyl-CoA carboxylase, regulating cholesterol levels and macrophage, and reviving beta cells in Type 2 diabetes. These results provided the foundation for developing new multi-disease-targeting drugs that can treat diabetes, obesity, cardiovascular disease, cancer, and other diseases by the stimulation of AMPK1 and AMPK2 targets.

Keywords: Diabetes; Adenosine monophosphate protein kinase; *Catharanthus roseus* alkaloids; Metformin; Docking study

***Corresponding author:**
Richa Goyal
(richagoyalaiims@gmail.com)

Citation: Goyal R, Kumar M, Mallick MA, 2023, An analysis of multi-disease targeting effect of phytochemicals by AMPK stimulation – Diabetes: A computational approach. *Gene Protein Dis,* 2(3): 0927.
https://doi.org/10.36922/gpd.0927

Received: May 8, 2023
Accepted: August 21, 2023
Published Online: September 12, 2023

1. Introduction

Diabetes mellitus is a disorder of metabolism and is characterized by hyperglycemia in which impaired carbohydrate, fat, and protein metabolism takes place due to insulin deficiency or insulin resistance[1,2]. The increased plasma glucose promotes glucotoxicity which leads to tissue injury, sclerosis of glomerular capillaries, retinopathy and blindness, neuropathy and peripheral vascular insufficiency, gangrene of limbs, and other pathological changes and secondary complications in multiple organ systems such as hypertension, vascular complications, early atherosclerosis, and heart attack[1-3]. The onset of diabetes cases is increasing and doubling every decade. The cases have been magnified from 30 million to 463 million from 1985 to 2019 and may become 578 million and 700 million by 2030 and 2045, respectively[3,4]. Diabetes is the second most common disease in India, where approximately 1 million diabetic cases are found in Jharkhand[5]. This disease is the ninth major cause of death worldwide, and the international expenses on diabetes and related health were 11% in 2013[3], and about 12% (> $550 billion) in 2019[6,7]. Diabetes mellitus is broadly categorized into two groups: Type I (5% of the diabetic population, genetic origin) and Type II (95% of the diabetic population). The majority of prevalent cases are related to Type II diabetes caused by environmental factors instead of genetic background[4,8].

Adenosine monophosphate protein kinase (AMPK) stimulation has a significant role in cellular energy homeostasis or metabolism. The target is present in several tissues including the liver, brain, cardiac muscle, and skeletal muscle[1,9]. AMPK is a member of the serine/threonine protein kinase family and is turned on by the stimuli that raise the cellular AMP: ATP ratio. AMPK is a heterotrimeric protein (αβγ). Its catalytic alpha subunit has two isoforms alpha-1 and alpha-2 which control the actions of several vital metabolic enzymes through phosphorylation[10,11]. Alpha-1 isoform of AMPK (AMPK1) plays an important role in safeguarding cells from stresses that bring out ATP depletion by turning off ATP-consuming biosynthetic pathways, regulating insulin release[12,13], immune-suppression, differentiation of myeloid-derived suppressor cells (MDSC)[14], prohibiting skeletal muscle hypertrophy[15], regulation of mTOR signaling pathway, and macrophage proliferation[16]. Alpha-2 isoform of AMPK (AMPK2) has a role in the invigilation of cellular energy status, regulates insulin sensitivity, recycling of insulin receptors, phosphorylation of metabolic enzymes, deactivates acetyl-CoA carboxylase (ACC) and beta-hydroxyl beta-methylglutaryl-CoA reductase (HMGCR) enzymes[17], control the biosynthesis of fatty acid and cholesterol, sustain myocardial energy homeostasis during ischemia insulin secretion[18], and transcriptional activation of preproinsulin gene[13]. Food-starved cells activate AMPK with sugar and the activated target promotes glucose uptake by modulating insulin secretion and decreasing hepatic gluconeogenesis[1]. Metformin (N,N-dimethyl biguanide) is an oral antihyperglycemic drug that stimulates AMPK[19,20], despite having side effects like late assimilation of hexoses (e.g., glucose), amino acids (AA), and intestinal absorption of vitamin B12, abdominal pain, vomiting, malnourishment, swelling, mild diarrhea, and metallic taste, as well as tiredness[21].

Plants have a large number of biomolecules having a broad range of structure variation and medicative actions with no or fewer side effects and toxicity in comparison to synthetic drugs, and the ability to care for a wide variety of illnesses and infections[7,22,23]. Literature reported about 1200 plants and their components having antihyperglycemic properties[24-26]. *Catharanthus roseus*, formerly known as *Vinca rosea*, commonly known as Madagascar periwinkle and sadabahar in India, belongs to the Apocynaceae family and has antitumor, antimicrobial, and antihyperglycemic effects due to the presence of a large number of secondary metabolites (eg; alkaloids, phenols, etc.). *C. roseus* alkaloids can reduce blood glucose levels and glucotoxicity-induced secondary complications[27-34]. The pharmacodynamics of *C. roseus* alkaloids at the molecular level is still unexplored. Hence, this study aimed to address the AMPK1/2 targets and *C. roseus* alkaloids interaction concerning control through an *in silico* method. The exploration of the appropriate structure of ligands and the mechanism of action provided evidence of whether the compounds interact with the target and are able to reduce high blood sugar levels better than metformin. Further wet lab and clinical validations (animal and human trials) will be performed to examine the effectiveness and potency of individual *C. roseus* alkaloids in natural conditions against Type II diabetes and other linked diseases.

2. Methodology

2.1. Drug target selection

With the help of the literature and human model criteria, AMPK (UniProt ID: Q13131, P54646) was chosen and unloaded from UniProt[17]. SWISS-MODEL servers were used for homology modeling of proteins[35].

2.2. Detection of physiochemical properties of the target protein

Online software ProtParam was used for target protein sequence analysis and its physical and chemical parameters determination[36]. The self-optimized prediction method (SOPMA) was used for the prediction of secondary structures

of proteins[37]. The target protein nature was determined by the SOSUI (engine ver. 1.11) server[38]. The SWISS-MODEL server website was used for downloading a protein crystal three-dimensional (3D) structure[35]. SWISS-MODEL is a comparison modeling assistant processor and was applied to overlap proteins to assume its structural sequence or arrangement, sequence identity, coverage, model quality estimation (QMEAN SCORES), process, and resolution facts of the target protein. Ramachandran Z-score and Ramachandran plot ZLab server were used to determine protein structure along with phi-psi background probabilities[39]. CASTp 3.0 software was used to identify active sites and functional domains of targets[40].

2.3. Target protein preparation

Swiss-Pdb Viewer (aka Deep View) version 4.1.1 was applied to reduce the target protein's energy[41]. UCSF Chimera tool is an extensible molecular modeling system and was used for the addition of hydrogen and charges, to adjust atomic-subatomic torsion, degree of free will/freedom and stereo-chemical difference, and Gasteiger charges[42]. Auto Dock Vina and Chimera were used for grid preparation with 20 Å × 20 Å × 20 Å grid dimensions[43], and the X, Y, and Z coordinates of the grid box were −76.407, −36.793, 13.714 for AMPK1, respectively, and −36.303, −82.337, −22.555 for AMPK2, respectively[40].

2.4. The ligand library preparation

An extensive literature survey related to *C. roseus* revealed its major antihyperglycemic properties related to alkaloids[29,30,44]. The PubChem database was used to search for alkaloid structure and control drug metformin. Eighty-five compounds were selected from the database to prepare a library and metformin (CID: 4091) as a control drug. Followed by SDF, files were converted and saved in PDB files by online software, Open Babel[45].

2.5. Absorption, distribution, metabolism, excretion, and toxicity test

Absorption and assimilation through the mouth and digestive tract, distribution, metabolism, excretion, and toxicity (ADMET) test reveal the pharmacokinetic properties of compounds, that is, what the body does with compounds. The application programs OSIRIS Property Explorer and DruLiTo software were used to test the ADMET property of all library compounds[46,47]. According to the drug-likeness rules, the log S value estimates solubility whereas the cLogP value tells about the Lipinski rule, total polar surface area (TPSA), molecular mass, drug's likeness properties (functional portion of ligand), and drug score. The pharmacodynamics activities of a compound (what the drug does in the body) are predicted by the bioavailability of a ligand at the target region, chances of passage of the blood-brain barrier, toxicity, malignancy, etc.

2.6. Leads groundwork for docking

Avogadro 1.2.on-win 32 was applied for force field and geometry maximization of ligand[48]. Hydrogen and charge were then added with the help of The UCSF Chimera tool[42]. Torsion or tortuosities and level of freedom, followed by stereo-chemical distinction of ligands, were accustomed, Gasteiger charges were considered, and the PDB file was saved.

2.7. Examination of target-lead interaction/docking studies

The finally prepared PDB file of the target protein and lead was opened in the UCSF Chimera tool. The Vina tool was used to calculate the potential energy of the docked compound and the file was saved in PDBQT format. BIOVIA Discovery Studio Visualizer 4.5 was used to visualize target-ligand interactions[49]. With the help of the Discovery studio bonds detail, two-dimensional (2D) and 3D complex structures were analyzed.

3. Results

3.1. Target protein examination

To prepare a strong and stable model of the target protein, the selected targets and their properties were examined critically. 5'-AMPK protein was selected as a target, and its catalytic subunits AMPK1 and AMPK2 sequence and active site were taken from and analyzed by UniProt site (Table 1). AMPK1 is a globular protein with an enzyme commission number (EC) of 2.7.11.1[38], whereas AMPK2 is a transmembrane protein with an EC number of 2.7.11.1[50].

The online tool ProtPram was used for AMPK1 and AMPK2 sequence analysis and their physicochemical properties identification (Table 1). ProtPram grand average of hydropathicity (GRAVY) value indicates the globular (<0 for hydrophilic protein) or membranous (>0 for hydrophobic protein) nature of the protein. The value for AMPK1 GRAVY was −0.461, whereas the value for AMPK2 GRAVY was −0.295. The SOPMA tool was used to know about the target proteins' secondary structures (Table 2). AMPK1 has more α-helixes but fewer β turns and random coil than AMPK2. The soluble (AMPK1) and transmembrane (AMPK2) nature of the target was identified by SOSUI prediction, software details are given in Table 3.

The soluble (AMPK1) and transmembrane (AMPK2) nature of the target was identified by SOSUI prediction, software details are given in Table 3. Followed by validation of the target protein was done by the Ramachandran plot, which shows phi (ϕ) and psi (ψ) angle and protein

backbone confirmation, and Ramachandran Z score (−3 to +3) supporting good quality protein backbone confirmation (Table 4). AMPK1 (Template-4rer.1.A) and AMPK2 (Template-7myj.2.A) have > 95% highly preferred region and the Z-score value lies between −3.6 and −2.1. Hence, it was predicted that the selected target protein has a good backbone. QMEAN scores (represents the degree of nativeness of protein) for AMPK1 are 0.68 ± 0.05 and for AMPK2 0.73 ± 0.05 (Table 5). Greater than 0.6 QMEAN scores for the model protein indicated good quality/more nativeness of protein, thus, the selected target can be used for model preparation.

The active site, surface features, and functional region of the target protein were identified by CASTp, an online tool, and it shows the occurrence of basic, acidic, and neutral types of amino acids in the active site. From the catalytic site amino acids, residues Arg199 for AMPK1 and Glu264 for AMPK2 were selected based on lesser binding energy among 5 random selections.

3.2. Examination of ligand library and lead identification

From the PubChem database[50], the SDF file containing the structure and details of *C. roseus* alkaloids and related

Table 1. Physiochemical property of target proteins (UniProt and ProtParam value)

Protein	UniProt/PubMed ID	Length (AA)	Molecular weight	pI	Instability index	Aliphatic index (AI)	GRAVY	Extinction coefficients (M^{-1} cm^{-1}, at 280 nm)	$T_{1/2}$
AMPK1	AAPK1_HUMAN/Q13131	559 AA	64 kDa	8.32	49.44	84.03	−0.461	60,905	30 h
AMPK2	AAPK2_HUMAN/P54646	552 AA	62 kDa	7.65	49.20	85.11	−0.295	53,290	30 h

Abbreviations: AA: Amino acid; pI: Isoelectric point; GRAVY: Grand average of hydropathicity; $T_{1/2}$: Estimated half-life (mammalian reticulocytes, *in vitro*)

Table 2. Secondary structure detail of target protein (SOPMA result)

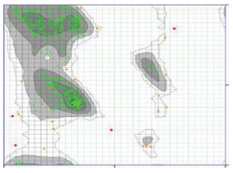
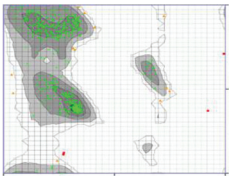
Proteins	Length (AA)	α-helix (Hh)	310 helix (Gg)	π-helix (Ii)	β-bridge (Bb)	Extended strand (Ee)	β-turn (Tt)	Bend region (Ss)	Random coil (Cc)	Ambiguous states	Other states
AMPK1	559 AA	30.95%	0.00%	0.00%	0.00%	17.35%	6.80%	0.00%	44.90%	0.00%	0.00%
AMPK2	552 AA	27.72%	0.00%	0.00%	0.00%	17.39%	8.15%	0.00%	46.74%	0.00%	0.00%

Abbreviation: SOPMA: Self-optimized prediction method

Table 3. Nature of target protein based on SOSUI prediction

Receptor	Protein nature	No	Region	Transmembrane sequence	Type
AMPK1	Soluble protein				
AMPK2	Membrane Protein	1	6–28	KVLLGLLGAAALVTIITVPVVLL	Primary

Table 4. Ramachandran Z-score and Ramachandran plot of target proteins

S. No.	Receptor	Ramachandran Z-score	Ramachandran plot		
1	AMPK1 (Template: 4rer. 1.A)	−3.583	Highly preferred observations (Green Crosses) Preferred observations (Brown Triangles) Questionable observations (Red Circles)	463 (95.859%) 16 (3.313%) 4 (0.828%)	
2	AMPK2 (Template: 7myj. 2.A)	−2.135	Highly preferred observations (Green Crosses) Preferred observations (Brown Triangles) Questionable observations (Red Circles)	464 (96.667%) 12 (2.500%) 4 (0.833%)	

Note: The Ramachandran Z-score showed how normal a model is as compared to the reference, a score of±3 indicated good protein backbone confirmation.

structures was downloaded and a total of 85 ligands were examined for ADMET testout and drug-likeness features. The qualified ligands and control compound details are shown in Table 6 and they were then used for lead preparation and docking study.

3.3. ADMET and drug-likeness calculation of ligands

Some online tools were used for the drug-likeness predictions. Out of 85 ligands, only five ligands, vindoline (CID: 425978), vindolinine (CID: 24148538), (+)-vindorosine (CID: 261578), Cr-1 (CID: 5315746), and Cr-2 (CID: 59908094) were qualified in the toxicity test (mutagenic, carcinogenic, irritation, and reproductive properties). The high, medium, and no-risk zones were shown in red, yellow, and green colors, respectively. The no-risk zone shows that the qualified biomolecules have low or non-toxic effects. The OSIRIS Property Explorer estimated values and DruLiTo software customized values details are presented in Table 7. The data indicated that the biomolecules have drug-like properties.

3.4. Molecular docking investigation

3.4.1. AMPK1 targets interaction with lead and control

Intermolecular interaction between metformin (control) and AMPK1 target has been confirmed by their docked poses (Figure 1) and had a binding affinity of -4 kcal/mol (see Table 8). Figure 1A shows the interaction between chain A of AMPK1, pocket residue ARG199, and the control drug metformin. The complex is formed between them with interactions of two conventional H-bonds of length 2.45 Å and 2.20 Å with Leu200; six attractive charge interactions with Glu205 and Glu194 at lengths of 5.57 Å, 4.27 Å, 3.95 Å, 3.06 Å, 4.79 Å, and 2.96 Å; a salt bridge of 1.84 Å with Glu205 residue; Phe180 of target forms two π-cation interactions of 4.16 Å and 4.91 Å; and an unfavorable donor-donor interaction of 2.55 Å with Ala202 (Table S1).

Figure 1B presents the binding of vindolinine (CID: 24148538) through the selected pocket site (residue Arg199) of AMPK1 chain A through non-covalent interactions and had a binding affinity of −7.4 kcal/mol (Table 8). In between the target and the lead, there are three conventional H-bond interactions with Arg182 and Leu462 of 2.66 Å, 3.08 Å, and 2.20 Å; two π-alkyl interactions with Leu462 and Tyr463 of 5.39 Å and 5.14 Å, respectively; one π-π stacked interaction with Tyr441 of 4.13 Å; and an unfavorable positive-positive interaction of 3.97 Å with Arg182 (Table S1).

Figure 1C shows the attachment of the lead vindoline (CID: 425978) with the active site of AMPK1 with non-covalent interactions and having a binding affinity of −6.3 kcal/mol. Their docking pose displays the presence of strong intermolecular interactions where Leu462 and Arg182 residues of the target form two conventional H-bond interactions of 2.31 Å and 2.30 Å, respectively; Gly198 and Gln461 residues form three carbon H-bond interactions of 3.62 Å, 3.59 Å, and 3.56 Å; Leu200 residue forms an alkyl interaction of 5.20 Å; and Tyr463 of AMPK1

Table 5. QMEAN scores for the model protein generated using the SWISS-MODEL server

Receptor	Template	Model quality estimation (QMEAN scores)	Sequence identity	Coverage	Method	Resolution
AMPK 1	4rer. 1.A	0.68±0.05	99.07	0.97	X-ray	4.05Å
AMPK 2	7myj. 2.A	0.73±0.05	99.82	1.00	X-ray	2.95Å

Note: QMEAN scores referred to the degree of nativeness of protein, a score of>0.6 indicated good quality model protein.

Table 6. Ligands selected for lead preparation

Ligand	PubChem CID	IUPAC name	Molecular weight (g/mol)
Vindolinine	CID: 24148538	Methyl (1R,9R,10R,12R,19S,20R)-20-methyl-8,16-diazahexacyclo [10.6.1.19,12.01,9.02,7.016,19] icosa-2,4,6,13-tetraene-10-carboxylate	336.4
Vindoline	CID: 425978	Methyl 11-acetyloxy-12-ethyl-10-hydroxy-5-methoxy-8-methyl-8,16 diazapentacyclo [10.6.1.01,9.02,7.016,19] nonadeca-2 (7),3,5,13-tetraene-10-carboxylate	456.5
(+)-Vindorosine	CID: 261578	1H-Indolizino (8,1-cd) carbazole-5-carboxylic acid, 3a-ethyl-3a, 4,5,5a, 6,11,12,13a-octahydro-4,5-dihydroxy-6-methyl-, methyl ester, 4-acetate	426.5
CR-1	CID: 5315746	Methyl 12-ethyl-10,11-dihydroxy-8-methyl-8,16-diazapentacyclo [10.6.1.01,9.02,7.016,19] nonadeca-2,4,6,13-tetraene-10-carboxylate	384.5
CR-2	CID: 59908094	(4-Tert-butyl-12-ethyl-10-formyl-10-hydroxy-5-methoxy-8-methyl-8, 16-diazapentacyclo [10.6.1.01,9.02,7.016,19] nonadeca-2,4,6,13-tetraen-11-yl) acetate	482.6
Metformin	CID: 4091	3-(diaminomethylidene)-1,1-dimethylguanidine	129.16

Table 7. ADMET and drug-likeness prediction results of qualified compounds (OSIRIS Property Explorer and DruLiTo tools result)

Compounds	PubChem CID	Toxicity Risk	cLogP	LogS	Molecular weight	TPSA	Drug likeness	Drug score	BBB
Vindolinine	CID: 24148538	G	1.7	−3.38	336	41.57	2.36	0.49	Pass
Vindoline	CID: 425978	G	1.32	−3.12	456	88.54	3.95	0.74	Fail
(+)-Vindorosine	CID: 261578	G	1.39	−3.1	426	79.31	3.86	0.78	Fail
Cr-1	CID: 5315746	G	0.91	−2.69	384	73.24	3.95	0.84	Pass
Cr-2	CID: 59908094	G	2.61	−4.38	482	79.31	0.58	0.5	Fail
Metformin (control drug)	CID: 4091	G	−1.54	−0.13	129	91.49	1.21	0.88	Pass

Abbreviations: cLogP (O/W), Logarithm of partition coefficient between n-octanol and water; LogS, aqueous solubility; TPSA, Topological polar surface area; G, Green signal (indicated no toxicity risk).

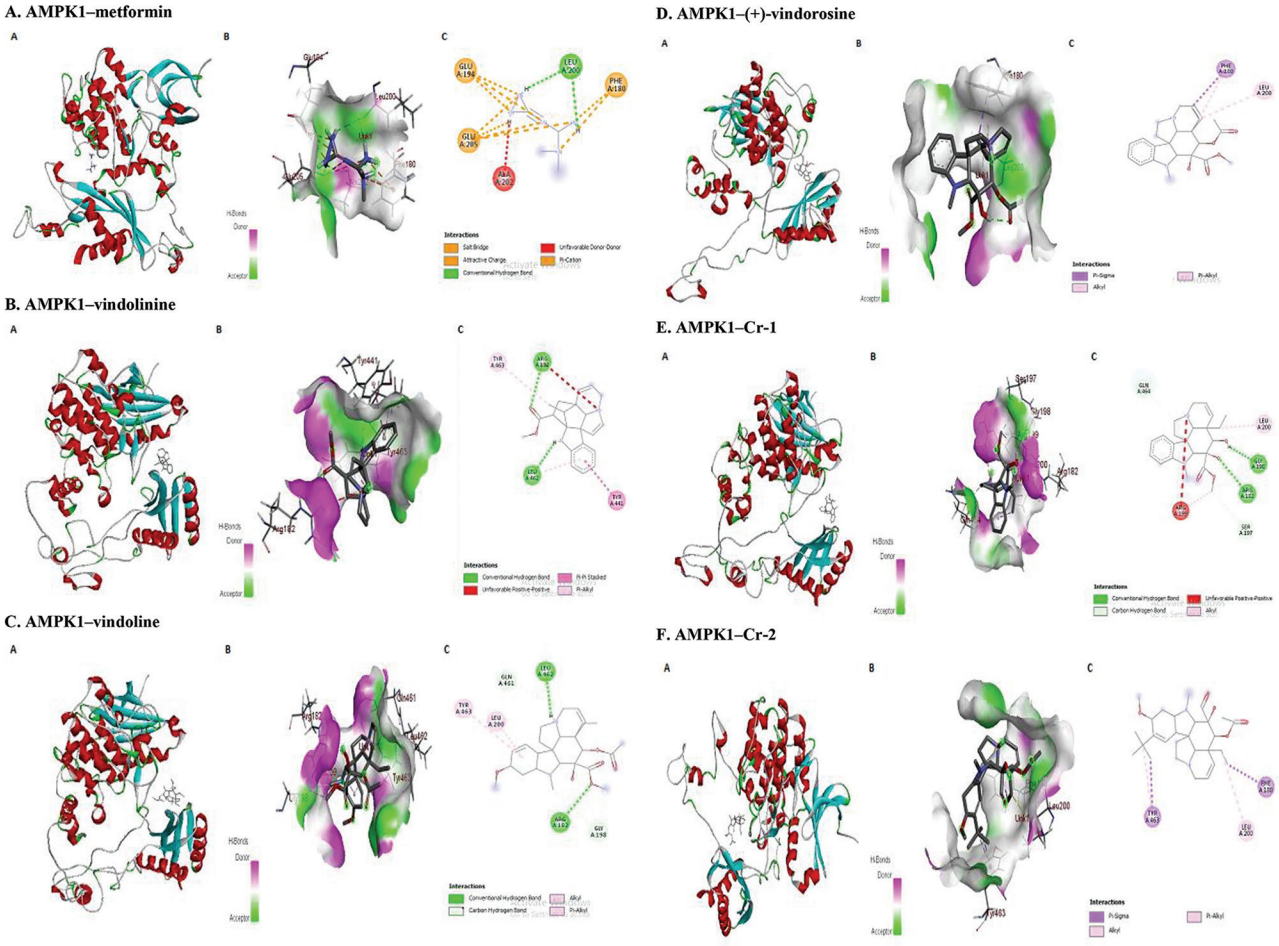


Figure 1. Molecular docking studies of AMPK1 with different ligands. (A) Interaction of AMPK1 with metformin (CID: 4091); (B) Interaction of AMPK1 with vindolinine (CID: 24148538); (C) Interaction of AMPK1 with vindoline (CID: 425978); (D) Interaction of AMPK1 with (+)-vindorosine (CID: 261578); (E) Interaction of AMPK1 with Cr-1(CID-5315746); and (F) Interaction of AMPK1 and Cr-2 (CID: 59908094). In subfigures, Image A shows about specific and reversible interaction of drug target AMPK1 and ligand/lead compound; Image B provides the insight of depth of the catalytic site of the target, similar to a real situation and distance between ligand and target around them, and target-ligand interaction within the active site of AMPK1 target based on H-bond donor and acceptor characteristic of amino acid residues; and Image C representing the 2D image, which is signifying the linkage between ligand and target amino acids residues. On the 2D map, the color lines indicated the following: Conventional H-bond interaction (green); carbon-hydrogen bond (sky blue); attractive charge interaction, π-cation interaction, and salt bridge (orange); alkyl interaction, π-alkyl interaction, and π-π stacked interaction (pink); conventional π-σ interaction (purple); and unfavorable positive-positive interaction (red).

Table 8. The binding energy of docking interaction between AMPK1 target and ligands (control and test compounds)

Group	Docked compound	Binding energy (kcal/mol)
Control	AMPK1–metformin	−4.0
Test	AMPK1–vindolinine	−7.4
	AMPK1–vindoline	−6.3
	AMPK1–(+)-vindorosine	−6.3
	AMPK1–Cr-1	−6.4
	AMPK1–Cr-2	−6.1

forms a π-alkyl interaction of 5.47 Å between the target and the lead (Table S1).

The molecular docked pose of (+)-vindorosine (CID: 261578) and AMPK1 (Figure 1D) showed that the lead attaches within a selected pocket site (residue Arg199) of AMPK1 chain A through non-covalent interactions and has a binding affinity of –6.3 kcal/mol. There is one intramolecular conventional H-bond interaction of 2.08 Å; a π-σ interaction with Phe180 residues of 3.5 Å; a π-alkyl interaction of 4.48 Å with a Phe180 residue; and an alkyl interaction with Leu200 residues of 4.61 Å (Table S1).

The docking pose (Figure 1E) reveals that the binding of Cr-1 (CID: 5315746) around Arg199 residues of chain A of AMPK1 through non-covalent links and had a binding affinity of −6.4 kcal/mol. There are two conventional H-bond interactions with Gly198 and Arg182 of 1.95 Å and 2.57 Å, respectively; three carbon H-bonds with Ser197, Gly198, and Gln464 of 3.54 Å, 3.67 Å, and 3.37 Å, respectively; an intramolecular π-σ interaction of 3.69 Å; two alkyl interactions with Arg199 and Leu200 of 4.42 Å and 4.44 Å, respectively; and a unfavorable positive-positive interaction with Arg199 of 4.11Å (Table S1).

Figure 1F shows the binding of Cr-2 (CID: 59908094) in the active site pocket residue Arg199 of chain A of AMPK1. The complex molecular docked pose shows interactions between them like two π-σ interactions through Tyr463 and Phe180 of 3.51 Å and 3.55 Å, respectively; single π-alkyl interaction with Tyr463 of 4.25 Å; and one alkyl interaction through Leu200 of 5.28 Å. The minimum potential binding energy between them is −6.1 kcal/mol (Table S1).

The potential binding energy or binding affinity value of a lead compound with AMPK1 extends between −7.4 and −6.1 kcal/mol whereas AMPK1 and metformin are −4.0 kcal/mol (Table 8). The results support the fact that the aforementioned compounds have lower free energy of interaction than the control drug and the biomolecules can become a potent AMPK1 stimulator.

3.4.2. AMPK2 targets interaction with leads and control compounds

Interaction between AMPK2 (chain A, pocket residues Glu264) and metformin is established and confirmed by docking pose in Figure 2A. Metformin is attached with AMPK2 residues Asp261, His247, and Thr243 through four usual H-bond interactions of 2.33 Å, 2.11 Å, 2.40 Å, and 2.40 Å; a carbon H-bond relations with residues Asp261 of 3.54Å; and two attractive charge interactions of 4.45 Å and 5.18 Å with Asp261 residues (Table S2). The Vina score of binding affinity between them is −4.2 kcal/mol (Table 9).

Vindoline (CID: 425978) interacts through chain A residue GLU264 of the AMPK2 and binding is confirmed by molecular docked pose (Figure 2B). The potential binding energy between vindoline and AMPK2 is −6.2 kcal/mol (Table 9). The lead is connected with AMPK2 residues with non-covalent interaction, such as Arg263 forms two H-bond of 2.21 Å and 2.29 Å; Lys260 and Arg263 residues form two carbon H-bond connections of 2.62 Å and 2.80 Å, respectively; residues Leu272 and Arg263 interact with two alkyl connections of 5.22 Å and 3.87 Å, respectively; residues Glu279 by an attractive charge interaction of 5.36 Å; and residues Asp280 with an unfavorable acceptor-acceptor interaction of 2.97 Å (Table S2).

The potential energy of binding (−5.6 kcal/mol) and conformer docked pose (Figure 2C) of vindolinine (CID: 24148538) and AMPK2 chain A, pocket residue Glu264, establishes the presence of stable molecular interaction between them. There is one carbon H-bond interaction with target residue Glu264 of 3.45 Å; two attractive charge interactions with Glu264 and Asp280 of 4.77 Å and 5.16 Å, respectively; a single π-anion interaction with Asp280 of 3.95Å; a single π-cation interaction with Lys269 of 3.97 Å; and one π-alkyl interaction with Arg263 between the target residues and the lead of 5.36 Å length (Table S2).

The lead compound (+)-vindorosine (CID:261578) and AMPK2 interaction docked pose (Figure 2D) has established that the compound positively enters and connected to the catalytic site Glu264 residue of chain A of AMPK2, with 3 conventional H-bond interaction around catalytic pocket residues Arg263 of 2.10Å, 2.41 Å, and 3.0 Å; one carbon H-bond interaction through Glu264 of 3.65 Å; two alkyl interactions through Lys260 of 4.17 Å and 3.77 Å; single attractive charge interaction with Glu264 of 4.62 Å; and a single π-anion bond contact with Glu279 of 3.86 Å. The minimum potential energy for binding between (+)-vindorosine and AMPK2 is −5.6 kcal/mol (Table S2).

The docked pose at the molecular level (Figure 2E) of AMPK2 and lead Cr-1 (CID: 5315746) has confirmed about definite interaction between them and the compound

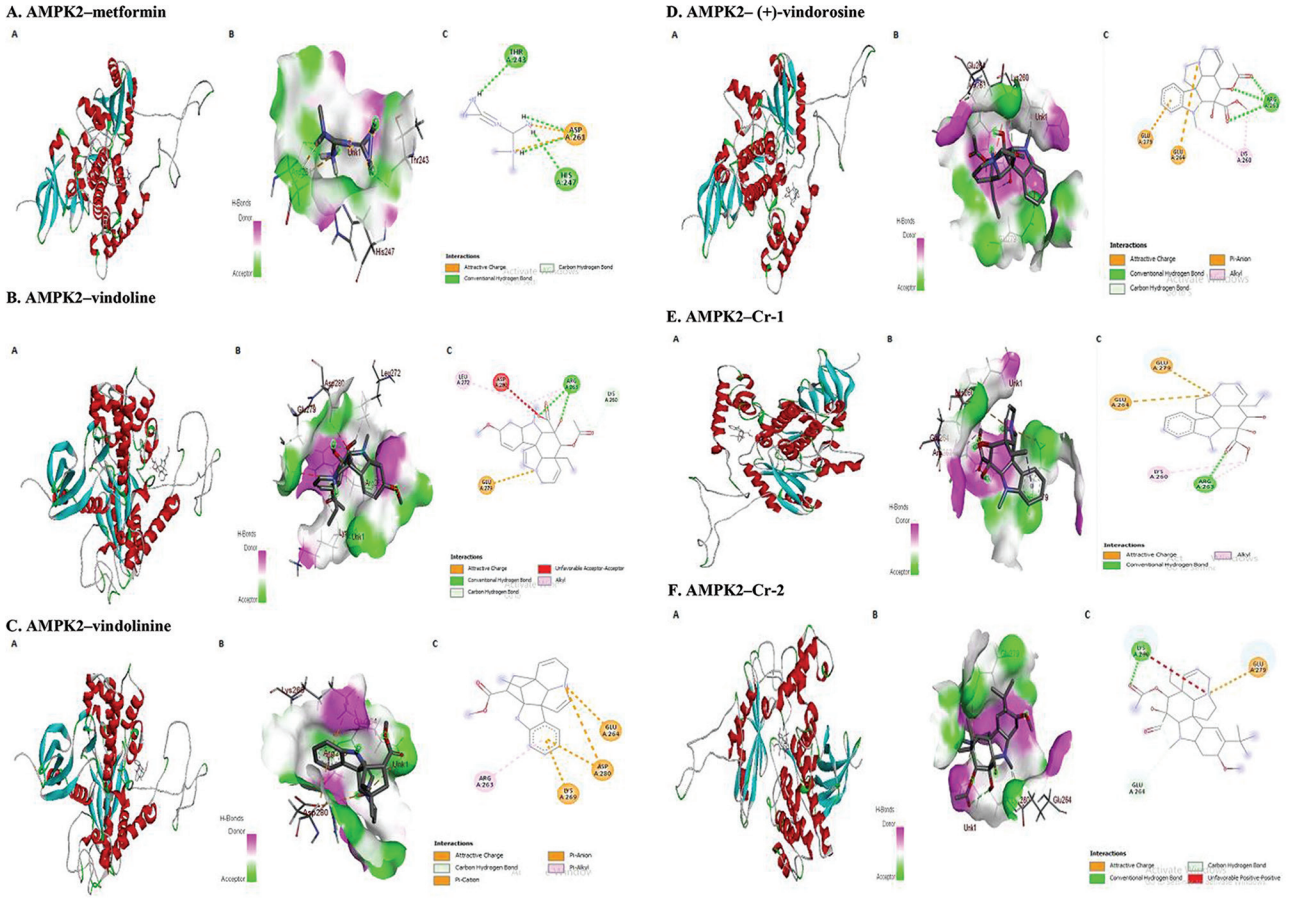


Figure 2. Docking studies of AMPK2 with different leads/ligands at the molecular level. (A) Interaction of AMPK2 with metformin (CID: 4091); (B) Interaction of AMPK2 with vindolinine (CID: 24148538); (C) Interaction of AMPK2 with vindoline (CID: 425978); (D) Interaction of AMPK2 with (+)-vindorosine (CID: 261578); (E) Interaction of AMPK2 with Cr-1(CID-5315746); and (F) interaction of AMPK2 and Cr-2 (CID: 59908094). In subfigures, Image A shows about specific and reversible interaction of drug target AMPK2 and ligand/lead compound; Image B provides the insight of depth of the catalytic site of the target, similar to a real situation and distance between ligand and target around them, and target-ligand interaction within the active site of AMPK2 target based on H-bond donor and acceptor characteristic of amino acid residues; and Image C representing the 2D image, which is signifying the linkage between ligand and target amino acids residues. On the 2D map, the color lines indicated the following: Conventional H-bond interaction (green); carbon-hydrogen bond (sky blue); attractive charge interaction, π-anion interaction, and π-cation interaction (orange); alkyl interaction and π-alkyl interaction (pink); unfavorable acceptor-acceptor interaction and unfavorable positive-positive interaction (red).

Table 9. The binding energy of docking interaction between AMPK2 target and ligands (control and test compounds)

Group	Docked compound	Binding energy (kcal/mol)
Control	AMPK2–metformin	−4.2
Test	AMPK2–vindoline	−6.2
	AMPK2–vindolinine	−5.6
	AMPK2–(+)-vindorosine	−5.6
	AMPK2–Cr1	−5.3
	AMPK2–Cr2	−5.3

positively entering and attaching to the catalytic site of the AMPK2 chain A, through two conventional H-bond interactions around catalytic pocket residues Arg263 of 2.63 Å and 2.32 Å; two alkyl interaction with Arg263 and Lys260 of 4.13 Å and 4.21 Å, respectively; two attractive charge interaction with residues Glu264 and Glu279 of 4.92 Å and 4.78 Å, respectively; and one intra molecular π-σ interaction of 3.73 Å. The potential binding affinity energy between them is −5.3 kcal/mol (Table S2).

The binding affinity conformer (Figure 2F) of the AMPK2 and lead Cr-2 (CID: 59908094) docking pose shows that the lead attached to the active location of chain A of AMPK2, compartment residues Glu264, with one intramolecular H-bond of 3.01 Å; conventional H-bond connection with Lys260 of 2.30 Å; two carbon H-bond links through Glu264 and Lys260 of 3.73 Å and 3.71 Å, respectively; single attractive charge interaction with

Glu279 of 3.45 Å; single unfavorable positive-positive contact with Lys260 of 3.81 Å. The binding energy between them is −5.3 kcal/mol (Table S2).

Comparison between the potential energy of metformin along with the biomolecules of *C. roseus* on the target, we found that the tested compounds' potential energy varies from −6.2 to −5.3 kcal/mol and has a lower potential energy than the control (−4.2 kcal/mol) (Table 9). Hence, it is predicted that the tested *C. roseus* alkaloids can bind strongly with the target and are more potent activators of AMPK2 than metformin.

4. Discussion

Diabetes mellitus initially promotes hyperglycemia but in chronic cases, it moves toward a multi-factorial disease situation. The current investigation supports that *C. roseus* alkaloids are powerful stimulators of AMPK and have antihyperglycemic properties[51]. The SOSUI prediction result shows that human AMPK1 (UniProt ID Q13131) is a soluble protein and made up of 559 AA, instead of 548 AA as reported in a previous paper[52], whereas AMPK2 (UniProt ID P54646) is a membrane-bound protein and made up of 552 AA (Table 1).

The ADMET test results (Table 7) have shown that putative drug candidates of *C. roseus* phytochemicals obey the Lipinski rule, that is, partition coefficient (cLogP) ≤ 5, which means the compounds have good hydrophilicity and can engross quickly in the cells. The biomolecules have a high opportunity of absorption in the oral and alimentary canal after administration. The standard parameter for the logarithmic assessment of aqueous solubility (LogS) is 0 to −4 at 25°C, pH = 7.5, and our result for the majority of ligands candidate is ≤ −4 mol/lit which means that the biomolecules have better absorption and distribution features. The molecular weight ≤ 500 for the ligands represents their small sizes leading to a high opportunity of arriving at the site of action and a high possibility of activity on the target site. The standard criteria for topological polar surface area (TPSA) are ≤140 Å whereas the result shows TPSA ≤100 Å, indicating a molecule has a low polar surface and the ability to infuse cells with high surface region for action, thus, they have a high transport properties or membrane permeability. Drug-likeness value varies from 0.58 to 3.95; the positive and higher value indicates molecules contain the majority of fragments having biological action or drug-likeness features. The drug-score result varies from 0.49 to 0.84, which means that ligands have the potential to meet the criteria as a drug. The blood-brain barrier (BBB) crossing ability was tested by DruLiTo software, which shows that vindolinine (CID: 24148538) and Cr-1 (CID: 5315746) had passed the test, whereas the others such as vindoline (CID: 425978), (+)-vindorosine (CID: 261578), and Cr-2 (CID: 59908094) were unable to be detected, means they may cross the BBB[46,47].

In the journey of *in silico* docking analysis, we found that metformin and the *C. roseus* alkaloids bind with chain A of AMPK1 pocket residue Arg199 and with chain A of AMPK2 pocket residue Glu264, other than substrate binding groves of AMPK (Thr172). Hence, metformin and the biomolecules act as modulators and indirect AMPK activators, which activate AMPK by enhancing the AMP: ATP ratio by inhibiting complex I of the respiratory chain of mitochondria and by intracellular accumulation of calcium[19,53,54].

The docking result of the present study indicates that vindolinine has a strong binding affinity with target AMPK1 and AMPK2 with binding energy values of −7.4 kcal/mol and −5.6 kcal/mol, respectively. Similarly, the binding energy of vindoline with target AMPK1 and AMPK2 is −6.3 kcal/mol and −6.2 kcal/mol, respectively; (+)-vindorosine is −6.3 kcal/mol and −5.6 kcal/mol, respectively; Cr-1 is −6.4 kcal/mol and −5.3 kcal/mol, respectively; and Cr-2 is −6.1 kcal/mol and −5.3 kcal/mol, respectively (Tables 8 and 9). The binding energy of the metformin is −4.0 kcal/mol with AMPK1 and −4.2 kcal/mol with AMPK2. Based on the outcome, it is concluded that the *C. roseus* alkaloids vindoline (CID: 425978), vindolinine (CID: 24148538), (+)-vindorosine (CID: 261578), Cr-1 (CID: 5315746), and Cr-2 (CID: 59908094) stimulate AMPK1 and AMPK2 more effectively than metformin. Numerous researches and literature found that the *C. roseus* plant has a hypoglycemic effect and is superior to biguanides (e.g., metformin). Furthermore, its extract decreases the risk of diarrhea-related toxicity (a side effect of metformin)[21].

As per the thermodynamic principle, "the lesser the Gibbs free energy or potential energy, the higher the complex stability". The descending order of binding affinity of AMPK1 with different ligands is vindolinine (−7.4 kcal/mol) > Cr-1 (−6.4 kcal/mol) > vindoline (−6.3 kcal/mol) = (+)-vindorosine (−6.3 kcal/mol) > Cr-2 (−6.1 kcal/mol) > metformin (−4.0 kcal/mol). Whereas the descending order of binding affinity of AMPK2 with different ligands is vindoline (-6.2 kcal/mol) > vindolinine (-5.6 kcal/mol) = (+)-vindorosine (-5.6 kcal/mol) > Cr-1 (-5.3 kcal/mol) = Cr-2 (-5.3 kcal/mol) > metformin (-4.2 kcal/mol). The general binding affinity of biomolecules with AMPK1 appeared to be greater than with AMPK2, which is likely due to the globular and soluble nature of AMPK1. Among all ligands, vindolinine (−7.4 kcal/mol) is the strongest stimulator of AMPK1 whereas vindoline (−6.2 kcal/mol) is of AMPK2.

The difference in potential energy of the AMPK-ligand complex may be due to the distance and disparity in the attractive and weaker interactions such as H-bonding, electrostatic forces, hydrophobic interaction, or unfavorable bonding. The result provides evidence that AMPK1 and AMPK2 complexes have identical specific activity but differences in binding affinity of ligands and also have differential substrate preference. All the qualified ligands/leads have correlated substrate specificity with both targets because the alpha-1 and alpha-2 subunits of AMPK have well-built homology (77% resemblance in amino acids) around the substrate binding groves (Thr172) and contribute similarly to the total AMPK action[10]. Dissimilarity in substrate liking based on binding energy value may be due to a difference in the degree of phosphorylation[52]. Diversity in the mechanism of action of both subunits is due to differences in downstream targets of the AMPK cascade[10,52]. AMPK1 and AMPK2 enzymes are involved in the direct phosphorylation of metabolic enzymes and transcription regulation and have distinct biological functions[17]; therefore, AMPK is known to be an important first-line target for the treatment of Type II diabetes and other diseases[55].

AMPK1 is widely present and expressed, regulates intracellular signaling through phosphorylation, and accounts for more than 90% of the AMPK activity[10,11]. The current study result shows that the *C. roseus* alkaloids are excellent indirect AMPK1 stimulators and activate the target by the release of sarcoplasmic calcium ion and disturbs energy balance which leads to increased AMP: ATP ratio[56]. It is known that AMPK1 regulates enzymes of glucose metabolism[16], and insulin release[13], but is not able to generate a sufficient signal for glucose uptake. However, in the case of low-intensity twitch contraction, it promotes glucose uptake up to 60% by glucose transporter 1 (GLUT1) and has a limited role in glucose metabolism[56]. Moreover, AMPK1 plays a major role in safeguarding cells from stresses that bring out ATP depletion through turning off ATP-consuming biosynthetic pathways[12]. In addition, it engages in the mammalian target of rapamycin (mTOR) signaling pathway[15], cytokine secretion, macrophage activation through regulation of arginine metabolism key enzymes, prostaglandin synthesis, immune suppression, differentiation of MDSC[14,16], prohibiting skeletal muscle hypertrophy, etc.[15,57]. The metabolic pathway was highlighted in Figure 3A based on the literature. According to the computational docking studies resulted in high binding affinity between the target and ligands in this study, it is predicted that by activating AMPK1 the biomolecules (vindoline, vindolinine, (+)-vindorosine, Cr-1, and Cr-2) can protect beta cells of pancreas and control diseases related to cellular stress conditions, cancer, and cardiovascular problems through downward cascading molecules and related mechanism. Therefore, its nature of action can be validated further by wet laboratory experiments and clinical trials. Moreover, some literature shows that vindoline and α/β-tubulin (PDB: 1Z2B) have a binding affinity of −7.28 kcal/mol. Therefore, vindoline can improve the protection mechanism of plants and have inhibitory outcomes on cancer cells. Alkaloids have antioxidant properties and can decrease glucotoxicity and reactive oxygen species (ROS) generation by quenching ROS and suppressing oxidative phosphorylation in stress conditions by exciting AMPK1[58].

AMPK2 is greatly expressed in cardiac muscle, skeletal muscles, and liver[11], and contributes < 10% activity of AMPK[10]. AMPK2 enzyme regulates metabolic health by cellular energy homeostasis, promoting glucose and fatty acid uptake, improving insulin sensitivity, reducing insulin resistance, recycling insulin receptors (INSR), and regulating GLUT4 transport[17]. Also, activated AMPK2 phosphorylates glycolysis pathway machinery 6-phosphofructokinase-2 (PFKB2), and fructose-2,6-biphosphatase 3 (PFKB3). In the liver, it stimulates glucose homeostasis by phosphorylating CREB-regulated transcription coactivator 2 (CRTC2)/target of rapamycin 2 (TORC2)[17]. Thus, AMPK2 activation by biomolecules reduces hepatic glucose production, increases glucose utilization in hepatocytes, and amplifies insulin-mediated peripheral glucose absorption by GLUT4 in skeletal muscle and fat cells and as a result, decreases blood sugar concentration[1,17,19-21,54]. Based on the literature, the metabolic pathway of AMPK2 was highlighted in Figure 3B and its nature of action can be confirmed further by wet laboratory experiments and clinical trials.

C. roseus ethanolic extract appreciably curtailed the elevated blood sugar level in addition to uplifting significantly antioxidant security systems and helping in the revival of beta cells[58]. An *in silico* study of alkaloid vindoline showed that it controls hyperglycemia and other glucotoxicity-induced pathogenesis by acting on multiple targets (AMPK, peroxisome proliferator-activated receptor γ [PPARγ], dipeptidyl peptidase 4 [DPP4], and α-glucosidase)[59]. Vindoline was examined on Wistar rats and the outcome showed that vindoline reduces significantly fasting blood glucose level (FBG) ($P < 0.05$) and lessens hyperglycemia-generated liver, kidney damage, and improves the *in vitro* insulin discharge in pancreatic tissues[60].

Chronic AMPK2 stimulation reduces or inhibits ACC activity and HMG-CoA reductase[61], induces fatty acid oxidation, and suppresses expression of lipogenic enzymes, leading to a decrease of membrane cholesterol (~10%) and triglyceride (~45%) levels[62], restoring F-actin

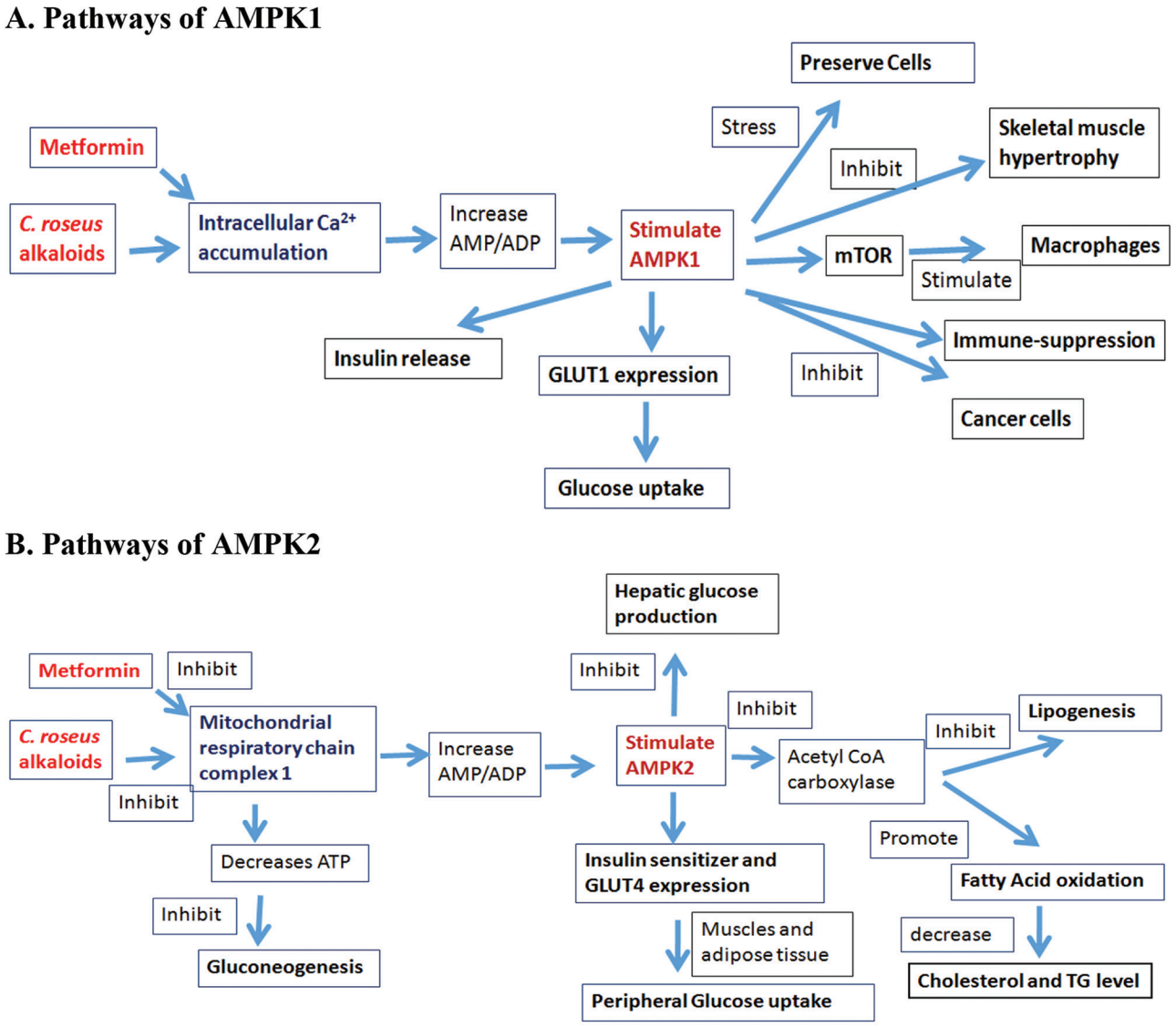


Figure 3. Schematic representation of activation pathways and action of target AMPK1 and AMPK2. (A) The projected mechanism of action of *C. roseus* alkaloids and metformin on activation AMPK1 and their effect on prevention of different diseases; (B) The proposed mechanism of action of the tested alkaloids and metformin on downregulation of hyperglycemia and hypercholesterolemia by activation of AMPK2.

structure[63,64], inhibiting cell proliferation through p53 and mTOR, inhibiting macrophage proliferation[16], and preventing glucotoxicity by decreasing NADPH oxidase 2 (NOX2)-mediated ROS production. In addition, plant alkaloid treatment improves the organization of the nephritic parenchyma and considerably reduces caspase 9 expression and peroxidation of lipids in diabetic ones[65]. AMPK2 stimulation can also protect against diabetic cardiomyopathy and cardiovascular disease[61].

Many literatures describe that the aqueous extract of *C. roseus* considerably reduces the blood glucose level and glycogen content of the liver and muscle[31]. Alkaloids such as vindoline, vindolinine, catharanthine, etc. of *C. roseus* extract have blood glucose lowering action, secretion boost of insulin from the β-cell of Langerhans, and improvement of enzymatic activities of glycogen synthase, glucose-6-phosphate-dehydrogenase, succinate dehydrogenase, and malate dehydrogenase (an extrapancreatic mechanism) reported in diabetic rats[28,66]. Test of the alkaloids on pancreatic β-TC6 or myoblast C2C12 cell lines reports that vindolidine, vindolicine, vindoline, and vindolinine alkaloids were elevating glucose intake in beta and muscle cells through hampering tyrosine phosphatase-1B (PTP-1B), and also promotes sensitivity of insulin and antioxidant potentials[44]. The plant alkaloids act on phosphatidylinositol 3,4,5-triphosphate [Ptdlns(3,4,5)P3 or

PIP[3], inactivate phosphor glucocorticoid-regulated kinase 3 (pSGK3), and activate glycogen synthesis in the liver[27]. Herbal alkaloids could stimulate insulin secretion from beta cells, reduce insulin resistance, stimulate glycogenesis and hepatic glycolysis, activate PPARγ, etc.[67]. There is a need for clinical tests and validate the effect of the biomolecules as AMPK2 stimulators and their role in heart diseases, metabolic syndrome, obesity, diabetes, and other diseases.

5. Conclusion

Docking studies results between *C. roseus* phytochemicals and targets AMPK1 and AMPK2 highlighting strong bonding affinity between ligands and targets and bring into focus several questions related to diverse medicinal roles of alkaloids vindolinine (CID: 24148538), vindoline (CID: 425978), (+)-vindorosine (CID: 261578), Cr-1 (CID: 5315746), and Cr-2 (CID: 59908094). AMPK2 is a major regulator of glucose metabolism along with lipid metabolism leading to an effective control of hyperglycemia, hypercholesterolemia, and obesity. The major role of AMPK1 is in cellular stress, macrophages, and cytokine-based immunity regulation pathways. The present study evidence supports the findings of alkaloids having a greater potential to trigger AMPK1/2 targets and related pathways better than metformin and may effectively control hyperglycemia and diabetic cardiomyopathy, metabolic syndrome, obesity, etc. The putative drug candidates of *C. roseus* alkaloids can be further tested experimentally and validated clinically. These findings provide the basis for the development of drug candidates that have the ability to target multiple diseases, including diabetes and its related secondary complications.

Acknowledgments

Richa Goyal is thankful to Dr. Ashish Kumar Biswas (HOD and Associate Professor, Pharmacology Department, Sheikh Bhikhari Medical College and Hospital Hazaribagh, Jharkhand, India) and Dr. Mukesh Nitin (Head of Department of Tech. Data Sci. & Biosciences, Digianalix, Ranchi, Jharkhand, India) for guidance and feedback.

Funding

None.

Conflict of interest

All the authors declared no conflicts of interest.

Author contributions

Conceptualization: Richa Goyal
Investigation: Richa Goyal
Writing – original draft: Richa Goyal
Richa Goyal *Writing – reviewing and editing*: Manoj Kumar, M. Anwar Mallick

Ethics approval and consent to participate

The current research article work was performed by computational tools, without using any scientific lab animals, human samples, or any human patients. Thus, no official permission was taken from our research team against the research article from the ethical committee of our university and government.

Consent for publication

Not applicable.

Availability of data

Not applicable.

References

1. Tripathi KD, 2010, Essentials of Medical Pharmacology. 6th ed. India: Jaypee Publication.
2. Hall JE, Guyton AC, 2011, Guyton and Hall Textbook of Medical Physiology. 12th ed. USA: Saunders Elsevier Publication.
3. Kasper DL, Fauci AS, Longo DL, *et al.*, 2015, Harrison's Principles of Internal Medicine. 19th ed. USA: McGraw Hill Publication.
4. Saeedi P, Petersohn I, Salpea P, *et al.*, 2019, Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract*, 157: 107843.
   https://doi.org/10.1016/j.diabres.2019.107843
5. Singh U, 2016, Prevalence of diabetes and other health related problems across India and worldwide: An overview. *J Appl Nat Sci*, 8(1): 500–505.
   https://doi.org/10.31018/jans.v8i1.825
6. Proença C, Ribeiro D, Freitas M, *et al.*, 2021, Flavonoids as potential agents in the management of Type 2 diabetes through the modulation of α-amylase and α-glucosidase activity: A review. *Crit Rev Food Sci Nutr*, 62: 3137–3207.
   https://doi.org/10.1080/10408398.2020.1862755
7. Kashtoh H, Baek KH, 2022, Recent updates on phytoconstituent alpha-glucosidase inhibitors: An approach towards the treatment of type two diabetes. *Plants* (*Basel*), 11: 2722.
   https://doi.org/10.3390/plants11202722
8. Lebovitz HE, 1997, Alpha-glucosidase inhibitors. *Endocrinol Metab Clin North Am*, 26(3): 539–551.
   https://doi.org/10.1016/s0889-8529(05)70266-8

9. Alam S, Sarker MM, Sultana TN, *et al*., 2022, Antidiabetic phytochemicals from medicinal plants: Prospective candidates for new drug discovery and development. *Front Endocrinol* (*Lausanne*), 13: 800714.

https://doi.org/10.3389/fendo.2022.800714

10. Woods A, Salt I, Scott J, *et al*., 1996, The alpha1 and alpha2 isoforms of the AMP-activated protein kinase have similar activities in rat liver but exhibit differences in substrate specificity *in vitro*. *FEBS Lett*, 397: 347–351.

https://doi.org/10.1016/s0014-5793(96)01209-4

11. Wang YG, Han XG, Yang Y, *et al*., 2016, Functional differences between AMPK α1 and α2 subunits in osteogenesis, osteoblast-associated induction of osteoclastogenesis, and adipogenesis. *Sci Rep*, 6: 32771.

https://doi.org/10.1038/srep32771

12. PRKAA1 Protein Kinase AMPK-activated Catalytic Subunit Alpha 1 [*Homo sapiens* (Human)]. Available from: https://www.ncbi.nlm.nih.gov [Last accessed on 2023 Apr 15].

13. Da Silva Xavier G, Leclerc I, Varadi A, *et al*., 2003, Role for AMP-activated protein kinase in glucose-stimulated insulin secretion and preproinsulin gene expression. *Biochem J*, 371(3): 761–774.

https://doi.org/10.1042/bj20021812

14. Trillo-Tinoco J, Sierra RA, Mohamed E, *et al*., 2019, AMPK alpha-1 intrinsically regulates the function and differentiation of tumor myeloid-derived suppressor cells. *Cancer Res*, 79(19): 5034–5047.

https://doi.org/10.1158/0008-5472.CAN-19-0880

15. Mounier R, Lantier L, Viollet B, *et al*., 2009, Important role for AMPKalpha1 in limiting skeletal muscle cell hypertrophy. *FASEB J*, 23(7): 2264–2273.

https://doi.org/10.1096/fj.08-119057

16. Phair IR, Nisr RB, Viollet B, *et al*., 2023, AMPK integrates metabolite and kinase-based immunometabolic control in macrophages. *Mol Metab*, 68: 101661.

https://doi.org/10.1016/j.molmet.2022.101661

17. UniProt Consortium, 2021, UniProt: The universal protein knowledgebase in 2021. *Nucl Acids Res*, 49: D480–D489.

https://doi.org/10.1093/nar/gkaa1100

18. PRKAA2 Protein Kinase AMP-activated Catalytic Subunit Alpha2 [*Homo sapiens* (Human)]. Available from: https://www.ncbi.nlm.nih.gov [Last accessed on 2023 Apr 15].

19. Foretz M, Guigas B, Bertrand L, *et al*., 2014, Metformin: From mechanisms of action to therapies. *Cell Metab*, 20(2): 953–966.

https://doi.org/10.1016/j.cmet.2014.09.018

20. Foretz M, Guigas B, Viollet B, 2019, Understanding the glucoregulatory mechanisms of metformin in Type 2 diabetes mellitus. *Nat Rev Endocrinol*, 15(10): 569–589.

https://doi.org/10.1038/s41574-019-0242-2

21. Ohadoma SC, Akpan JL, Odey PA, *et al*., 2021, Mechanistic considerations of *Catharanthus roseus* on the hypoglycemic activity of alpha glucosidase inhibitors and biguanides: A review. *J Pharm Adv Res*, 4(10): 1390–1398.

22. Francini F, Schinella GR, Ríos JL, 2019, Activation of AMPK by medicinal plants and natural products: Its role in Type 2 diabetes mellitus. *Mini Rev Med Chem*, 19(11): 880–901.

https://doi.org/10.2174/1389557519666181128120726

23. Srivastava V, Yadav A, Sarkar P, 2022, Molecular docking and ADMET study of bioactive compounds of *Glycyrrhiza glabra* against main protease of SARS-CoV2. *Mater Today Proc*, 49: 2999–3007.

https://doi.org/10.1016/j.matpr.2020.10.055

24. Adhikari B, 2021, Roles of alkaloids from medicinal plants in the management of diabetes mellitus. *J Chem*, 10: 2691525.

https://doi.org/10.1155/2021/2691525

25. Bennouna J, Delord JP, Campone M, *et al*., 2008, Vinflunine: A new microtubule inhibitor agent. *Clin Cancer Res*, 14: 1625–1632.

https://doi.org/10.1158/1078-0432.CCR-07-2219

26. Newman DJ, Cragg GM, 2007, Natural products as sources of new drugs over the last 25 years. *J Nat Prod*, 70(3): 461–477.

https://doi.org/10.1021/np068054v

27. Balaji H, 2014, Versatile therapeutic effects of *Vinca rosea* linn. *Int J Pharm Sci Health Care*, 1(4): 56–76.

28. Mohan SC, Anand T, Priyadharshini GS, *et al*., 2015, GC-MS analysis of phytochemicals and hypoglycemic effect of *Catharanthus roseus* in alloxan-induced diabetic rats. *Int J Pharm Sci Rev Res*, 31(1): 123–128.

29. Tolambiya P, Mathur S, 2016, A study on potential phytopharmaceuticals assets in *Catharanthus roseus* L. (Alba). *Int J Life Sci Biotechnol Pharm Res*, 5: 1–6.

https://doi.org/10.18178/ijlbpr.5.1.1-6

30. Nisar A, Mamat AS, Hatim MI, *et al*., 2016, An updated review on *Catharanthus roseus*: Phytochemical and pharmacological analysis. *Indian Res J Pharm Sci*, 9: 631–653.

31. John J, 2017, Evaluation of hypoglycemic effect of *Aloe vera* on allaxon induced diabetic rats. *Int J Inf Res Rev*, 4(3): 3865–3868.

32. Kalhotra P, Chittepu CS, Osorio-Revilla G, *et al*., 2020, Phytochemicals in garlic extract inhibit therapeutic enzyme DPP-4 and induce skeletal muscle cell proliferation: A possible mechanism of action to benefit the treatment of

diabetes mellitus. *Biomolecules*, 10: 305.

https://doi.org/10.3390/biom10020305

33. Bharadwaj S, Dubey A, Yadava U, *et al.*, 2021, Exploration of natural compounds with anti-SARS-CoV-2 activity via inhibition of SARS-CoV-2 Mpro. *Brief Bioinform*, 22(2): 1361–1377.

https://doi.org/10.1093/bib/bbaa382

34. Yousif E, Sherif R, Abeer AE, *et al.*, 2021, Vinca (*Catharanthus roseus*) extracts attenuate alloxan-induced hyperglycemia and oxidative stress in journal rats. *Am Food Sci Technol*, 9(4): 161–172.

https://doi.org/10.12691/ajfst-9-4-8

35. Schwede T, Kopp J, Guex N, *et al.*, 2003, SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res*, 31(13): 3381–3385.

https://doi.org/10.1093/nar/gkg520

36. Gasteiger E, Hoogland C, Gattiker A, *et al.*, 2005, Protein identification and analysis tools on the expasy server. In: Walker JM, editor. The Proteomics Protocols Handbook. United States: Humana Press. p571-607.

https://doi.org/10.1385/1-59259-890-0:571

37. Geourjon C, Deléage G, 1995, SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci*, 11(6): 681–684.

https://doi.org/10.1093/bioinformatics/11.6.681

38. Hirokawa T, Boon-Chieng S, Mitaku S, 1998, SOSUI: Classification and secondary structure prediction system for membrane proteins. *Bioinformatics*, 14(4): 378–379.

https://doi.org/10.1093/bioinformatics/14.4.378

39. Hollingsworth SA, Karplus PA, 2010, A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomol Concepts*, 1(3–4): 271–283.

https://doi.org/10.1515/BMC.2010.022

40. Binkowski TA, Naghibzadeh S, Liang J, 2003, CASTp: Computed atlas of surface topography of proteins. *Nucleic Acids Res*, 31(13): 3352–3335.

https://doi.org/10.1093/nar/gkg512

41. Guex N, Peitsch MC, 1997, SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis, 18: 2714–2723.

https://doi.org/10.1002/elps.1150181505

42. Pettersen EF, Goddard TD, Huang CC, *et al.*, 2004, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, 25(13): 1605–1612.

https://doi.org/10.1002/jcc.20084

43. Eberhardt J, Santos-Martins D, Tillack AF, *et al.*, 2021, AutoDock Vina 1.2.0: New docking methods, expanded force field, and python bindings. *J Chem Inf Model*, 61: 3891–3898.

https://doi.org/10.1021/acs.jcim.1c00203

44. Tiong SH, Looi CY, Hazni H, *et al.*, 2013, Antidiabetic and antioxidant properties of alkaloids from *Catharanthus roseus* (L.) G. Don. *Molecules*, 18: 9770–9784.

https://doi.org/10.3390/molecules18089770

45. O'Boyle NM, Banck M, James CA, *et al.*, 2011, Open babel: An open chemical toolbox. *J Cheminform*, 3: 33.

https://doi.org/10.1186/1758-2946-3-33

46. Sander T, 2001, OSIRIS Property Explorer. Switzerland: Organic Chemistry Portal. Available from: https://www.organic-chemistry.org/prog/peo [Last accessed on 2022 Dec 25].

47. Bickerton GR, Paolini GV, Besnard J, *et al.*, 2012, Quantifying the chemical beauty of drugs. *Nat Chem*, 4: 90–98.

https://doi.org/10.1038/nchem.1243

48. Hanwell MD, Curtis DE, Lonie DC, *et al.*, 2012, Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. J Cheminform, 4: 17.

https://doi.org/10.1186/1758-2946-4-17

49. BIOVIA, Dassault Systems, 2021, BIOVIA Discovery Studio Visualizer 4.5, 21.1. San Diego: Dassault System. Available from: https://biovia-discovery-studio-2021-client.software.informer.com [Last accessed on 2023 Jan 02].

50. Kim S, Chen J, Cheng T, *et al.*, 2021, PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res*, 49(D1): D1388–D1395.

https://doi.org/10.1093/nar/gkaa971

51. Sharma B, Mittal A, Dabur R, 2018, Mechanistic approach of anti-diabetic compounds identified from natural sources. *Chem Biol Lett*, 5(2): 63–99.

52. Stapleton D, Mitchelhill KI, Gao G, *et al.*, 1996, Mammalian AMP-activated protein kinase subfamily. *J Biol Chem*, 271(2): 611–614.

https://doi.org/10.1074/jbc.271.2.611

53. Kim J, Yang G, Kim Y, *et al.*, 2016, AMPK activators: Mechanisms of action and physiological activities. *Exp Mol Med*, 48: e224.

https://doi.org/10.1038/emm.2016.16

54. Tarasiuk O, Miceli M, Di Domizio A, *et al.*, 2022, AMPK and diseases: State of the art regulation by AMPK-targeting molecules. *Biology* (*Basel*), 11: 1041.

https://doi.org/10.3390/biology11071041

55. Coughlan KA, Valentine RJ, Ruderman NB, *et al.*, 2014, AMPK activation: A therapeutic target for Type 2 diabetes? *Diabetes Metab Syndr Obes*, 7: 241–253.

https://doi.org/10.2147/DMSO.S43731

56. Jensen TE, Schjerling P, Viollet B, *et al*., 2008, AMPK α1 activation is required for stimulation of glucose uptake by twitch contraction, but not by H_2O_2, in mouse skeletal muscle. *PLoS One*, 3(5): e2102.

https://doi.org/10.1371/journal.pone.0002102

57. Sertel S, Fu Y, Zu Y, *et al*., 2011, Molecular docking and pharmacogenomics of *Vinca* alkaloids and their monomeric precursors, vindoline and catharanthine. *Biochem Pharmacol*, 81(6): 723–735.

https://doi.org/10.1016/j.bcp.2010.12.026

58. Goboza M, Meyer M, Aboua YG, *et al*., 2020, *In vitro* antidiabetic and antioxidant effects of different extracts of *Catharanthus roseus* and its indole alkaloid, vindoline. *Molecules*, 25: 5546.

https://doi.org/10.3390/molecules25235546

59. Goyal R, Kumar M, Mallick MA, 2023, Drug Innovation Studies Targeting Diabetes: A Computational Docking Approach on Muti-drug Targets Including COVID Inhibitors. North Carolina: Research Square.

https://doi.org/10.21203/rs.3.rs-2457415/v1

60. Goboza M, Aboua YG, Chegou N, *et al*., 2019, Vindoline effectively ameliorated diabetes-induced hepatotoxicity by docking oxidative stress, inflammation and hypertriglyceridemia in Type 2 diabetes-induced male Wistar rats. *Biomed Pharmacother*, 112: 108638.

https://doi.org/10.1016/j.biopha.2019.108638

61. Heidary MR, Samimi Z, Asgary S, *et al*., 2022, Natural AMPK activators in cardiovascular disease prevention. *Front Pharmacol*, 12: 738420.

https://doi.org/10.3389/fphar.2021.738420

62. Foretz M, Even PC, Viollet B, 2018, AMPK activation reduces hepatic lipid content by increasing fat oxidation *in vivo*. *Int J Mol Sci*, 19: 2826.

https://doi.org/10.3390/ijms19092826

63. Zhou G, Myers R, Li Y, *et al*., 2001, Role of AMP-activated protein kinase in mechanism of metformin action. *J Clin Invest*, 108: 1167–1174.

https://doi.org/10.1172/JCI13505

64. Habegger KM, Hoffman NJ, Ridenour CM, *et al*., 2012, AMPK enhances insulin-stimulated GLUT4 regulation via lowering membrane cholesterol. *Endocrinology*, 153(5): 2130–2141.

https://doi.org/10.1210/en.2011-2099

65. Oguntibeju OO, Aboua Y, Goboza M, 2019, Vindoline-a natural product from *Catharanthus roseus* reduces hyperlipidemia and renal pathophysiology in experimental Type 2 diabetes. *Biomedicines*, 7: 59.

https://doi.org/10.3390/biomedicines7030059

66. Mishra JN, Verma NK, 2017, A brief study on *Catharanthus roseus*: A review. *Int J Res Pharm Pharm Sci*, 2(2): 20–23.

67. Sujatha S, Alhaji NM, 2022, Mechanism of action of alkaloids in the management of diabetes mellitus. *J Adv Sci Res*, 13(3): 4–7.

https://doi.org/10.55218/JASR.202213302

ACCSCIENCE PUBLISHING

Gene & Protein in Disease

ORIGINAL RESEARCH ARTICLE

In silico mutation analysis of the SARS-CoV-2 Spike glycoprotein in the Omicron (B.1.1.529) variant isolated from the Iraqi patients

Dana Khdr Sabir*

Charmo Center for Research, Training, and Consultancy, Charmo University, Chamchamal, Sulaymaniyah, Kurdistan Region, Iraq

Abstract

Since its first breakout in December 2019, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has impacted the lives of millions of people worldwide. The virus has been rapidly mutating, and the accumulation of various mutations has precipitated the emergence of several new variants. The Omicron variant (B.1.1.529 lineage) was first identified in Botswana and South Africa back in November 2021. Since then, several Omicron sub-lineages have emerged as a result of hypermutations. In this study, a computational analysis of the 381 spike glycoprotein (S protein) of the SARS-CoV-2 Omicron variants isolated from Iraqi patients was performed. The full-length S protein sequences (1273 amino acids) were obtained from the publicly accessible Global Initiative on Sharing All Influenza Data database. A total of 60 mutation sites were recognized: 49 substitution sites, ten deletions, and one insertion. K417N and N440K were the most prevalent mutations (n = 379, 99.4%), followed by G339D (n = 377, 98.9%) and S373P and S375F (both n = 367, 96.3%). Both BA.1.1 (n = 198, 52%) and BA.1 (n = 91, 14%) were the predominant variant types encountered throughout this study. The current work offers the data of SARS-CoV-2 Omicron variants derived from the Iraqi patients. The data from this study could assist in the molecular design of more potent vaccines and/or antiviral drugs against the virus and also provide a fundamental understanding of SARS-CoV-2 evolution with concerns about their pathogenicity.

Keywords: Severe acute respiratory syndrome coronavirus 2; Omicron; Iraq; Mutations; Variant; Spike glycoprotein

***Corresponding author:**
Dana Khdr Sabir
(dana.sabir@charmouniversity.org)

Citation: Sabir DK, 2023, *In silico* mutation analysis of the SARS-CoV-2 Spike glycoprotein in the Omicron (B.1.1.529) variant isolated from the Iraqi patients. *Gene Protein Dis,* 2(3): 1646. https://doi.org/10.36922/gpd.1646

Received: August 21, 2023
Accepted: September 20, 2023
Published Online: September 29, 2023

1. Introduction

The coronavirus disease (COVID-19) is a highly contagious infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which was first reported in late December 2019 in China[1,2]. The disease has impacted almost all sectors of life[3,4], infecting more than 750 million cases and causing 7 million deaths worldwide[5]. In Iraq alone, 2.5 million COVID-19 cases and more than 25000 COVID-19-related deaths have been reported[6,7].

SARS-CoV-2 virus belongs to the beta-coronavirus genus, consisting of non-segmented, positive-sense, single-stranded RNA[8-10]. This virus has a genome size of ~30

kbp that encodes for a number of accessory proteins and four major structural proteins, such as spike (S) protein, nucleocapsid (N) protein, membrane protein (M), and envelope (E) protein[11-13]. Comprising 1273 amino acids, S protein is a trimeric protein having several domains that are known to facilitate the entry of the virus into the host cells through the attachment and fusion into the angiotensin-converting enzyme 2 (ACE2) receptor of the host cells[14].

On completing the genetic sequencing of the virus for the first time, several variants of the virus with varying levels of pathogenicity and transmissibility were reported around the world[15]. The Omicron variant (B.1.1.529 lineage) was first identified in Botswana and South Africa on November 24, 2021 and then classified as a variant of concern (VOC) on November 26, 2021[16,17]. This variant is characterized by high environmental durability, high transmissibility, firm binding to human ACE2 receptor, attenuated viral replication, and resistance against therapeutic antibodies produced as a result of vaccination[18,19]. This variant was primary accountable for the fourth wave of the COVID-19 crisis in several countries[20].

As of February 2023, approximately 15 million SARS-CoV-2 whole genome sequences were available on the Global Initiative on Sharing All Influenza Data (GISAID) database[21]. Numerous studies have been conducted to comprehend the evolution process of SARS-CoV-2 and to, mostly by means of genomic sequencing, classify the virus into corresponding clades that may possess characteristics phenotypes[6]. Given the rampage caused by the Omicron variant around the world and the irreversible health damage it engenders, it is of great importance to undertake in-depth research to understand its evolutionary pattern, mutations carried by the sub-variants, and clinical manifestations of the infectious disease. Therefore, this study aimed to investigate the different SARS-CoV-2 Omicron variants that appeared in Iraq and assess the impact of these variants on the viral transmissibility and virulence.

2. Materials and methods

2.1. Spike protein sequences

All the 381 S protein sequences of the Iraqi SARS-CoV-2 Omicron variants were extracted from GISAID databank on January 23, 2023. The accessory data included collection date, accession ID, amino acid substitutions, gender, age, city, Pango lineage, originating and submitting laboratory, and authors' name (see Raw Data File). The information was, then, wrangled and analyzed manually using Microsoft Excel.

2.2. Pango lineages and S protein mutation analysis

Variants and sub-variants of the SARS-CoV-2 Omicron were determined based on the original GISAID classification. Mutations of each S protein sequences were identified by manual checking of the recorded sequence data, and the sequences were aligned with the first sequence of SARS-CoV-2 isolate from Wuhan Hu-1/China (accession number: NC_045512.2).

2.3. Protein visualization and construction of the phylogenic trees

Data regarding the 3D structure of the S protein and its common mutation sites were obtained the ViralZone database (https://viralzone.expasy.org/)[22]. A phylogenic tree was constructed based on the 381 Omicron variants sequences using the Nextstrain tool (https://nextstrain.org/)[23].

3. Results

3.1. Pango lineage

The sequence analysis revealed that six SARS-CoV-2 Omicron variants circulated in Iraq within the period of this study (November 2021 – November 2022) (Figure 1).

Based on the mutation sequence analysis on the GISAID website, the most common lineage (52%) was BA.1.1, followed by unassigned Omicron lineage accounting for 24% of the total cases. The lineages of BA.1, BA.2, and BA.1.17.2 accounted for 14%, 3%, and 2%, respectively.

The percentage of the different Omicron variants that appeared during the period of the study was calculated. As shown in Figure 1, all the reported Omicron variants in November 2021 belonged to the BA.1 Pango lineage. Between December and February 2022, BA.1.1 was one of the prevalent assigned lineages. Despite accounting

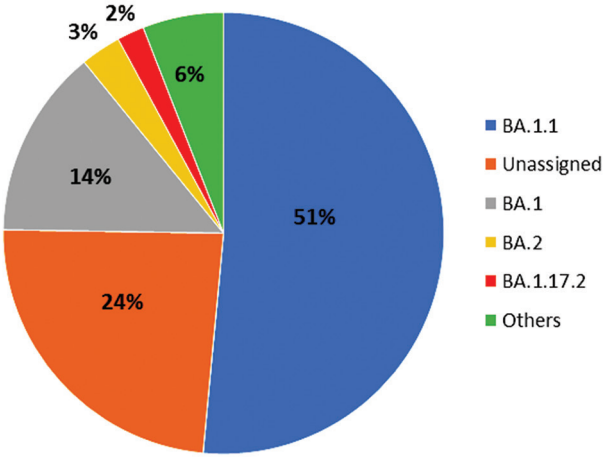


Figure 1. Percentage of the different SARS-CoV-2 Omicron variants circulated in Iraq from November 2021 to November 2022. The "others" refers to the Pango lineage of the Omicron variant which did not fit into any of the Pango lineages based on the computational analysis of the mutations.

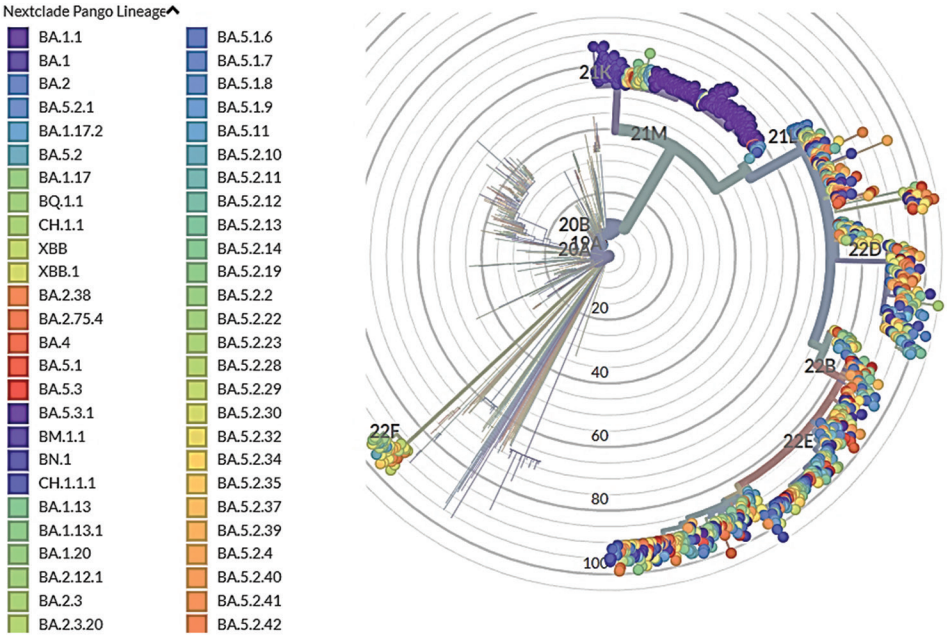


Figure 2. SARS-CoV-2 phylogenetic tree of Omicron variants of Iraq. Nextstrain tool was used to build the phylogenetic tree utilizing all 381 Omicron variant sequences deposited in the GISAID database. Pango lineages are indicated with different colors, while the phylogenic tree illustrating the Nextstrain clades is portrayed in gray color.

for only 4% of all cases in January 2022, BA.2 became the predominant lineage (100%) that infected all cases in May 2022. Afterward, BA.2 lineage was not reported in any isolates until November 2022. On the other hand, BA.5.2 was the most prevalent lineage in June, August, and September 2022. November 2022 saw the most diverse blend of lineages, and several new variants, such as CV.1, BN.1, XBB.1, and XBB.2, were reported for the first time in Iraq (Figure 3A).

The weekly morbidity and mortality data of COVID-19 cases in Iraq, which the data were obtained from Worldometers.info database (www.worldometers.info/coronavirus/country/iraq/)[24], were analyzed and visualized (Figure 3B and C). As shown in Figure 3B, two waves of COVID-19 in Iraq were noted during the span of this study. The majority of the confirmed COVID-19 cases were recorded during the period between December 2021 and February 2022, and toward the end of March 2022, a less severe form of the disease prevailed as the most common form of COVID-19, with the same trend persisting until the end of July 2022 (Figure 3B). Notably, the substantial portion of the COVID-19 deaths was recorded in February and March of 2022 (Figure 3C).

The evolutionary relationship between the different lineages of the 381 SARS-CoV-2 sequences from Iraq and other known lineages was analyzed and visualized using the Nextstrain database (https://nextstrain.org/) (Figure 2).

3.2. Amino acid mutations

SARS-CoV-2 is an RNA virus that can rapidly mutate during cell infection and replication. Our results showed that there are 54 substitutions, ten deletions, and one insertion mutation within the genome of 381 sequences of the SARS-CoV-2 Omicron variants among the Iraqi individuals (Figure 4). Among the mutations, 13 substitutions and all ten amino acid deletions are located in the N-terminal domain (NTD) of the protein (region from 13 to 305 amino acids of the protein). The top three prevalent mutation sites were A67V (67%), L212I (63.0%), and G142D (55.0%) (Figure 5), and H69del and V70del deletions were detected in 70% of the Omicron sequences (Figure 6). The prevalence rates of the other mutations are shown in Figures 5 and 6.

Twenty-five mutations reside in the receptor-binding domain (RBD), including G339D, K417N, and N440K with a prevalence rate of 99%, followed by S373P and S375F with 96% (Figure 5). G446S had a prevalence rate of 80%, while a combination of substitution mutations, including S477N, T478K, E484A, Q498R, N501Y, and Y505H, had a prevalence rate of 75% (Figure 5).

Other known mutations are located at the S1 and S2 subunits of S protein, all with the mutation rates of >60%. Notably, D614G, a signature mutation of SARS-CoV-2, was reported in only 78% of the isolates (Figure 5). Furthermore, three substitution mutations are located at

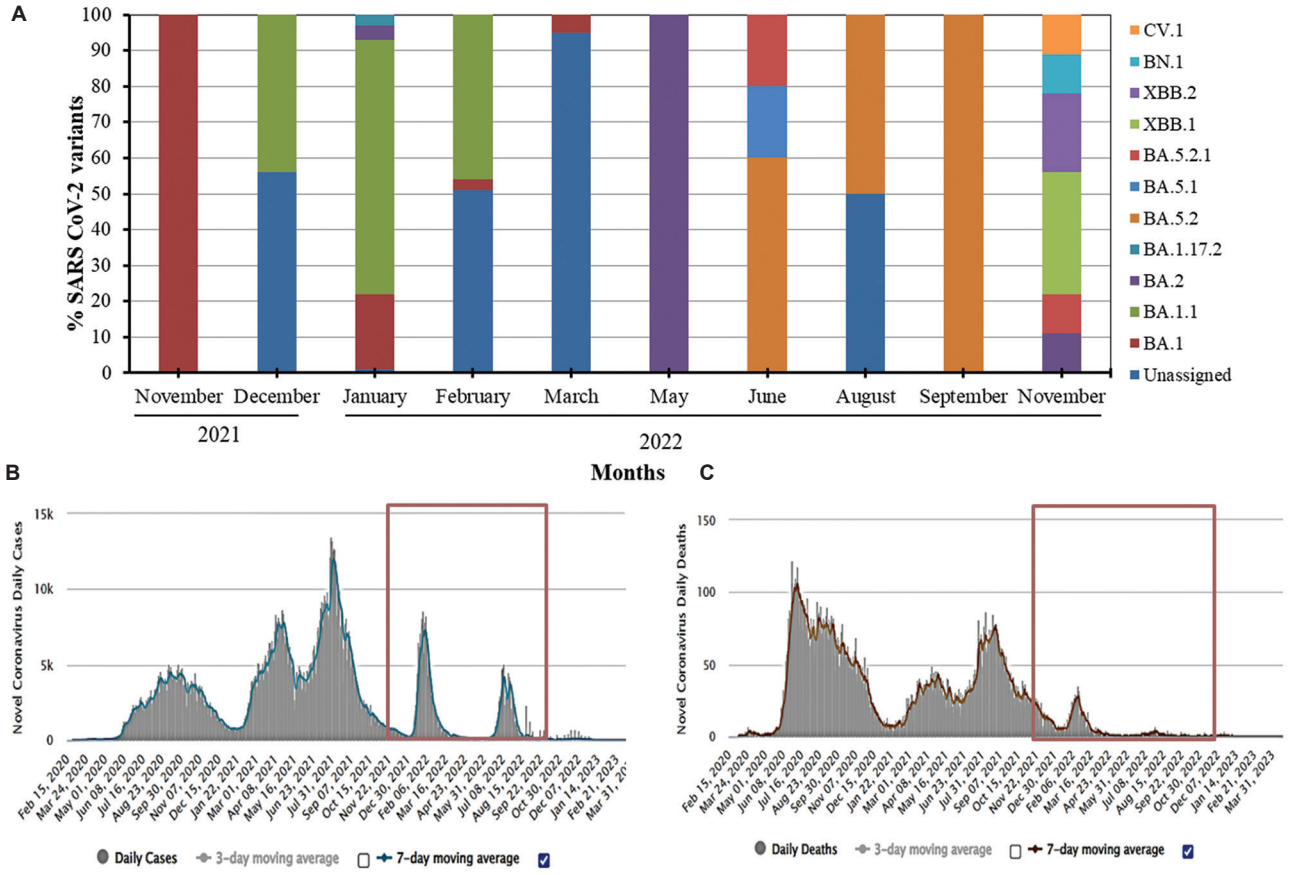


Figure 3. (A) Percentage of the different SARS-CoV-2 Omicron lineages circulating in Iraq during the span of this study. (B and C) Daily morbidity and mortality cases in Iraq, respectively. The graphs in (B) and (C) were obtained from the Worldometers website (https://www.worldometers.info/coronavirus/country/iraq/; Data accessed on May 31, 2023)[24].

the heptad repeat (HR)-1 domain of the S2 subunit of the S protein, namely, Q954H (73%), N969K (73%), and L981F (66%) (Figure 5). It was also observed that 5 of the 60 site mutations of the S protein had more than one substitution, including V213G (Prevalence rate 12%) and V213E = 13%. Furthermore, there were G399D (Prevalence rate = 15%) and G399H=16% (Figure 5).

4. Discussion

Similar to other RNA viruses, SARS-CoV-2 has a high mutation rate, which has led to the emergence of different variants of the virus with varying levels of pathogenicity and transmissibility[25,26]. The Omicron variant (B.1.1.529), which was first reported in November 2021 in Botswana, was categorized as a VOC by the World Health Organization on November 26, 2021[26].

4.1. Omicron variants

The genome of the Omicron variant contains a wide range of deletions and insertions relative to the first SARS-CoV-2, including 31 – 37 mutations located in the S protein[27].

The result of this study has shown that BA.1.1 (52%), BA.1 (14%), BA.2 (3%), and BA.17.2 (2%) were the most prevalent Omicron variants that circulated in Iraq from late 2021 to the end of November 2022. The emergence of new highly mutated sub-lineages of BA.1, BA.1.1, and BA.2 raised concerns about the pathogenicity of these variants[28]. BA.2 variant was first reported on October 22, 2021[29], and was shown to have a 30% increased transmissibility in comparison to BA.1[26]. However, this study showed that BA.2 was less prevalent among Iraqi patients in comparison to the BA.1 and BA.1.1 variants (Figure 1). This perhaps explains the lower COVID-19 mortality and morbidity rates recorded in Iraq during the period of this study in comparison to neighboring countries, such as Iran and Turkey (Figure 3B and C)[30]. Moreover, it has been shown that the mortality and morbidity rates during the Omicron wave were reduced by 20–80% in comparison with those during the Delta wave[26].

Although BA.1, BA.1.1, and BA.2 have shown to share most of the Omicron mutations, these sub-lineages differ in some distinctive mutations. For instance, R346K is a

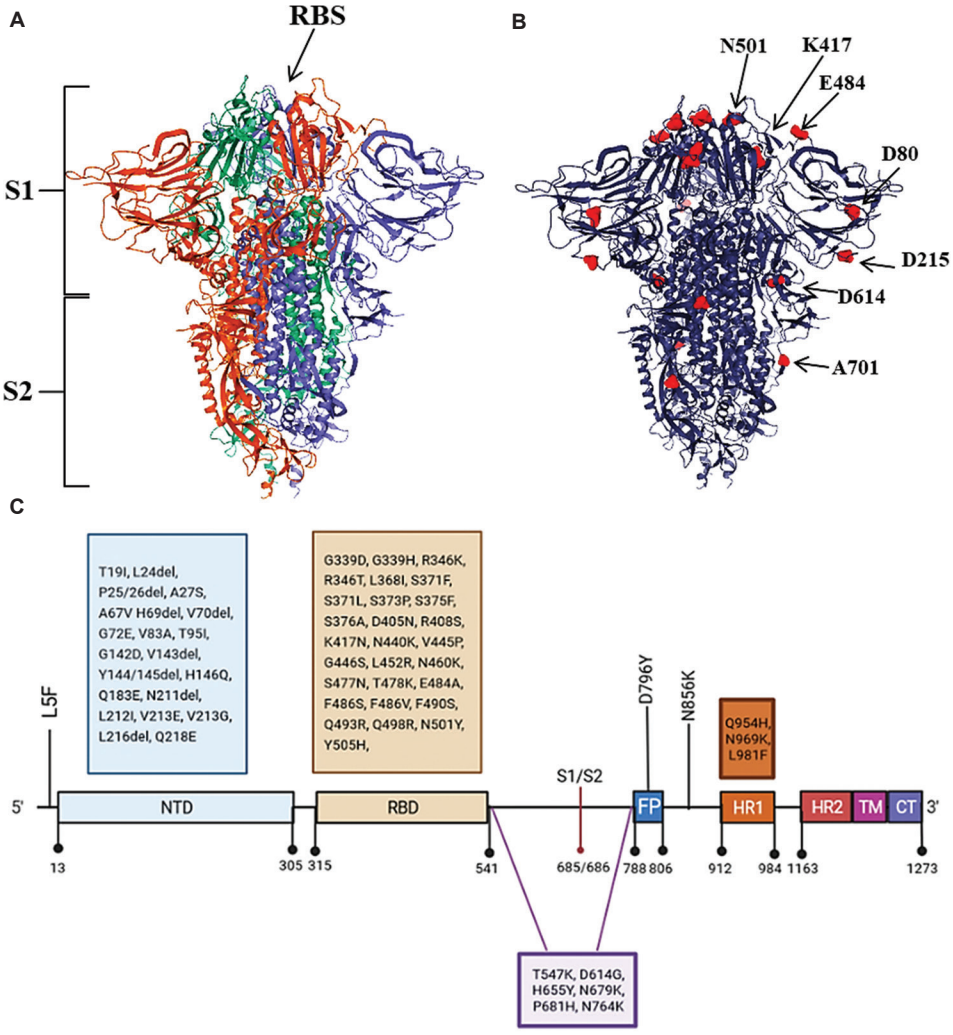


Figure 4. (A) Schematic representation of the SARS-CoV-2 S protein. (B) Locations of some mutations on the protein. (C) Approximate locations of the substations and deletions in the different domains of the S protein.
Abbreviations: NTD: N-terminal domain; RBD: Receptor-binding domain; S1/S2: Protease cleavage site; FP: Fusion peptide; HR1: Heptad Repeat 1; HR2: Heptad repeat 2; TM: Transmembrane domain; CP: Cytoplasmic peptide.

signature substitution of BA.1.1 sub-lineage, whereas T19I, L24S, del25/27, V213G, T376A, and R408S are signature mutations in BA.2 isolates[26].

Monthly SARS-CoV-2 Omicron variant analysis revealed that the different lineages of the virus appeared in different periods of the study (Figure 3). In November 2021, BA.1 was the predominant variant; this variant is characterized by the possession of 35 mutations, of which 15 mutations are located in the receptor-binding site of the S protein[31]. At the end of February 2022, the Omicron BA.1 and its derivates, including BA.1.1, and BA.1.17.2, were found among the sequenced samples. However, starting May 2022, BA.1 was replaced by BA.2, BA.5, and its derivate as the predominant variants. This phenomenon lends its support in explaining the higher transmissibility of BA.2 relative to BA.1[26]. By the end of November 2022, several SARS-CoV-2 variants, such as CV.1, BA.2, BN.1, BA.5.2.1, XBB.1, and XBB.2, were recorded in Iraq (Figure 3A). The emergence of such diverse variants of the SARS-CoV-2 within 1 year can be rationalized by the polymerase mutations of the virus occurring at a rapid pace[32]. In addition, genetic recombination was also reported in individuals coinfected with different SARS-CoV-2 variants[33].

The phylogenetic tree analyses of the sequences showed that the majority of the studied sequences were clustered at the 21K clade, followed by fewer samples clustered at the 21L clade. Most importantly, both clades diverged from 21M, which is a designated clade of the Omicron variant of SARS-CoV-2[34] (Figure 2).

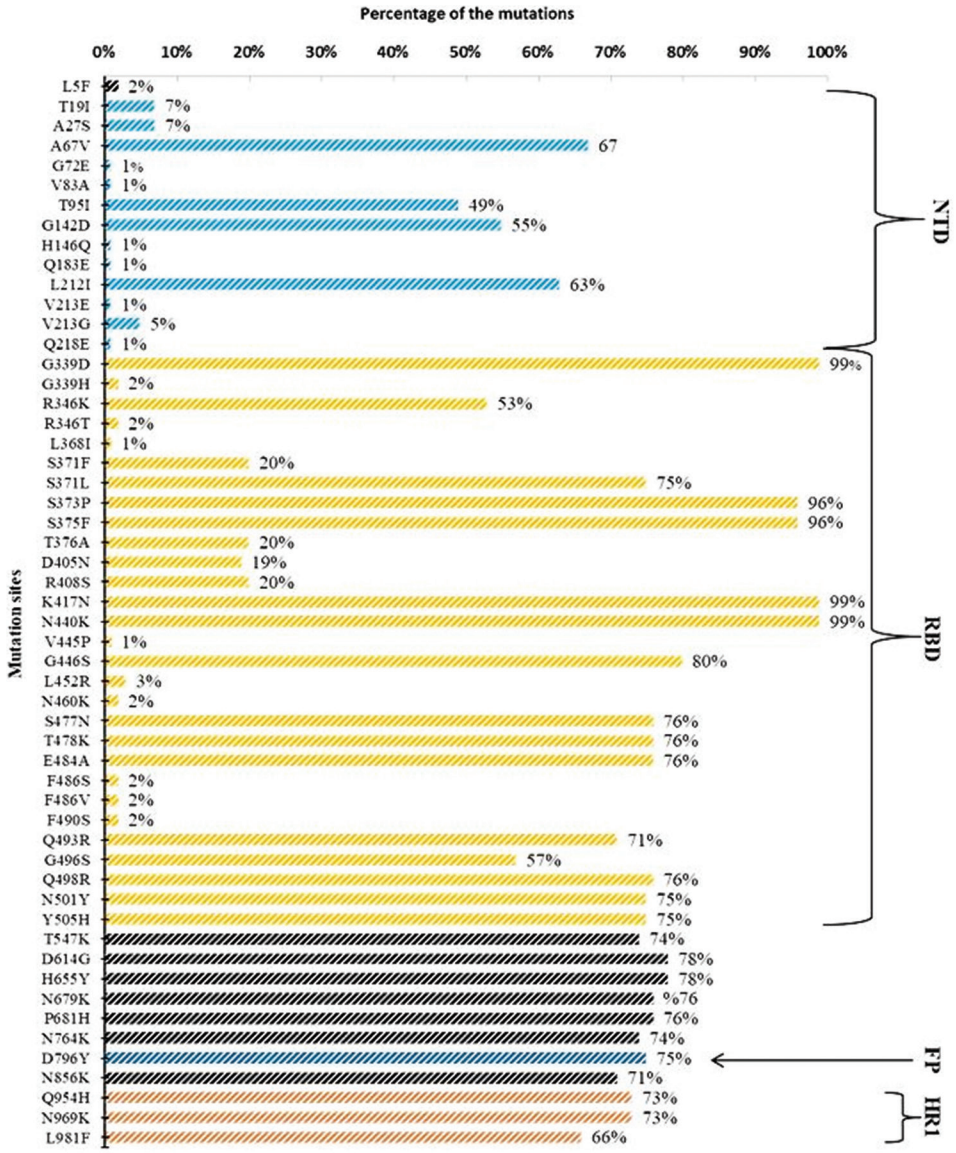


Figure 5. Percentage of the amino acid substitutions of the S protein among the Iraqi Omicron variants that emerged from November 2021 to November 2022. Abbreviations: NTD: N-terminal domain; RBD: Receptor-binding domain; FP: Fusion peptide; HR1: Heptad Repeat 1.

4.2. Mutation in Iraqi Omicron variants

In our study, 23 out of 60 mutations (38%), including ten deletions and one substitution, were located in the NTD of the S protein (Figures 4–6). The top three most prevalent substitution mutations were A67V (67%), L212I (63%), and G142D (55%). A few studies have illustrated the importance of the NTD mutations, especially T19I and P25del, V143del, Y144del, and Y145del in conferring the immune escape capacity on the variants against the neutralizing antibodies produced by the host cells[35,36]. Interestingly, T95I which accounted for 49% of the total cases (Figure 5), as well as del25 and del143–145 which account for prevalence rates of 7%, 49%, 50%, and 49%, respectively, were among the Omicron variants detected in Iraq (Figure 6). Furthermore, the G142D mutation, which was found in 55% of the available sequences (Figure 5), can cause alteration to the binding sites of the protein, which help foster resistance of the protein to monoclonal antibody[37,38]. Other deletions that were detected among the S protein sequences, which can interfere with the pathogenicity of the virus[39], were H69/V70 (70%) and Y144/145del deletion (50%) (Figure 6).

The RBD of the Omicron variant is a highly mutable region[40,41], and 42% ($n = 25$) of all the mutations are located

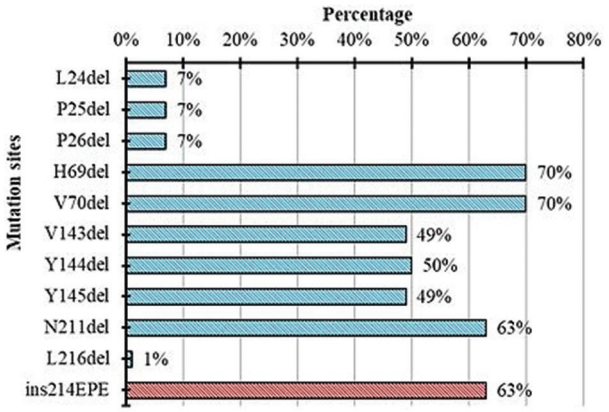


Figure 6. Percentage of amino acid deletion mutations and insertion mutations in the N-terminal of the S protein of the SARS-CoV-2 Omicron variants.

in this domain of the protein (Figures 4 and 5). Among the mutations, G339D, S371D, S375F, K417N, and N440K had prevalence rates of 99% (Figure 5). The important role of the RBD in the pathogenicity of the virus is well documented since this region is where the S protein directly binds to the ACE2[42]. It has also been shown that mutations located between positions 338 and 506 in the RBD are responsible for altering the binding affinity of the S protein with ACE2[43,44]. In addition, S371L results in higher antibody resistance, whereas S373P, 375F, and K417N reduce protein stability and enhance infectivity[45]. N440K and G446S mutations have been shown to increase antibody resistance and infection vulnerability[46]. Similarly, it has been reported that reduced protein stability and enhanced infection rate are the prime consequences following S477N, T478K, E484A, Q496S, Q498R, and Y505H substitutions[45,47]. Other substitutions have also been found to structurally alter the viral protein and thus its capabilities in invading and infecting the host in different aspects. For instance, N501Y could alter the structural conformation of S protein at RBD and invade the host's immune system[48]; L452R could reduce its ability in antibody neutralization[49]; and T478K could facilitate the immune escape of the virus[50].

Glycine substituted for aspartic acid at position 614 in the S protein (D614G), first reported in the early 2020, enhances the transmissibility and infectivity of SARS-CoV-2[51,52]. This mutation also increases S protein flexibility as a result of the open conformation of RBD domain[51,53]. Nevertheless, only 78% of the sequenced samples harbor this mutation. In addition, H655Y, N679K, and P681H mutations that are located at the S1/S2 boundary may affect the splitting process of S1 and S2 subunits. Although P681H might improve the efficiency of proteolytic cleavage of the S protein in the alpha variants at a lesser extent, its influence on viral fitness is not very significant[54]. On the other hands, N679K and P681H mutations have been found to regulate the fusogenic ability of Omicron variants[55].

Q954H and N969K mutations, which are located in the HR domain, have been found in all Omicron sub-lineages, and only affect viral fusion and infection ability, since these mutations do not impact the HR1-HR2 connections[56].

5. Conclusion

This study analyzed all the available data about the diversity of SARS-CoV-2 Omicron sub-lineages detected in Iraq. The BA.1, BA.1.1, BA.2, and BA.5.1 variants are among the Omicron sub-lineages commonly found in the Iraqi patients, while the less common ones include XBB.1 and XBB.2. The BA.1 had been the most common variant among the samples before it was replaced by the more transmissible sub-lineages such as BA.2 and BA.5 starting May 2022. Overall, a total of 60 mutation sites were recognized: 49 substitution sites, ten deletions, and one insertion from the Iraqi Omicron variants. Signature mutations of Omicron variants, such as G339D, S371L, S373P, S375F, E484A, Q496S, and Q498R, are beneficial to the viral fitness in terms of the transmissibility of the variants. Of note, a lack of SARS-CoV-2 genome sequences across the entire period of the study is the limitation of this study, potentially compromising the generalizability of the current findings to similar contexts. As part of its future direction, the current study is set to evaluate the impact of different vaccines on the viral variants analyzed.

Acknowledgments

None.

Funding

None.

Conflict of interest

The author declares no competing interests.

Author contributions

This is single-authored article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The accessory data are given in the Raw Data File.

References

1. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, 2020, The species Severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol*, 5: 536–544.

   https://doi.org/10.1038/s41564-020-0695-z

2. Zhou P, Yang XL, Wang XG, *et al.*, 2020, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579: 270–273.

   https://doi.org/10.1038/s41586-020-2012-7

3. Waris A, Atta UK, Ali M, *et al.*, 2020, COVID-19 outbreak: Current scenario of Pakistan. *New Microbes New Infect*, 35: 100681.

   https://doi.org/10.1016/j.nmni.2020.100681

4. Waris A, Ali M, Khan AU, *et al.*, 2020, COVID-19 incidence in Pakistan: Gender disparity. *Iran J Psychiatry Behav Sci*, 14: e105990.

5. World Health Organization, 2023, COVID-19 Weekly Epidemiological Update. Geneva: World Health Organization.

6. World Health Organization, 2022, World Health Organization (WHO) Coronavirus (COVID-19) Dashboard. World Health Organization. Available from: https://covid19who.int [Last accessed on 2022 Feb 28].

7. Alsayed R, Ali A, Makia R, *et al.*, 2020, Challenges facing Iraq to tackle the spread of COVID-19: An overview. *J Univ Anbar Pure Sci*, 14: 22–27.

8. Fayad N, Abi Habib W, Kandeil A, *et al.*, 2021, SARS-CoV-2 variants in Lebanon: Evolution and current situation. *Biology* (*Basel*), 10: 531.

   https://doi.org/10.3390/biology10060531

9. Kim D, Lee JY, Yang JS, *et al.*, 2020, The architecture of SARS-CoV-2 transcriptome. *Cell*, 181: 914–921.e10.

   https://doi.org/10.1016/j.cell.2020.04.011

10. Wang R, Hozumi Y, Yin C, *et al.*, 2020, Mutations on COVID-19 diagnostic targets. *Genomics*, 112: 5204–5213.

    https://doi.org/10.1016/j.ygeno.2020.09.028

11. Sabir DK, 2022, Analysis of SARS-COV2 spike protein variants among Iraqi isolates. *Gene Rep*, 26: 101420.

    https://doi.org/10.1016/j.genrep.2021.101420

12. Yang H, Rao Z, 2021, Structural biology of SARS-CoV-2 and implications for therapeutic development. *Nat Rev Microbiol*, 19: 685–700.

    https://doi.org/10.1038/s41579-021-00630-8

13. Tilli CMLJ, Van Steensel MA, Krekels GA, *et al.*, 2005, Molecular aetiology and pathogenesis of basal cell carcinoma. *Br J Dermatol*, 152: 1108–1124.

    https://doi.org/10.1111/j.1365-2133.2005.06587.x

14. Brüssow H, 2022, COVID-19: Omicron-the latest, the least virulent, but probably not the last variant of concern of SARS-CoV-2. *Microb Biotechnol*, 15: 1927–1939.

    https://doi.org/10.1111/1751-7915.14064

15. Lauring AS, Hodcroft E, 2021, Genetic variants of SARS-CoV-2-what do they mean? *JAMA*, 325: 529–531.

    https://doi.org/10.1001/jama.2020.27124

16. Ledford H, 2021, How severe are Omicron infections. *Nature*, 600: 577–578.

    https://doi.org/10.1038/d41586-021-03794-8

17. Burki TK, 2022, Omicron variant and booster COVID-19 vaccines. *Lancet Respir Med*, 10: e17.

    https://doi.org/10.1016/S2213-2600(21)00559-2

18. da Silva SJR, Kohl A, Pena L, *et al.*, 2022, Recent insights into SARS-CoV-2 omicron variant. *Rev Med Virol*, 33: e2373.

    https://doi.org/10.1002/rmv.2373

19. Zamel D, Khan AU, 2022, Covid-19 vaccines: Overview, comparative analysis and dynamics. *Acta Sci Biotechnol Vol*, 3: 24–33.

20. Karim SSA, Karim QA, 2021, Omicron SARS-CoV-2 variant: A new chapter in the COVID-19 pandemic. *Lancet*, 398: 2126–2128.

    https://doi.org/10.1016/S0140-6736(21)02758-6

21. Sun J, Xie T, Jamal M, *et al.*, 2020, CLEC3B as a potential diagnostic and prognostic biomarker in lung cancer and association with the immune microenvironment. *Cancer Cell Int*, 20: 106.

    https://doi.org/10.1186/s12935-020-01183-1

22. Hulo C, de Castro E, Masson P, *et al.*, 2011, ViralZone: A knowledge resource to understand virus diversity. *Nucleic Acids Res*, 39: D576–D582.

    https://doi.org/10.1093/nar/gkq901

23. Hadfield J, Megill C, Bell SM, *et al.*, 2018, Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics*, 34: 4121–4123.

    https://doi.org/10.1093/bioinformatics/bty407

24. Worldometers.info. Available from: https://www.worldometers.info/coronavirus/country/iraq [Last accessed on 2023 May 31].

25. Sabir DK, Khwarahm NR, Ali SM, *et al.*, 2020, Children protection against COVID-19 at the pandemic outbreak. *J Immunol Sci*, 4: 8–12.

    https://doi.org/10.29245/2578-3009/2020/2.1188

26. Zhou Y, Zhi H, Teng Y, 2023, The outbreak of SARS-CoV-2 Omicron lineages, immune escape, and vaccine effectivity. *J Med Virol*, 95: e28138.

https://doi.org/10.1002/jmv.28138

27. Tian D, Sun Y, Xu H, *et al*., 2022, The emergence and epidemic characteristics of the highly mutated SARS-CoV-2 Omicron variant. *J Med Virol*, 94: 2376–2383.

https://doi.org/10.1002/jmv.27643

28. Majumdar S, Sarkar R, 2022, Mutational and phylogenetic analyses of the two lineages of the Omicron variant. *J Med Virol*, 94: 1777–1779.

https://doi.org/10.1002/jmv.27558

29. Xia S, Wang L, Zhu Y, *et al*., 2022, Origin, virological features, immune evasion and intervention of SARS-CoV-2 Omicron sublineages. *Signal Transduct Target Ther*, 7: 241.

30. Abdullah HM, Ali SM, Sabir DK, 2020, Data suggesting that COVID-19 may have existed in the kurdistan region of Iraq at the time of the outbreak in wuhan province of China. *J Kermanshah Univ Med Sci*, 24: e105758.

31. Shrestha LB, Foster C, Rawlinson W, *et al*., 2022, Evolution of the SARS-CoV-2 omicron variants BA. 1 to BA. 5: Implications for immune escape and transmission. *Rev Med Virol*, 32: e2381.

https://doi.org/10.1002/rmv.2381

32. Markov PV, Ghafari M, Beer M, *et al*., 2023, The evolution of SARS-CoV-2. *Nat Rev Microbiol*, 21: 361–379.

https://doi.org/10.1038/s41579-023-00878-2

33. He Y, Ma W, Dang S, *et al*., 2022, Possible recombination between two variants of concern in a COVID-19 patient. *Emerg Microbes Infect*, 11: 552–555.

https://doi.org/10.1080/22221751.2022.2032375

34. Balakrishnan KN, Yew CW, Chong ET, *et al*., 2023, Timeline of SARS-CoV-2 transmission in Sabah, Malaysia: Tracking the molecular evolution. *Pathogens*, 12: 1047.

https://doi.org/10.3390/pathogens12081047

35. Fan Y, Li X, Zhang L, *et al*., 2022, SARS-CoV-2 Omicron variant: Recent progress and future perspectives. *Signal Transduct Target Ther*, 7: 141.

https://doi.org/10.1038/s41392-022-00997-x

36. West AP Jr., Wertheim JO, Wang JC, *et al*., 2021, Detection and characterization of the SARS-CoV-2 lineage B.1.526 in New York. *Nat Commun*, 12: 4886.

https://doi.org/10.1038/s41467-021-25168-4

37. Klinakis A, Cournia Z, Rampias T, 2021, N-terminal domain mutations of the spike protein are structurally implicated in epitope recognition in emerging SARS-CoV-2 strains. *Comput Struct Biotechnol J*, 19: 5556–5567.

https://doi.org/10.1016/j.csbj.2021.10.004

38. Suryadevara N, Shrihari S, Gilchuk P, *et al*., 2021, Neutralizing and protective human monoclonal antibodies recognizing the N-terminal domain of the SARS-CoV-2 spike protein. *Cell*, 184: 2316–2331.e15.

https://doi.org/10.1016/j.cell.2021.03.029

39. Meng B, Kemp SA, Papa G, *et al*., 2021, Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Rep*, 35: 109292.

https://doi.org/10.1016/j.celrep.2021.109292

40. Hong Q, Han W, Li J, *et al*., 2022, Molecular basis of receptor binding and antibody neutralization of Omicron. *Nature*, 604: 546–552.

https://doi.org/10.1038/s41586-022-04581-9

41. Cao Y, Yisimayi A, Jian F, *et al*., 2022, BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. *Nature*, 608: 593–602.

https://doi.org/10.1038/s41586-022-04980-y

42. Barton MI, MacGowan SA, Kutuzov MA, *et al*., 2021, Effects of common mutations in the SARS-CoV-2 Spike RBD and its ligand, the human ACE2 receptor on binding affinity and kinetics. *Elife*, 10: e70658.

https://doi.org/10.7554/eLife.70658

43. Cao Y, Wang J, Jian F, *et al*., 2022, Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature*, 602: 657–663.

https://doi.org/10.1038/s41586-021-04385-3

44. Verma S, Patil VM, Gupta MK, 2022, Mutation informatics: SARS-CoV-2 receptor-binding domain of the spike protein. *Drug Discov Today*, 27: 103312.

https://doi.org/10.1016/j.drudis.2022.06.012

45. Kumar S, Thambiraja TS, Karuppanan K, *et al*., 2022, Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein. *J Med Virol*, 94: 1641–1649.

https://doi.org/10.1002/jmv.27526

46. Ali F, Kasry A, Amin M, 2021, The new SARS-CoV-2 strain shows a stronger binding affinity to ACE2 due to N501Y mutant. *Med Drug Discov*, 10: 100086.

https://doi.org/10.1016/j.medidd.2021.100086

47. Quarleri J, Galvan V, Delpino MV, 2022, Omicron Variant of the SARS-CoV-2: A Quest to Define the Consequences of its High Mutational Load. Berlin: Springer. pp1–4.

48. Chakraborty C, Saha A, Sharma AR, *et al*., 2021, D614G mutation eventuates in all VOI and VOC in SARS-CoV-2: Is it part of the positive selection pioneered by Darwin? *Mol Ther Nucleic Acids*, 26: 237–241.

https://doi.org/10.1016/j.omtn.2021.07.011

49. Mittal A, Khattri A, Verma V, 2022, Structural and antigenic variations in the spike protein of emerging SARS-CoV-2 variants. *PLoS Pathog*, 18: e1010260.

https://doi.org/10.1371/journal.ppat.1010260

50. Tian D, Sun Y, Zhou J, *et al*., 2021, The global epidemic of the SARS-CoV-2 Delta variant, key spike mutations and immune escape. *Front Immunol*, 12: 751778.

https://doi.org/10.3389/fimmu.2021.751778

51. Jackson CB, Zhang L, Farzan M, *et al*., 2021, Functional importance of the D614G mutation in the SARS-CoV-2 spike protein. *Biochem Biophys Res Commun*, 538: 108–115.

https://doi.org/10.1016/j.bbrc.2020.11.026

52. Wang C, Zheng Y, Niu Z, 2021, The virological impacts of SARS-CoV-2 D614G mutation. *J Mol Cell Biol*, 13: 712–720.

https://doi.org/10.1093/jmcb/mjab045

53. Ozono S, Zhang Y, Ode H, *et al*., 2021, SARS-CoV-2 D614G spike mutation increases entry efficiency with enhanced ACE2-binding affinity. *Nat Commun*, 12: 848.

https://doi.org/10.1038/s41467-021-21118-2

54. Lubinski B, Fernandes MHV, Frazier L, *et al*., 2022, Functional evaluation of the P681H mutation on the proteolytic activation of the SARS-CoV-2 variant B.1.1.7 (Alpha) spike. *iScience*, 25: 103589.

55. Saito A, Irie T, Suzuki R, *et al*., 2022, Enhanced fusogenicity and pathogenicity of SARS-CoV-2 Delta P681R mutation. *Nature*, 602: 300–306.

https://doi.org/10.1038/s41586-021-04266-9

56. Xia S, Liu M, Wang C, *et al*., 2020, Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res*, 30: 343–355.

https://doi.org/10.1038/s41422-020-0305-x

ACCSCIENCE PUBLISHING

Gene & Protein in Disease

ORIGINAL RESEARCH ARTICLE

Testosterone as a biomarker of colorectal cancer in the South Indian population

Mohd Younis[1,2,3]*, Sevgi Gezici[4], Amrit Sudershan[2], Sanjeev Kumar Digra[5], Ashma Gupta[6], Arun Meyyazhagan[7], Parvinder Kumar[2,3]*, and Vijaya Anand[1]*

[1]Department of Human Genetics and Molecular Biology, Bharathiar University, Coimbatore, Tamil Nadu, India

[2]Institute of Human Genetics, University of Jammu, Jammu, Jammu and Kashmir, India

[3]Department of Zoology, University of Jammu, Jammu, Jammu and Kashmir, India

[4]Department of Medical Biology, Faculty of Medicine, Gaziantep University, Gaziantep, Turkey

[5]Department of Pediatrics, SMGS Hospital, Government Medical College, Jammu, Jammu and Kashmir, India

[6]Department of Zoology, Government College for Women, Gandhi Nagar, Jammu, Jammu and Kashmir, India

[7]Department of Translation Medicine and Surgery, University of Perugia, Perugia, Italy

Abstract

Colorectal cancer (CRC) is a highly life-threatening disease associated with a significant mortality rate. It has been proposed that testosterone levels may play a role in predisposing individuals to such devastating conditions. Testosterone primarily governs the maturation of the male reproductive system while also exerting physiological effects in both genders. In the present study, we investigated testosterone levels in CRC patients among the South Indian population. Blood samples were collected in the hospitals in Tamil Nadu, South India, and a cohort of healthy controls was selected for comparative analysis. A total of 130 subjects participated in the study, consisting of 65 CRC patients and an equal number of healthy controls. Approximately 7 mL of blood was collected from each subject for radioimmunoassay. The results of radioimmunoassay on the blood samples were analyzed using SPSS to assess differences between discrete and continuous data variables. Chi-square and *t*-tests were conducted for statistical evaluation. CRC patients exhibited significantly ($P < 0.0001$) reduced mean testosterone levels (06.68 ± 2.15 nmol/L) compared to controls (22.54 ± 8.85 nmol/L). Further stratification by smoking status revealed that non-smoker CRC patients had lower testosterone levels (06.81 ± 2.21 nmol/L) than non-smoking controls (10.15 ± 2.48 nmol/L), with a statistically significant difference ($P < 0.0001$). Adjusting for alcohol consumption, CRC patients displayed decreased mean testosterone levels (06.31 ± 2.30 nmol/L) compared to controls (07.96 ± 2.45 nmol/L), and this difference was found to be significant ($P < 0.022$). These findings support the notion that reduced testosterone levels serve as a critical risk biomarker in the pathogenesis of CRC.

Keywords: Alcohol; Biomarker; Colorectal cancer; Screening; Smoking; Testosterone

***Corresponding authors:**
Mohd Younis
(younisgenetic@gmail.com)
Parvinder Kumar
(parvinderkb2003@gmail.com)
Vijaya Anand
(avamiet@yahoo.com)

Citation: Younis M, Gezici S, Sudershan A, *et al.*, 2023, Testosterone as a biomarker of colorectal cancer in the South Indian population. *Gene Protein Dis,* 2(3): 1082.
https://doi.org/10.36922/gpd.1082

Received: June 14, 2023
Accepted: September 19, 2023
Published Online: September 29, 2023

1. Introduction

Cancer is a devastating disease that has been the subject of continuous study for over a century. According to the International Agency for Research on Cancer, cancer was responsible for nearly 1.9 million incidence cases and 0.9 million deaths in 2020. One specific example of cancer is colorectal cancer (CRC), which affects the colon and the rectum[1,2]. Examining the connection between testosterone and CRC within the South Indian population is imperative, given the unique characteristics of this demographic and the potential implications for cancer prevention and treatment. Such research can substantially enhance our understanding of CRC and enable the customization of healthcare strategies to address the specific needs of this region.

Various factors have been identified as influential in determining cancer susceptibility[3]. As an illustration, testosterone levels are widely recognized for their significance, particularly in the development of the male reproductive system[4]. Testosterone is primarily synthesized in the smooth endoplasmic reticulum of Leydig cells, which are located in the testes and produced through the traditional pathway from cholesterol[5]. In the testicular interstitial tissue, Leydig cells serve the primary function of producing testosterone to support spermatogenesis and various extratesticular androgenic, and anabolic functions. The transport of testosterone to its target sites relies on the serum glycoprotein called sex hormone-binding globulin (SHBG)[6], which is synthesized in the liver. SHBG forms strong bonds with three sex hormones: estrogen, dihydrotestosterone, and testosterone, facilitating these three hormones throughout the body. Reference ranges for testosterone levels can vary significantly, with reported limits for men ranging from as low as 10 to as high as 35 nmol/L, and for women, from as low as 0.5 to as high as 2.4 nmol/L[7-10].

Recently, emerging evidence has established a statistically significant positive relationship between circulating endogenous sex hormone levels and the risk of developing CRC[6,11,12]. In men, higher levels of total testosterone have been associated with a decreased risk of CRC, and there appears to be an inverse relationship between the estradiol-to-testosterone ratio and CRC[13]. Both genetic and environmental factors, which can influence sex hormone levels, have been linked to an increased risk of CRC[14-17]. Within the Women's Health Initiative Clinical Trial (WHI-CT), a nested research study revealed that exposure to the highest and lowest SHBG concentration groups was associated with a more than twofold increased risk of CRC[6,18]. However, several smaller studies examining the relationship between circulating SHBG concentrations and CRC have yielded conflicting results[13,19]. To the best of our knowledge, no epidemiologic research has examined the relationship between testosterone levels and the risk of CRC in the South Indian population. This study aimed to evaluate the relationships between testosterone levels and the likelihood of CRC among patients in this region. Investigating the relationship between testosterone levels and CRC in the South Indian population holds significant promise for advancing medical knowledge and patient care. It has the potential to provide population-specific insights, contribute to personalized medicine, improve risk assessment, guide treatment strategies, inform preventive measures, advance scientific understanding, and have broader implications for global health. This research has the potential to improve the quality of life for individuals in South India and contribute to the global effort to combat CRC.

2. Materials and methods

2.1. Subject recruitment

In the present study, blood samples of CRC patients were collected from hospitals in Tamil Nadu, South India. CRC was diagnosed based on distinct peripheral blood cytology, standard clinic-hematological criteria, and baseline clinical findings that were recorded and confirmed by an oncologist. Patients diagnosed at various stages of the illness were enrolled in the study, and diagnostic criteria were applied. Each participant underwent a personal interview and completed an open-ended questionnaire to obtain pertinent clinical information, including factors such as age, smoking behaviors, and alcohol history. The study was conducted based on the medical data provided by the practitioners. The samples, collected by the doctors for the investigation, were further used in subsequent experiments.

2.2. Sample collection

Approximately 7 mL of blood samples were collected from patients diagnosed with CRC through venipuncture at hospitals. Blood samples were also voluntarily collected from healthy individuals using the same method to serve as controls for comparative analysis. The overarching goal of employing the same blood collection method is to ensure that control subjects closely resemble patients in all relevant aspects except for the specific exposure or condition under investigation. This approach allows researchers to make valid comparisons and derive meaningful conclusions from their study findings. Venipuncture is a common medical procedure that demands careful attention to detail and strict adherence to safety protocols to ensure the proper collection and handling of blood samples while minimizing discomfort or complications for the patient. The blood was

drawn into ethylene diamine tetraacetic acid-filled sterile tubes and then delivered to the immunobiology laboratory. Serum and plasma were subsequently separated from the blood samples and assayed for biochemical parameters, specifically focusing on the levels of testosterone.

2.3. The testosterone assay

2.3.1. Enzyme immunoassay (EIA)

Testosterones were directly measured in the serum using radioimmunoassay kits (Catalog number: 07BC-1115, Bangalore Genei Private Limited) with reference values[20]. The detection method is based on the competitive binding of testosterone in the test specimen with a testosterone-horseradish peroxidase (HRP) conjugate for a consistent amount of rabbit anti-testosterone. During the incubation, testosterone standards, control samples, patient samples, testosterone-HRP conjugate reagent, and rabbit anti-testosterone reagent were all incubated for 90 min within goat anti-rabbit IgG-coated wells. Subsequently, the wells were rinsed to remove unbound testosterone-peroxidase conjugate, and the addition of tetramethylbenzidine (TMB) solution induced the development of a blue color. TMB is a chromogenic substrate used in immunohistochemical staining processes and as a visualizing reagent in enzyme-linked immunosorbent tests. To measure the absorbance spectrophotometrically at 450 nm, the color development was halted. The standard curve served to determine the testosterone concentration of the samples and controls in relation to the standards.

2.3.2. Reagents composition

Microtiter plates with 96 antibody-coated wells, with goat anti-rabbit IgG coating, were used. The reference standard set (1 mL/vial) contains preservative-free human serum with testosterone concentrations of 0, 0.1, 0.5, 2.0, 6.0, and 18.0 ng/mL. The pink-colored, 7mL rabbit anti-testosterone reagent comprises anti-testosterone in a bovine serum albumin buffer with preservatives. The testosterone-HRP conjugate reagent (12 mL/vial, blue color), designed for the quantitative determination of testosterone in human serum or plasma, comprises only HRP-conjugated testosterone. Testosterone control sets one and two each include approximately 0.6 and 11 ng/mL of testosterone in human serum (0.5 mL/vial). The 11 mL container of TMB reagent contains 3,3',5,5'-TMB stabilized in a buffer solution. In addition, a bottle of stop solution containing 11 mL of diluted hydrochloric acid (1N HCl) is included. Before use, all reagents should be equilibrated to room temperature (18 – 25℃) and gently inverted or swirled to ensure proper mixing. Testosterone concentrations in samples were quantified through dilution with diluents. To prevent exposure to humid air, a desiccant was included to seal the microtiter plate in a bag. The unopened kit was stored at 2 – 8℃.

2.3.3. Assay procedure

The desired number of coated wells were placed in the holder. Next, 10 µL each of standards, samples, and controls were dispensed into separate wells. Subsequently, 100 µL of testosterone-HRP conjugate reagent and 50 µL of rabbit anti-testosterone reagent were added to each respective well. After thorough mixing for 30 s, the mixture was incubated for 90 min at 37°C. The incubation mixture was then removed from the wells by flicking, followed by rinsing each well five times with deionized water. Subsequently, each well received 100 µL of TMB reagent, which was gently mixed for 10 s and then incubated for 20 min at room temperature. The reaction was stopped by gently adding 100 µL of the stop solution (1N HCl) to each well for 30 seconds. The optical density must be read at 450 nm using a microtiter well reader within 15 min of the complete conversion of all blue colors to yellow. Figure 1 illustrates each step of the process, from sample collection through spectrophotometric analysis.

2.4. Statistical analysis

Descriptive data analysis utilized frequency distribution and mean with standard deviation (SD) statistics. To assess significant differences between discrete and continuous data variables, Chi-square and *t*-tests were used in inferential statistics. ANOVA was used to assess group mean differences. All statistical calculations were performed using the International Business Machines-Statistical Package for the Social Sciences (IBM-SPSS) tool.

3. Results

In the present study, a total of 130 participants were involved, with 65 being diagnosed with CRC, while the remaining 65 served as healthy, cancer-free controls. These participants were further categorized into two groups, namely Group I (comprising subjects under 50 years old) and Group II (comprising subjects over 50 years old), based on "age" criteria. The average age of patients in Group I was 42.62 ± 6.10 (n = 24), while in Group II, it was 6.40 ± 2.03 (n = 41), representing 36.92% and 63.08% of the total, respectively. The age range of CRC-diagnosed patients ranged from 37 – 78 years, while that of the control subjects ranged from 36 – 76 years. There was a deviation of ±2 years in the age distribution between the CRC-diagnosed patient groups and controls.

Patients were also divided based on "disease stage" criteria, including Dukes stage A, which encompassed 8 subjects (12.31%); Dukes stage B, comprising 14 subjects

(21.54%); Dukes stage C, involving 25 subjects (38.46%); and Dukes stage D, which included 18 subjects (27.69%), as depicted in Figure 2A. Furthermore, CRC-diagnosed patients were categorized into six groups based on their exposure to certain risk factors, such as smokers (37 subjects, 56.92%), non-smokers (28 subjects, 43.08%), alcohol consumers (43 subjects, 66.15%), and non-alcohol consumers (22 subjects, 33.85%), as illustrated in Table 1.

Concerning the level of testosterone, CRC-diagnosed patients exhibited a decreased mean testosterone level (06.68 ± 2.15 nmol/L) with a mean age of 54.63 ± 10.24, in contrast to the controls, who had a mean testosterone level

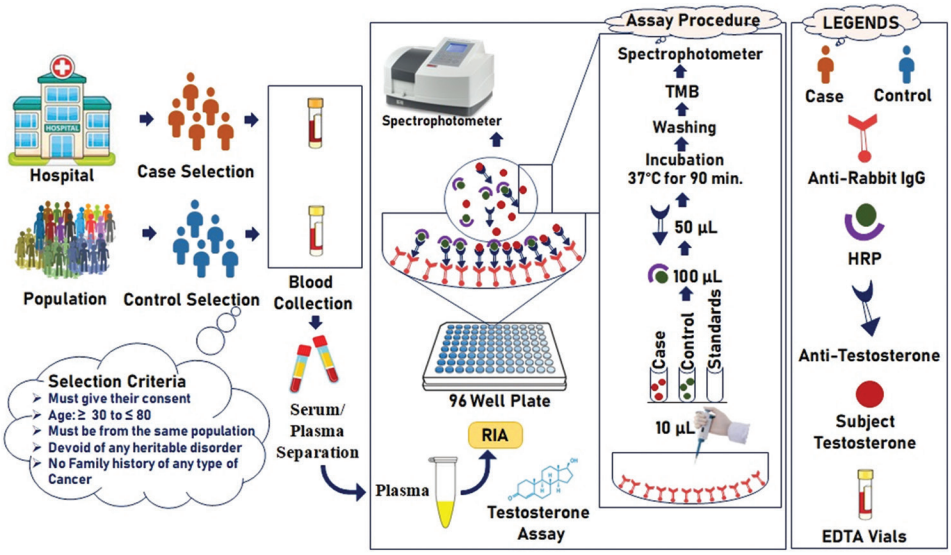


Figure 1. Flow diagram depicting the study procedure.
Abbreviations: EDTA: Ethylenediaminetetraacetic acid; HRP: Horseradish peroxidase; RIA: Radioimmunoassay; TMB: Tetramethylbenzidine.

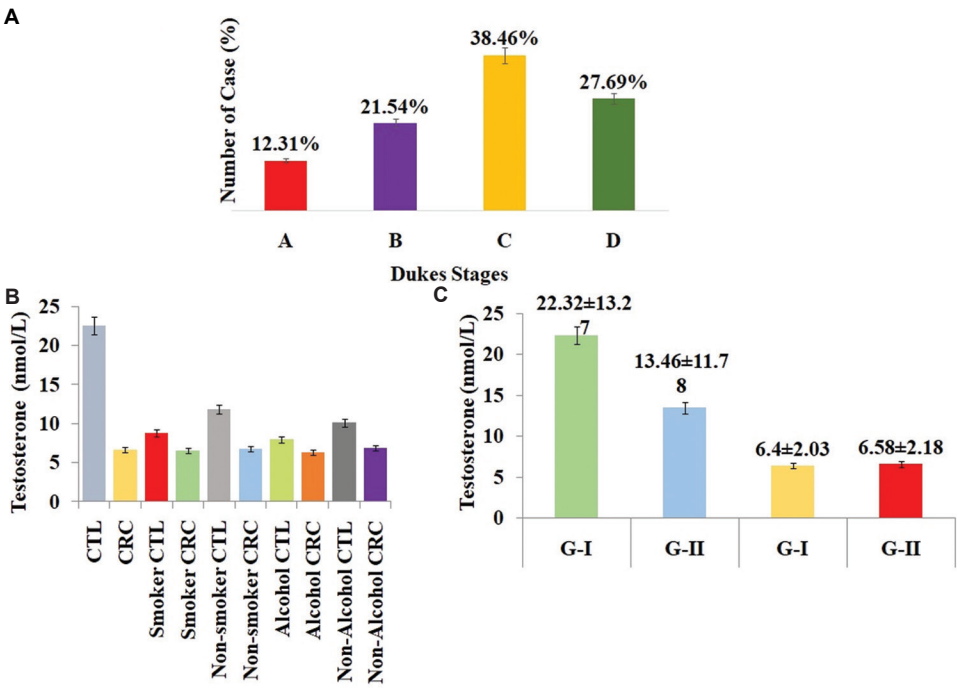


Figure 2. (A) Frequency distribution of CRC-diagnosed patients based on Dukes stages. (B) Testosterone levels in different groups. (C) Testosterone levels adjusted for age.
Notes: CRC: Colorectal cancer patient group; CTL: Control group; G-I: Group I; G-II: Group II.

Table 1. Hormonal frequency distribution of testosterone in colorectal cancer patients and controls based on certain habits (smoker and alcohol)

Variable	Groups	Age range (years)	No. of subjects (n)	Testosterone (nmol/L)	*P*-value
	CTL	36 – 76 (54.63±9.83)	65	22.54±8.85	<0.0001*
	CRC	37 – 78 (54.72±10.24)	65	06.68±2.15	
Smoker	Smoker (CTL)	40 – 67 (55.08±9.22)	12	07.82±1.78	0.07
	Smoker (CRC)	38 – 78 (57.70±9.93)	37	06.57±2.13	
	Non-Smoker (CTL)	36 – 76 (53.79±12.14)	53	11.86±2.65	<0.0001*
	Non-Smoker (CRC)	37 – 67 (50.92±9.36)	28	06.81±2.21	
Alcohol consumer	Alcohol consumer (CTL)	38 – 70 (55.40±10.88)	15	07.96±2.45	<0.022
	Alcohol consumer (CRC)	37 – 78 (55.54±11.05)	43	06.31±2.30	
	Non-alcohol Consumer (CTL)	36 – 76 (51.58±9.64)	50	10.15±2.48	<0.0001*
	Non-alcohol consumer (CRC)	37 – 68 (53.95±8.00)	22	06.87±2.00	

Notes: CRC: Colorectal cancer; CTL: Controls. Values are presented as mean±SD unless specified. *Statistically significant compared to controls (p<0.0001).

of 22.54 ± 8.85 nmol/L with a mean age of 54.63 ± 9.83, as shown in Figure 2B. A highly significant statistical difference was observed between the serum testosterone levels of CRC-diagnosed patients and controls ($t = 14.0399$, $P < 0.0001$), as presented in Table 1. Smoking was found to be significantly ($P < 0.0001$) linked with CRC, with an odds ratio (OR) of 5.8363 and a 95% confidence interval (CI) of 2.6329 – 12.9373 when considering the risk of CRC in the presence of smoking, as shown in Table 1. In addition, alcohol was considered more impactful than smoking, increasing the CRC risk by 6.5-fold (OR: 6.5152, 95% CI: 3.0088 – 14.1077, $P < 0.0001$).

In addition, we observed a slight difference in testosterone levels between the patient group and the control group, both of which consisted of smokers. This difference, however, was not found to be statistically significant ($t = 1.82$, $P = 0.07$). However, when comparing the testosterone levels of subjects in the non-smoker group (patient group: 6.81 ± 2.21 nmol/L; control group: 10.15 ± 2.48 nmol/L), a statistically significant difference was identified ($P < 0.0001$). CRC-diagnosed patients who were adjusted for alcohol consumption exhibited a decreased mean testosterone level (06.31 ± 2.30 nmol/L) in contrast to the controls (7.96 ± 2.45 nmol/L), and this difference was statistically significant ($P < 0.022$), as shown in Table 1. The mean testosterone levels were decreased in CRC-diagnosed patients, who were also smokers and alcohol consumers (06.68 ± 2.15, 06.57 ± 2.13, and 06.31 ± 2.30 nmol/L, respectively) when compared to their respective controls, who were non-smokers, and non-alcohol consumers (22.54 ± 8.85, 11.86 ± 2.65, and 10.15 ± 2.48 nmol/L, respectively), as indicated in Table 1 and Figure 2B. Based on the aforementioned information, it is evident that testosterone levels were statistically lower in CRC patients who were smokers and alcohol consumers when compared to control subjects who were non-smokers and non-alcohol consumers.

Regarding the grouping of subjects by age, CRC-diagnosed patients within Group I, with a mean age of 43.87 ± 3.62, exhibited mean testosterone levels of 6.40 ± 2.03 nmol/L. In contrast, the control group, with a mean age of 42.62 ± 6.10, displayed higher testosterone levels (28.29 ± 6.10 nmol/L), as shown in Figure 2C. This observed difference was substantial and statistically significant ($P < 0.0001$), as indicated in Table 2. Group II, representing a mean age of 62.79 ± 6.79 in CRC-diagnosed patients and 61.08 ± 6.32 in controls, showed a significant difference ($P < 0.0001$) in the mean testosterone levels in CRC-diagnosed patients and controls, i.e., 6.58 ± 2.18 nmol/L and 20.39 ± 7.20 nmol/L, respectively, as demonstrated in Figure 2C.

After conducting the intergroup significant difference test, we also examined intragroup differences among patients at different Dukes stages. In Dukes stage A, the mean testosterone level was 7.43 ± 1.86 nmol/L, while in Dukes stage B, it was 6.00 ± 2.31 nmol/L. Dukes stage C had a mean level of 6.92 ± 2.54 nmol/L, and Dukes stage D showed a mean level of 5.85 ± 2.42 nmol/L. In Group II, which includes patients at Dukes stages C and D, the testosterone level was lower in Dukes stage D, although this difference was not statistically significant ($P = 0.17$). Similarly, in Group I, the mean difference in testosterone levels was not statistically significant ($P = 0.15$), as shown in Table 2. The mean values of testosterone levels in CRC-diagnosed patients at Dukes stage D within Group II were statistically lower compared to both Group I and Group II of the controls ($P < 0.001$). Comparing the

Table 2. Hormonal frequency distribution of testosterone in colorectal cancer patients and controls based on their age

Group	Particulars	Age range (years)	No. of subjects (n)	Testosterone (nmol/L)	*t*-test/ANOVA	*P*-value
I	CTL	36 – 49 (42.62±6.10)	24	28.29±6.10	16.68	<0.0001*
	CRC	37 – 49 (43.87±3.62)	24	6.40±2.03		
II	CTL	51 – 75 (61.08±6.32)	41	20.39±7.20	11.7546	0.0001*
	CRC	51 – 78 (62.79±6.79)	41	6.58±2.18		
I	Dukes stage A (CRC)	38 – 48 (42.37±3.62)	8	7.43±1.86	1.35	0.26
I	Dukes stage B (CRC)	37 – 49 (44.62±3.02)	14	6.0±2.31		
II	Dukes stage C (CRC)	48 – 71 (57.87±5.59)	25	6.92±2.54		
II	Dukes stage D (CRC)	57 – 78 (64.87±6.87)	18	5.85±2.42		

Notes: CRC: Colorectal cancer: CTL: Controls. Dukes Stages: A–D; Group I: Age<50; Group II: Age>50; Values are presented as mean±SD unless specified. *Statistically significant compared to controls (P<0.05).

testosterone levels of Group I at Dukes stage A and B with the controls in Group I, a statistically significant difference was observed (P < 0.0001). Similarly, when comparing the differences between Group II at Dukes stages C and D and the controls in Group II, a significant difference was also observed (P < 0.0001). Interestingly, when utilizing ANOVA, no significant difference was observed among different Dukes stages. This indicates that all observed differences within the Dukes stages were negligible, as shown in Table 2. ANOVA is a justified and widely accepted statistical method in scientific research due to its ability to handle multiple groups, control for experiment-wise error, provide valuable insights into variability, and offer flexibility for various experimental designs. It helps researchers make informed decisions and draw meaningful conclusions from their data, contributing to the advancement of scientific knowledge.

4. Discussion

CRC is a complex illness influenced by various risk factors, including environmental exposure to chemicals, toxins, and carcinogens, as well as endogenous factors. One such endogenous factor is reduced testosterone levels, which play a crucial role in controlling cell proliferation and apoptosis in healthy cells[21], as shown in Figure 3. As testosterone levels decrease due to various factors, such as aging, alcohol consumption, and smoking, the risk of CRC increases. However, there is limited knowledge regarding how testosterone affects cancer cells or the potential inhibitory effects of anti-androgens[22,23]. Recent research has proposed testosterone as a potential tumor biomarker for CRC[24], alongside established tumor markers such as a carcinoembryonic antigen.

In the present study, when comparing the mean testosterone levels among CRC participants who consume alcohol and smoke to those in the control group who were non-smokers, non-drinkers, and control groups, we observed significantly lower testosterone levels (P < 0.05), which is consistent with findings from other investigations[25,26]. The testosterone levels in the CRC-diagnosed patients in this study aligned with those reported in other studies, both globally and across age groups, reaffirming the validity of our results[27]. Furthermore, several other studies have suggested that elevated testosterone levels in CRC patients at Dukes stages B and C may be indicative of more aggressive cancers[28-30]. Increased testosterone levels have been associated with cancer progression and are used as a prognostic indicator for advanced CRC, with marker sensitivity increasing with tumor stages[31-33]. However, it is worth noting that a population-based study involving 3635 males revealed no correlation between testosterone levels and the incidence of colon cancer, which contradicts the findings of our study[34].

Pregnenolone, 17-hydroxypregnenolone, and dehydroepiandrosterone are intermediates in one of the pathways through which testosterone can be synthesized from cholesterol[5,35]. In males, the smooth endoplasmic reticulum of Leydig cells in the testes is primarily responsible for testosterone production. In females, testosterone is primarily produced by the ovaries, suprarenal glands, and adipose tissue. In obese individuals, testosterone is often converted to estradiol by the enzyme aromatase. In addition, androstenedione and testosterone can undergo extra glandular aromatization to produce estrone and estradiol, respectively[36]. SHBG is a hepatically produced glycoprotein that serves as the primary transport protein for testosterone, playing a crucial role in regulating its bioactivity, as shown in Figure 3. Altered levels of SHBG also contribute to CRC risk[19,37,38]. One important site where this regulation occurs is within the colon epithelial cells. When the SHBG complex with testosterone reaches these cells, testosterone is released. Due to its lipophilic nature, it can easily pass through the cell membrane[12,39,40]. Once inside the cytoplasm, it binds to

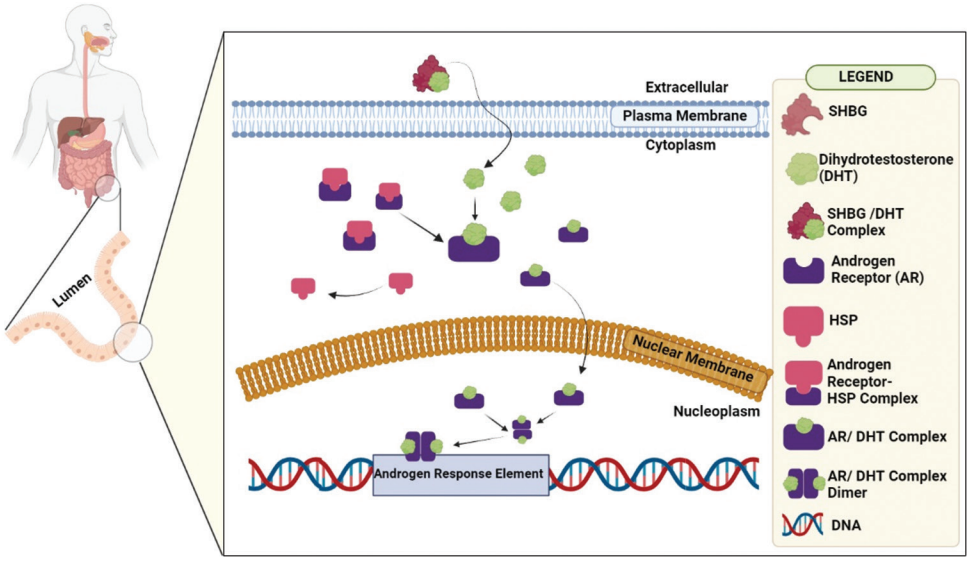


Figure 3. Testosterone signaling in colon epithelial cell: Testosterone is synthesized in the tissue of Leydig cells, where it is then transported by steroid hormone binding globin (SHBG) via the bloodstream to different target sites. One of these sites is within colon epithelial cells. Once the SHBG complex with testosterone reaches this site, testosterone is released. Due to its lipophilic nature, it readily passes through the cell membrane. Once inside the cytoplasm, it binds to its receptor, known as the androgen receptor (AR), a DNA-binding protein complexed with heat shock protein. The testosterone-AR complex is subsequently transported to the nucleus, where it forms a dimer and binds to the androgen response element site, thereby regulating the expression of genes critical for cell proliferation regulation, as well as cell death. In colon cancer, the disruption of this signaling is attributed to low testosterone levels.

its receptor, the androgen receptor (AR), which is a DNA-binding protein complexed with heat-shock protein[38,41,42]. The testosterone-AR complex is then transported to the nucleus, where it forms a dimer and binds to the androgen response element site, regulating the expression of genes critical for cell proliferation, regulation, and cell death[43]. The gender disparity in CRC risk can be attributed to both genetic and environmental variables, both of which have the potential to affect sex hormone levels[16,44,45]. Low testosterone levels result in the disruption of this signaling cascade, as illustrated in Figure 3.

5. Conclusion

The present study supports the notion that a decreased level of testosterone serves as a critical risk biomarker for CRC pathogenesis. The examination of such biomarkers contributes to a comprehensive understanding of the relevance of efficient diagnosis and early screening. Therefore, further molecular and genetic investigations are needed to determine the specific role of testosterone in the human colonic mucosa.

Acknowledgments

The authors are very grateful to all the patients, family members, and patient support groups for their active participation and cooperation during the course of the research study. The authors also thank the hospitals, clinical, and research fraternity for joining hands together to solve the colorectal cancer mysteries in South India. Mohd Younis and Vijaya Anand thank Bharathiar University and DST PURSE-II for constant support, and encouragement.

Funding

This study was supported by DST-PURSE-II, Bharathiar University, Government of India. The fellowship was provided by DST New Delhi and Bharathiar University India.

Conflict of interest

The authors affirm that they do not have any competing interests.

Author contributions

Conceptualization: Mohd Younis, Vijaya Anand
Formal Analysis: Parvinder Kumar, Sanjeev Kumar Digra
Investigation: Mohd Younis, Vijaya Anand
Methodology: Ashma Gupta, Arun Meyyazhagan
Writing – original draft: Mohd Younis
Writing – review & editing: Sevgi Gezici, Amrit Sudershan

Ethics approval and consent to participate

The study procedure and institutional human ethical clearance were obtained from the Bharathiar University Human Ethical Committee (Ethical Reference No. BUHEC-006/2018), and the Helsinki Declaration was followed throughout the study.

Availability of data

All data analyzed were incorporated into the manuscript.

References

1. Siegel RL, Miller KD, Wagle NS, *et al.*, 2023, Cancer statistics, 2023. *CA Cancer J Clin*, 73(1): 17–48.

   https://doi.org/10.3322/caac.21763

2. Xi Y, Xu P, 2021, Global colorectal cancer burden in 2020 and projections to 2040. *Transl Oncol*, 14(10): 101174.

   https://doi.org/10.1016/j.tranon.2021.101174

3. Mohd Y, Kumar P, Bhotla KH, *et al.*, 2021, Transmission jeopardy of Adenomatosis polyposis coli and methylenetetrahydrofolate reductase in colorectal cancer. *J Renin Angiotensin Aldosterone Syst*, 2021: 7010706.

   https://doi.org/10.1155/2021/7010706

4. Bracci M, Zingaretti L, Martelli M, *et al.*, 2023, Alterations in pregnenolone and testosterone levels in male shift workers. *Int J Environ Res Public Health*, 20(4): 3195–4011.

   https://doi.org/10.3390/ijerph20043195

5. Roshan MH, Tambo A, Pace NP, 2016, The role of testosterone in colorectal carcinoma: Pathomechanisms and open questions. *EPMA J*, 7(1): 22.

   https://doi.org/10.1186/s13167-016-0071-5

6. Murphy N, Strickler HD, Stanczyk FZ, *et al.*, 2015, A prospective evaluation of endogenous sex hormone levels and colorectal cancer risk in postmenopausal women. *J Natl Cancer Inst*, 107: 210–227.

   https://doi.org/10.1093/jnci/djv210

7. Vesper HW, Botelho JC, 2010, Standardization of testosterone measurements in humans. *J Steroid Biochem Mol Biol*, 121(3–5): 513–519.

   https://doi.org/10.1016/j.jsbmb.2010.03.032

8. Steinberger E, Ayala C, His B, *et al.*, 1998, Utilization of commercial laboratory results in management of hyperandrogenism in women. *Endocr Pract*, 4(1): 1–10.

   https://doi.org/10.4158/EP.4.1.1

9. Rosner W, Auchus RJ, Azziz R, *et al.*, 2007, Position statement: Utility, limitations, and pitfalls in measuring testosterone: An endocrine society position statement. *J Clin Endocrinol Metab*, 92(2): 405–413.

   https://doi.org/10.1210/jc.2006-1864

10. Bhasin S, Zhang A, Coviell A, *et al.*, 2008, The impact of assay quality and reference ranges on clinical decision making in the diagnosis of androgen disorders. *Steroids*, 73(13): 1311–1317.

    https://doi.org/10.1016/j.steroids.2008.07.003

11. Ahmad R, Singh JK, Wunnava A, *et al.*, 2021, Emerging trends in colorectal cancer: Dysregulated signaling pathways (Review). *Int J Mol Med*, 47(3): 14.

    https://doi.org/10.3892/ijmm.2021.4847

12. Murphy N, Xu L, Zervoudakis A, *et al.*, 2017, Reproductive and menstrual factors and colorectal cancer incidence in the women's health initiative observational study. *Br J Cancer*, 116(1): 117–125.

    https://doi.org/10.1038/bjc.2016.345

13. Lin JH, Zhang SM, Rexrode KM, *et al.*, 2013, Association between sex hormones and colorectal cancer risk in men and women. *Clin Gastroenterol Hepatol*, 11(4): 419–424.e1.

    https://doi.org/10.1016/j.cgh.2012.11.012

14. Shiels MS, Rohrmann S, Menke A, *et al.*, 2009, Association of cigarette smoking, alcohol consumption, and physical activity with sex steroid hormone levels in US men. *Cancer Causes Control*, 20: 877–886.

    https://doi.org/10.1007/s10552-009-9318-y

15. Sainz A, Rudolph R, Hein M, *et al.*, 2011, Association of genetic polymorphisms in ESR2, HSD17B1, ABCB1, and SHBG genes with colorectal cancer risk. *Endocr Relat Cancer*, 18(2): 265–276.

    https://doi.org/10.1530/ERC-10-0264

16. Brenner H, Chen C, 2018, The colorectal cancer epidemic: Challenges and opportunities for primary, secondary and tertiary prevention. *Br J Cancer*, 119: 785–792.

    https://doi.org/10.1038/s41416-018-0264-x

17. Younis M, Sujeetha P, Sangeetha T, *et al.*, 2021, Molecular heterogeneity in colorectal cancer: Insights from genetic profiling. *J Adv Appl Sci Res*, 3(5): 6–15.

    https://doi.org/10.46947/joaasr352021118

18. Dimou N, Mori N, Harlid S, *et al.*, 2021, Circulating levels of testosterone, sex hormone binding globulin and colorectal cancer risk: Observational and Mendelian randomization analyses. *Cancer Epidemiol Biomarkers Prev*, 30(7): 1336–1348.

    https://doi.org/10.1158/1055-9965.EPI-20-1690

19. Mori N, Sawada N, Iwasaki M, *et al.*, 2019, Circulating sex hormone levels and colorectal cancer risk in Japanese postmenopausal women: The JPHC nested case-control study. *Int J Cancer*, 145: 1238–1244.

    https://doi.org/10.1002/ijc.32431

20. Yeap BB, Almeida OP, Hyde Z, *et al.*, 2007, In men older than 70 years, total testosterone remains stable while free testosterone declines with age. The health in men study. *Eur J Endocrinol*, 156(5): 585–594.

    https://doi.org/10.1530/EJE-06-0714

21. Orsted DD, Nordestgaard BG, Bojesen SE, 2014, Plasma testosterone in the general population, cancer prognosis and cancer risk: A prospective cohort study. *Ann Oncol*, 25(3): 712–718.

https://doi.org/10.1093/annonc/mdt590

22. Krasanakis T, Nikolouzakis TK, Sgantzos M, *et al*., 2019, Role of anabolic agents in colorectal carcinogenesis: Myths and realities (Review). *Oncol Rep*, 42(6): 2228–2244.

https://doi.org/10.3892/or.2019.7351

23. Ben-Batalla I, Vargas-Delgado ME, Von Amsberg G, *et al*., 2020, Influence of androgens on immunity to self and foreign: Effects on immunity and cancer. *Front Immunol*, 11: 1184.

https://doi.org/10.3389/fimmu.2020.01184

24. Yang W, Giovannucci EL, Hankinson SE, *et al*., 2020, Endogenous sex hormones and colorectal cancer survival among men and women. *Int J Cancer*, 147(4): 920–930.

https://doi.org/10.1002/ijc.32844

25. Khaw KT, Dowsett M, Folkerd E, *et al*., 2007, Endogenous testosterone and mortality due to all causes, cardiovascular disease, and cancer in men: European prospective investigation into cancer in Norfolk (EPIC-Norfolk) prospective population study. *Circulation*, 116: 2694–2701.

https://doi.org/10.1161/circulationaha.107.719005

26. Younis M, Arun M, Iqbal J, *et al*., 2018, Analysis of genetic alterations in colorectal cancer (CRC) patients in South Indian population. *Int J Hum Genet*, 18(3): 210–218.

https://doi.org/10.31901/24566330.2018/18.3.705

27. Harman SM, Metter EJ, Tobin JD, *et al*., 2001, Longitudinal effects of aging on serum total and free testosterone levels in healthy men. Baltimore longitudinal study of aging. *J Clin Endocrinol Metab*, 86: 724–731.

https://doi.org/10.1210/jcem.86.2.7219

28. Chen CC, Yang SH, Lin JK, *et al*., 2005, Is it reasonable to add preoperative serum level of CEA and CA19-9 to staging for colorectal cancer? *J Surg Res*, 124(2): 169–174.

https://doi.org/10.1016/j.jss.2004.08.013

29. Weissenberger C, Von Plehn G, Otto F, *et al*., 2005, Adjuvant radiochemotherapy of stage II and III rectal adenocarcinoma: Role of CEA and CA 19-9. *Anticancer Res*, 25(3A): 1787–1793.

30. Mohd Y, Balasubramanian B, Meyyazhagan A, *et al*., 2021, Extricating the association between the prognostic factors of colorectal cancer. *J Gastrointest Cancer*, 52: 1022–1028.

https://doi.org/10.1007/s12029-020-00535-4

31. Alberg AJ, Gordon GB, Hoffman SC, *et al*., 2000, Serum dehydroepiandrosterone and dehydroepiandrosterone sulfate and the subsequent risk of developing colon cancer. *Cancer Epidemiol Biomarkers Prev*, 9(5): 517–521.

32. Hundt S, Haug U, Brenner H, 2007, Blood markers for early detection of colorectal cancer: A systematic review. *Cancer Epidemiol Biomarkers Prev*, 16(10): 1935–1953.

https://doi.org/10.1158/1055-9965.EPI-06-0994

33. Locker GY, Hamilton S, Harris J, *et al*., 2006, ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol*, 24(33): 5313–5327.

https://doi.org/10.1200/JCO.2006.08.2644

34. Hyde Z, Flicker L, McCaul KA, *et al*., 2012, Associations between testosterone levels and incident prostate, lung, and colorectal cancer. A population-based study. *Cancer Epidemiol Biomarkers Prev*, 21(8): 1319–1329.

https://doi.org/10.1158/1055-9965.EPI-12-0129

35. Chakraborty S, Pramanik J, Mahata B, 2021, Revisiting steroidogenesis and its role in immune regulation with the advanced tools and technologies. *Genes Immun*, 22(3): 125–140.

https://doi.org/10.1038/s41435-021-00139-3

36. Lee HK, Lee JK, Cho B, 2013, The role of androgen in the adipose tissue of males. *World J Mens Health*, 31(2): 136–140.

https://doi.org/10.5534/wjmh.2013.31.2.136

37. Chen E, Xu X, Liu T, 2018, Hereditary nonpolyposis colorectal cancer and cancer syndromes: Recent basic and clinical discoveries. *J Oncol*, 2018: 3979135.

https://doi.org/10.1155/2018/3979135

38. Li H, Pham T, McWhinney BC, *et al*., 2016, Sex hormone binding globulin modifies testosterone action and metabolism in prostate cancer cells. *Int J Endocrinol*, 2016: 6437585.

https://doi.org/10.1155/2016/6437585

39. Roshni K, Younis M, Ilakkiyapavai D, *et al*., 2018, Anticancer activity of biosynthesized silver nanoparticles using *Murraya koenigii* leaf extract against HT-29 colon cancer cell line. *J Cancer Sci Ther*, 10(4): 72–75.

https://doi.org/10.4172/1948-5956.1000521

40. Saddozai UA, Wang F, Akbar MU, *et al*., 2021, Identification of clinical relevant molecular subtypes of pheochromocytoma. *Front Endocrinol* (*Lausanne*), 12: 605797.

https://doi.org/10.3389/fendo.2021.605797

41. Saddozai UA, Wang F, Cheng Y, *et al*., 2020, Gene expression profile identifies distinct molecular subtypes and potential therapeutic genes in Merkel cell carcinoma. *Transl Oncol*, 13(11): 100816.

https://doi.org/10.1016/j.tranon.2020.100816

42. Akbar MU, Badar M, Zaheer M, 2022, Programmable drug release from a dual-stimuli responsive magnetic metal-

organic framework. *ACS Omega*, 7(36): 32588–32598.

https://doi.org/10.1021/acsomega.2c04144

43. Sawicki T, Ruszkowska M, Danielewicz A, *et al*., 2021, A review of colorectal cancer in terms of epidemiology, risk factors, development, symptoms and diagnosis. *Cancers* (*Basel*), 13(9): 2025.

https://doi.org/10.3390/cancers13092025

44. Saddozai UA, Wang F, Khattak S, *et al*., 2022, Define the two molecular subtypes of epithelioid malignant pleural mesothelioma. *Cells*, 11(18): 2924.

https://doi.org/10.3390/cells11182924

45. Baraibar I, Ros J, Saoudi N, *et al*., 2023, Sex and gender perspectives in colorectal cancer. *ESMO Open*, 8(2): 101204.

https://doi.org/10.1016/j.esmoop.2023.101204

ACCSCIENCE PUBLISHING

Gene & Protein in Disease

ORIGINAL RESEARCH ARTICLE

Identification of hotspots in synthetic peptide inhibitors of the FOXO4:p53 interaction

Ran Zhang[1], Kai Gao[1], Afsaneh Sadremomtaz[1,2,3], Angel J. Ruiz-Moreno[1], Alessandra Monti[4], Zayana M. Al-Dahmani[1], Benjamin B. Gyau[1], Nunzianna Doti[4], and Matthew R. Groves[1]*

[1]XB20 Drug Design, Groningen Research Institute of Pharmacy, University of Groningen, AD Groningen, The Netherlands

[2]Department of Nanoengineering, North Carolina Agricultural and Technical State University, Greensboro, North Carolina, USA

[3]Department of Nanoengineering, Joint School of Nanoscience and Nanoengineering, North Carolina Agricultural and Technical State University, Greensboro, North Carolina, USA

[4]Institute of Biostructures and Bioimaging (IBB), National Council of Research, Napoli, Italy

Abstract

Forkhead box protein O4 (FOXO4) plays a pivotal role in cellular senescence by binding to and inactivating p53. Consequently, misregulation of the FOXO4:p53 complex is associated with numerous diseases. Targeting the FOXO4-p53 interface has been achieved using a synthetic D-retro-inverso (DRI) peptide derived from the forkhead-homology domain of FOXO4 (FOXO4-FDH), also known as DRI (FOXO4-FHD residues 91–124). However, a comprehensive understanding of the key amino acids driving the interaction between DRI and p53 remains incomplete. While previous publications have demonstrated a robust interaction between the forkhead homology domain of FOXO4 (FOXO4-FHD) and the transactivation domain of p53 (p53-TAD), emerging evidence suggests that the interaction within the binary complex forms a highly interconnected network, including a predicted interaction between FOXO4-FHD and the DNA-binding domain of p53 (p53-DBD). In this study, we investigated the DRI: p53-DBD interaction by measuring the binding affinities of DRI and the native peptide of FOXO4, from which it is derived, to p53-DBD using microscale thermophoresis and computational modeling. Our *in vitro* measurements reveal that DRI binds to p53-DBD with high affinity (K_d ~50 nM), while the native peptide exhibits significantly weaker binding affinity (K_d ~2.5 mM), implying distinct modes of interaction. Subsequently, we created an *in silico* model of the DRI: p53-DBD interaction, which we analyzed to identify potential interaction hotspots. The analysis of this model suggests that a truncated DRI peptide (FOXO4-FHD amino acids 101–109) retains the majority of the binding affinity, as subsequently demonstrated *in vitro* (K_d ~40 nM). Collectively, this data furnishes molecular-level insights that contribute to the understanding of the interplay of the amino acids between DRI and p53, further supporting the notion of domain rearrangement or refolding during the formation of the FOXO4:p53 complex. In addition, this data provides an additional basis for the design of small molecules aimed at inhibiting this interaction.

Keywords: FOXO4-DRI: p53 interaction; Protein–protein interaction; Computational docking; Microscale thermophoresis

***Corresponding author:**
Matthew R. Groves
(m.r.groves@rug.nl)

Citation: Zhang R, Gao K, Sadremomtaz A, *et al.*, 2023, Identification of hotspots in synthetic peptide inhibitors of the FOXO4:p53 interaction. *Gene Protein Dis,* 2(3): 1491.
https://doi.org/10.36922/gpd.1491

Received: August 10, 2023
Accepted: September 11, 2023
Published Online: September 29, 2023

1. Introduction

Cell senescence refers to the phenomenon of cell cycle arrest that occurs in response to cellular stress. This arrest is accompanied by halted cell growth, an upregulated expression of cytokines, and an enhanced secretory capacity[1]. This phenomenon is also known as the senescence-associated secretory phenotype (SASP), which is driven by various families of factors, including interleukins, chemokines, growth factors, proteases, and their regulators[2]. The effects of SASP are two sided. On the one hand, it can inhibit tumor growth and promote the restoration and regeneration of damaged tissues. Conversely, the presence of abnormal cytokines can accelerate the aging process, contribute to tumorigenesis, and facilitate tumor metastasis[3,4]. Therefore, cell senescence represents an important risk factor for cancer, as the accumulation of mutations and inflammation caused by senescence can induce oncogenesis[1,5].

p53 is a critical tumor suppressor that plays an important role in tumorigenesis and development[6]. The human p53 protein consists of 393 amino acids. It contains five domains: the N-terminal transactivation domain, a proline-rich domain, the DNA-binding domain (FOXO4-DBD), the tetramerization region, and a carboxy-terminal regulatory domain[7]. p53 becomes activated in response to various intracellular and extracellular signals, exerting its regulatory influence on tumor development through DNA damage repair, cell cycle arrest, apoptosis, senescence, and cellular stress responses[8,9]. Functionally, when p53 is inactive, it not only loses its cancer-suppressing function but also contributes to the promotion of carcinogenesis[10].

Research has shown that FOXO4 is upregulated in senescent cells but is only marginally expressed in most other tissues[11]. FOXO4 plays a crucial role in maintaining survival in senescent cells by repressing p53-induced apoptotic response[12]. DNA-SCARS, or DNA segments with chromatin alterations reinforcing senescence, is known to fuse with promyelocytic leukemia (PML) bodies. This interaction leads to the recruitment of p53, thereby forming a complex with FOXO4[13]. FOXO4 consists of several functional domains that contribute to its biological activities, including the forkhead domain (FH), nuclear localization signal (NLS), nuclear export sequence (NES), C-terminal transactivation domains (TRD), and conserved regions (CR).

Baar *et al.* reported the development of a synthetic D-retro-inverso (DRI) FOXO4-based peptide known as FOXO4-DRI, which is composed of D-amino acids arranged in a reversed sequence[11] and corresponds to residues 93–119 of the FOXO4-FHD. This peptide has the ability to bind to p53 and disrupt the FOXO4-p53 interaction. This disruption leads to the disorganization of the FOXO4-PML-DNA-SCARS complex, releasing p53 from this complex and making it available to trigger p53-mediated apoptosis. The synthetic peptide FOXO4-DRI can thereby reduce senescence and features of frailty in a fast-aged mice model (XpdTTD/TTD). In addition, experiments also show that FOXO4-DRI can restore the loss of renal function in both naturally and fast-aged mice. As a result, the inhibition of protein–protein interaction (PPI) by FOXO4-DRI exhibits promising potential in the treatment of senescence-related diseases[11].

However, when compared with small molecules, peptides have certain disadvantages, including limited membrane-penetrating ability, poor *in vivo* stability, low metabolic stability, and reduced oral availability[13,14]. In addition, the production cost and market price of peptide-based drugs tend to be higher than those of small molecules. These limitations have constrained the development and utilization of peptide-based drugs. However, therapeutic peptides feature unique biophysical properties compared to small molecules. Their larger size makes it easier to inhibit PPIs, and their higher specificity can help avoid side effects stemming from off-target effects[14]. Given the demonstrated strong potential of FOXO4-DRI, the development of a small molecule mimic targeting the FOXO4-DRI: p53 interaction could provide a preferable route to antagonizing the FOXO4:p53 interaction. The design of such a small molecule would benefit from an improved understanding of the DRI: p53 interaction.

While the interaction between FOXO4-DRI and p53 holds significant promise as a therapeutic target, there is currently no direct evidence regarding the molecular interaction and binding site on p53 for the DRI. Mandal *et al.* reported that the interaction between two domains, p53-TAD (transactivation domain) and FOXO4-FHD (forkhead domain), plays a crucial role in the formation of the FOXO4-p53 complex[15]. Le *et al.* showed that a designed peptide (with a K_d value of 0.66 nM) can disrupt the FOXO4-p53 interaction by preferentially binding to FOXO4[16]. Kim *et al.* demonstrated that the binding between two pairs of protein domains, FOXO4-FHD: p53-TAD (with a K_d value of 2.19 μM) and FOXO4-CR3:p53-DBD (with a Kd value of 6.96 μM), mediates the FOXO4-p53 interaction[17]. In addition, the N-terminal region of FOXO4, the CRD (C-terminal regulatory domain) of p53, and the DBD (DNA-binding domain) of both proteins are involved in stabilizing the FOXO4-p53 complex[15]. Given that the DRI is a synthetic D-amino acid retro-inverse sequence of residues 93–119 of FOXO4, it is highly likely to possess a distinct molecular shape and charge distribution from that of the L-amino acids of the

same sequence in wild-type FOXO4. This suggests that the interactions of the DRI may not be similar to those of the equivalent FOXO4 sequence in the wild-type FOXO4:p53 complex. These differences in molecular features are likely to strongly impact the interactions between the DRI and its p53 target.

In this study, we explored whether the differences in binding energy could be determined between the DRI and its L-amino peptide counterpart in relation to the p53-DBD. In addition, we present the use of an *in silico* model for the DRI: FOXO4 interaction, which generated data on the likely key interactions between FOXO4 and p53-DBD. This model aided in the design of shorter peptides to test hypotheses generated from the model. Subsequent *in vitro* experiments demonstrated that the newly developed shorter peptides contain the key interaction elements, as they retain almost complete binding affinity of the DRI to p53-DBD. These experiments provide data to aid in the design of smaller molecules (cyclic peptides or synthetic molecules) aimed at improving the pharmacological properties of inhibitors targeting the p53:FOXO4 PPI.

2. Materials and methods

2.1. Protein expression and purification

The p53-DBD gene was cloned into the pETM-11[18] expression vector using the NcoI and HindIII restriction sites and confirmed by sequencing. A single colony was inoculated into 10 ml LB culture medium containing 50 μg/ml kanamycin and 35 μg/ml chloramphenicol. This culture was incubated overnight at 37°C with shaking at 180 rpm. Subsequently, the 10 ml overnight culture was transferred into 1 L of LB medium supplemented with 50 μg/ml kanamycin and 35 μg/ml chloramphenicol. The culture was maintained at 37°C with continuous shaking at 180 rpm until the OD600 reached approximately 0.6. At this point, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM, and the culture was further incubated at 18°C for 16 h. Cells were harvested by centrifuging the culture at 6000 rpm for 30 min at 4°C and then resuspended in lysis buffer (100 mM HEPES at pH 7.5, 100 mM NaCl, and 5 mM β-mercaptoethanol). The cells were sonicated in an ice-water bath for 6 min. The supernatant was collected through high-speed centrifugation at 18000 rpm for 45 min at 4°C and subsequently loaded onto a 5 ml HisTrap HP (Cytiva) column that had been pre-equilibrated with lysis buffer. The column was then washed with 10 column volumes of washing buffer (100 mM HEPES at pH 7.5, 100 mM NaCl, 5 mM β-mercaptoethanol, and 25 mM imidazole). Protein was eluted using a linear gradient with an elution buffer (100 mM HEPES at pH 7.5, 100 mM NaCl, 5mM β-mercaptoethanol, and 500 mM imidazole). Fractions containing the target protein were identified using SDS-PAGE and subsequently pooled and concentrated using an Amicon Ultra-15 concentration unit. Further purification was carried out using a Superdex 75 (GE Healthcare) column with SEC buffer (20 mM Bis-Tris at pH 7.8, 150 mM NaCl, and 10 mM β-mercaptoethanol). Fractions containing the protein of interest were once again identified using SDS-PAGE, pooled, and concentrated to a final concentration of 10 mg/ml. The samples were then flash-frozen in liquid nitrogen and stored at −80°C until further use.

2.2. Peptides synthesis and characterization

All peptides (Table 1) were synthesized using solid-phase peptide synthesis, following the Fmoc (N-9- fluorenylmethyloxycarbonyl) strategy as described in the literature[19,20]. The synthesis was carried out on Wang resin (loading at 0.3 mmol/g for long peptides and 1.2 mmol/g for short ones) with a fully automated peptide synthesizer, Syro I from MultiSynTech GmbH (Witten, Germany). Each peptide was cleaved from the solid support using a mixture composed of 90% trifluoroacetic acid (TFA), 5% triisopropylsilane (TIS), and 5% H2O (v/v/v) and stirred for about 3 h at room temperature. The resulting peptides were precipitated in cold diethyl ether, and the pellets were resuspended in a mixture of water and acetonitrile (H_2O/ACN = 75:25 v/v) and then lyophilized. Peptide purification was carried out using a NUCLEODUR HTec C18 column (5 μm, 150×21 mm) at a flow rate of 12 ml/min. Elution was carried out using a gradient of solvent B (CH_3CN + 0.1% TFA) from 5% to 70% over 20 min, with solvent A (H2O+0.1% TFA) as the other component. The purity and identity of the peptides were verified using analytical liquid chromatography with mass spectrometry (LC-MS), and the results are presented in Table 1. LC-MS analysis of peptides was carried out using a C18 Waters XBridge column (3 μm, 4.6 × 5.0 mm). A linear gradient of 0.05% TFA in CH_3CN from 5 to 70% over 15 min was applied at a flow rate of 0.2 ml/min. The purity of all peptides exceeded 95% and determined by calculating the ratio of the peak area of the target peptide to the sum of all areas in UV chromatograms at 210 nm.

2.3. Microscale thermophoresis

The binding affinity of peptides toward p53 was analyzed using the microscale thermophoresis (MST) technique. For the MST experiments, purified p53 protein was diluted to a concentration of 50 nM and labeled using the His-Tag Labeling Kit RED-tris-NTA (Monolith Cat#-L018), following the manufacturer's recommended instructions. The labeled protein was pretested using Monolith™ NT.115 MST

Table 1. Amino acid sequences and theoretical and experimental m/z (monoisotopic) of peptides tested in this work

Entry	Sequence	Theoretical (MW)	Experimental (MW)
FOXO4-DRI	NH_2-tlrkepaseiaqsileaysqngwanrr-OH	3087.54	3087.546
FOXO4-DRI_short	NH_2-leaysqngw-OH	1066.44	1066.461
FOXO4	NH_2-RRNAWGNQSYAELISQAIESAPEKRLT-OH	3087.54	3087.546
FOXO4_short	NH_2-WGNQSYAEL-OH	1066.44	1066.461

Notes: All peptides are free at N-terminus and COOH at C-terminus. D-amino acids are shown in lowercase to distinguish them from L-amino acids (uppercase).

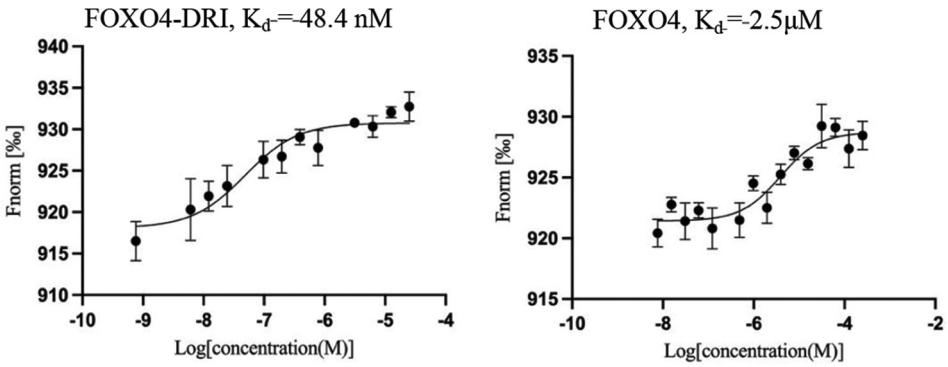


Figure 1. Microscale thermophoresis experiments were conducted to assess the binding affinities of peptides FOXO4 and FOXO4-DRI to the DNA binding domain of p53 (p53-DBD). Under identical conditions, the FOXO4-DRI binds with a binding affinity (K_d) of ~50 nM (A), whereas the equivalent wild-type peptide shows a binding affinity (K_d) of 2.5 mM (B).

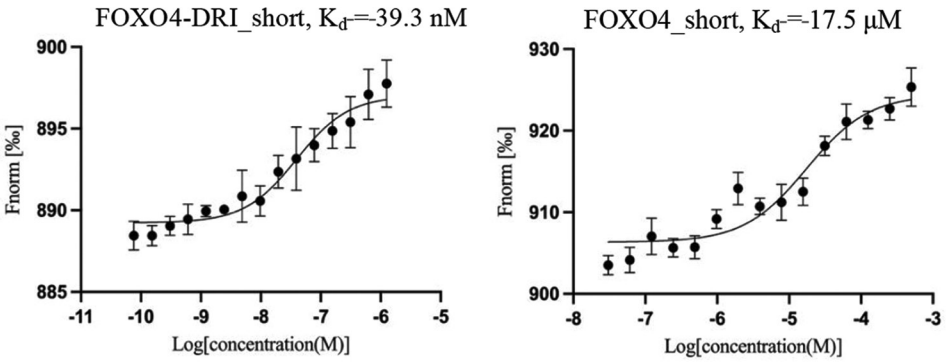


Figure 2. A microscale thermophoresis experiments to assess the binding affinities of truncated peptides FOXO4_short and FOXO4-DRI-short to the DNA binding domain of p53 (p53-DBD). (A) Under identical conditions to those shown in Figure 1, the FOXO4-DR_short binds with a measured K_d of ~50 nM, similar to that seen for FOXO4-DRI. (B) The equivalent shorter wild-type peptide (FOXO4_short) shows a minor reduction in binding affinity of ~20 mM when compared with that observed in FOXO4 in Figure 1.

Standard Capillaries (Monolith Capillaries Cat#-MO-K022) on a Monolith NT.115[21]. The peptides were prepared in PBS-T buffer (PBS pH7.4, 0.05% [v/v] Tween 20) to achieve a final concentration of 2.5 µM. A peptide serial dilution was prepared by performing 16 consecutive 2-fold dilutions. Subsequently, 10 µl of the labeled protein sample was added to 10 µl of peptide serial dilution, thoroughly mixed, and incubated at room temperature for 20 min. The mixture was then centrifuged at 13000 rpm for 15 min. The resulting samples were loaded into capillary trays for data measurement using 40% excitation power and medium MST power at 25°C. These experiments were conducted in triplicate, and the data were analyzed using the manufacturer's provided software. Errors in individual measurements used to calculate the binding affinities are shown in Figures 1 and 2.

2.4. Protein-protein docking

Protein–protein docking was performed using HADDOCK[22]. The crystal structure of a monomer of the p53 DBD was used as a receptor (PDB:3Q05), while

the crystal structure of the DBD of FOXO4 (PDB:3L2C) was employed as the ligand[23,24]. For both receptor and ligand, DNA, solvent molecules, and ions were removed, and hydrogens were added to all residues. For docking, the chosen binding site on the receptor encompassed residues 94–148 and 238–327, while for the ligand, residues corresponding to FOXO4-DRI (comprising residues 93–119 of FOXO4 in 3L2C) were selected to form the PPI. The p53-FOXO4 pose with the best complementarity between receptor and ligand was chosen for further analysis. Following docking, the structural flexibility and large-scale conformational transitions of the p53-FOXO4 complex were evaluated using the CABS-flex 2.0 server[25].

The docking structure of p53 complexed with FOXO4 was used to generate binding modes for FOXO4-based peptides. First, secondary structures for the peptides were built using Chimera X's structure building tool[26], ensuring proper torsions between amino acids by referencing the dynameomics rotamer library. The resulting structures were aligned with the corresponding peptide sequence in the FOXO4 DBD, considering its interaction with p53. Finally, the p53-peptide complexes and the FOXO4:p53 complex were refined through an analysis of structural flexibility and large-scale conformational transition using the CABS-flex 2.0 server[25].

3. Results

3.1. DRI binds significantly more tightly to p53-DBD than its L-amino acid equivalent

Synthetic peptides corresponding to the published sequence of the DRI (FOXO4-DRI) and the equivalent wild-type sequence (FOXO4) were generated as previously described. The binding affinity of these peptides against purified p53-DBD was measured using MST, as detailed earlier, generating experimental measurements of their binding affinity. The K_d values for the peptides "FOXO4" and its DRI equivalent "FOXO4-DRI" are 2.5 μM and 48.4 nM, respectively (Figure 1). The significant difference in K_d values between the DRI and native peptides under identical conditions provides further evidence that FOXO4-DRI interacts with p53-DBD with high affinity. The apparent strength of this interaction forms the basis for the ability of FOXO4-DRI to effectively compete with the relatively weaker interactions observed between FOXO4-FHD: p53-TAD (2.19 mM), FOXO4-CR3:p53-DBD (6.96 μM), the N-terminal region of FOXO4:p53-CRD, and FOXO4-DBD: p53-DBD. In contrast, the equivalent wild-type peptide (FOX04) exhibits weaker binding affinity to p53-DBD, thus displaying reduced competitiveness within the overall FOXO4:p53 complexation.

3.2. Modeling of the interaction between DRI and p53-DBD

Based on the measured binding affinities between the DRI and the equivalent L-amino acids of FOXO4, we continued with the identification of potential molecular interaction sites. We employed computational docking to generate models of the DRI: p53-DBD interaction, utilizing the p53 DBD (PDB:3Q05) and FOXO4 DBD (PDB:3L2C) as our initial models (Figure 3). These models facilitated the identification of a putative binding surface capable of accommodating both FOXO4 and FOXO4-DRI (Figure 3A and B). However, the predicted conformations of these two peptides exhibited significant differences, resulting in associated differences in predicted binding energies – the wild-type peptide corresponding to the DRI exhibited a predicted ΔG of binding of -6.4 kcal/mol, while the DRI displayed a correspondingly higher ΔG of binding at -7.1 kcal/mol. In addition, our computational modeling of the FOXO4-DRI synthetic peptide interacting with the p53 surface yielded a rearranged conformation when compared to the wild-type sequence (as expected from NMR analysis of the FOXO4:p53 interaction[21]) (Figure 3C). Upon closer examination (Figure 4), two significant pockets were identified on the surface of p53-DBD and predicted to bind both peptides. Pocket 1 (composed of amino acids Q7, K71, M76, S176, and E178) was shallower and less extended than Pocket 2 (composed of amino acids R17, L21, P35, A36, L37, N38, and S176) (Figure 4).

The two pockets together constitute the predicted core zone of the binding site for the DRI. S176 and N38 are located between the pockets, forming a barrier that separates the two pockets into two independent regions. This amino

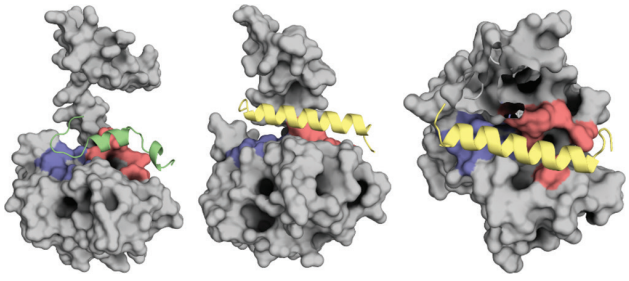


Figure 3. Molecular models of the FOXO4:p53-DBD and FOXO4-DRI: p53-DBD complexes. (A) Predicted binding mode of the wild-type FOXO4 peptide that corresponds to the FOXO4-DRI with p53-DBD (peptide in green, ΔG: −6.4 kcal/mol, K_d: 2.2×10^{-5} kcal/mol). The computational model suggests that the FOXO4 peptide interacts with 2 pockets on the surface of p53-DBD: Pocket 1 (blue) and Pocket 2 (red). (B) The predicted binding mode of the FOXO4-DRI (yellow, ΔG: −7.1 kcal/mol, K_d: 5.8×10^{-6} kcal/mol) suggests significant rearrangement of the FOXO4-DRI but maintains interactions with both pockets 1 and 2. (C) A top view of the predicted binding mode of the FOXO4-DRI shown in (B).

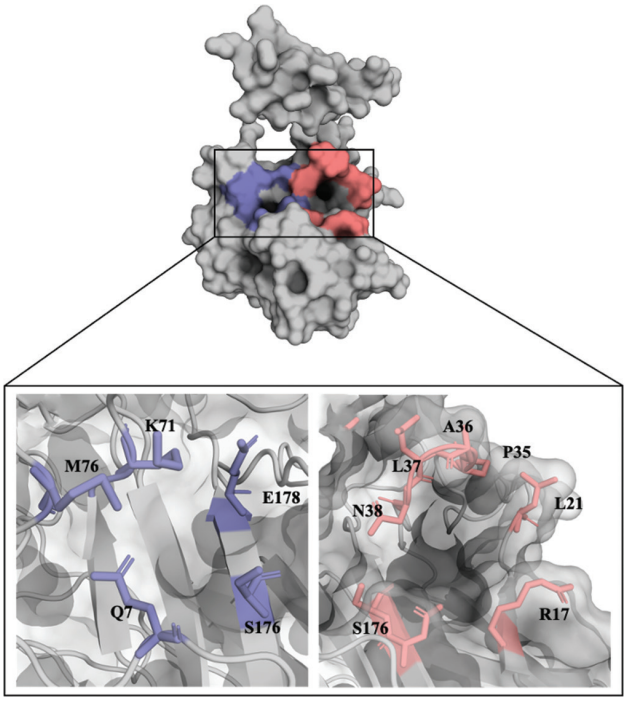


Figure 4. Analysis of the 2 predicted binding pockets of FOXO4-DRI. (A) Surface view of Pocket 1 (blue) and Pocket 2 (red) on the suggested FOXO4-DRI: p53-DBD interface. (B and C) Key amino acids in the binding sites are highlighted in stick representation. Analysis indicates that Pocket 2 is deeper and more extended than Pocket 1, suggesting it may drive the interaction between the peptides and p53-DBD.

acid "barrier" could potentially impact the interaction of a long peptide with p53, particularly if the peptide needs to interact across the two pockets and lacks sufficient internal flexibility to address both pockets simultaneously. In contrast, shorter peptides could potentially be shielded from this effect by maintaining strong binding interactions without incurring the potential entropic costs associated with a loss of solution flexibility upon binding. This suggests that a shorter peptide addressing only Pocket 2 may retain significant binding affinity and be easier to mimic using small molecule approaches.

3.3. A truncated DRI retains a strong affinity with p53-DBD

According to the above information and the known FOXO4:p53 interaction[11], two shorter peptides were synthesized. Peptide "FOXO4 short" corresponds to the region of FOXO4 that interacts with Pocket 2, composed of L-amino acids in the native ordering, whereas peptide "FOXO4-DRI short" corresponds to the DRI form of this peptide.

MST assays were performed with a serial peptide concentration from 0.25 mM to 7.6 nM while maintaining a p53 working concentration of 25 nM. The K_d values for FOXO4 short and FOXO4-DRI short were determined to be 17.5 μM and 39.3 nM, respectively (Figure 2). Interestingly, the majority of the binding *in vitro* is retained when truncating the 27 L-amino acids of FOXO4-DRI (Figure 4A) to the 9 amino acids of FOXO4-DRI_short (Figure 2A). This observation suggests that the majority of the interactions between the DRI peptide and p53-DBD occur with Pocket 2 (Figures 1B and 2B) of p53-DBD. A similar effect is observed in the K_d values of FOXO4 and FOXO4_short, where the majority of the binding interaction is preserved upon truncation to the 9 amino acids of FOXO4_short. This retention of binding affinity may be attributed to the insufficient binding of FOXO4-DRI to p53, likely due to reduced flexibility in the longer helix, as suggested by molecular modeling (Figure 3B and C). As depicted in Figure 3A, the conformation of bound FOXO4, composed of L-amino acids, is predicted to consist of an α helix and a loop, suggesting good flexibility. In contrast, the FOXO4-DRI (Figure 3B and C), composed of D-amino acids, retains an overall helical conformation – suggesting a lower entropic penalty upon binding to p53 than that displayed in its L-amino acid equivalent.

4. Discussion

The FOXO4:p53 complex plays a crucial role in maintaining the vitality of senescent cells. Baar *et al.* demonstrated that a DRI peptide could disrupt the FOXO4:p53 interaction, releasing p53 and triggering p53-induced apoptosis[11]. In this study, we assessed the ability of the DRI to interact with p53-DBD and used computer simulations to predict the FOXO4-DRI: p53-DBD interaction. These simulations suggested that the DRI peptide may interact with two pockets on the surface of p53-DBD, termed Pocket 1 and Pocket 2. These models guided the design of a truncated DRI peptide to test the hypothesis that interaction with Pocket 2 is the driving factor in this proposed model. The findings of this study indicate that a truncated D-retro-inverso analog of the FOXO4-DRI peptide exhibits significant binding affinity toward p53-DBD. Compared to native L-peptides, the D-retro-inverso peptide exhibits both enhanced biological stability and stronger binding affinity, resulting in significantly different biological activity[27].

Our biophysical experiments have confirmed our hypothesis that a truncated FOXO4 peptide would retain most of the binding features exhibited by the longer FOXO4-DRI. Notably, there is no significant loss in binding when the peptide is truncated from 27 to 9 amino acids. Our model suggests that this shorter region retains interactions with Pocket 2, and we are currently conducting X-ray soaking experiments using p53 crystals to validate this conformation. Furthermore, FOXO4-DRI_short

exhibits a slight improvement in activity when compared to FOXO4-DRI. It is possible that the longer FOXO4-DRI peptide, while capable of more interactions with the p53 surface, may be impacted by the need for increased peptide flexibility during binding to p53. In addition, the substitution of D-amino acids may allow the peptide chain to adopt different folding structures distinct from those formed by L-amino acid peptides, further promoting their binding to the active pocket.

The binding affinities suggest that the configuration, order difference, and flexibility of the peptide can significantly affect peptide affinity. This finding further supports the potential of truncated peptide design, based on conserved hot spots, in combination with DRI approaches as a promising strategy for addressing other PPI surfaces. Moreover, our predicted model, together with the activity assays, strongly indicates that Pocket 1 and Pocket 2 are the hot spot regions of the FOXO4:p53 interface. This insight provides a potential mechanism to support the design of further modulators targeting the FOXO4-DRI: p53 interaction.

5. Conclusion

Compared to the L peptide, the DRI exhibits a significantly higher *in vitro* binding affinity to p53-DBD, strongly suggesting distinct binding modes between the two species and p53-DBD. Computational calculations support this observation. A truncated form of DRI contains the crucial region for specifically targeting the hot spot regions, as the shorter peptide retains almost full binding affinity to p53-DBD. While retaining binding affinity, the potential to significantly reduce the size of FOXO4-DRI provides a basis for designing small molecule mimics of DRI, which may lead to improvements in drug-like behavior.

Acknowledgments

None.

Funding

The authors acknowledge financial support from the Chinese Scholarship Council (RZ and KG).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: All authors
Investigation: Ran Zhang, Kai Gao, Afsaneh Sadremomtaz, Angel Ruiz-Moreno, Alessandra Monti, Zayana Al-Dahmani, Benjamin B. Gyau
Formal analysis: Ran Zhang, Kai Gao, Afsaneh Sadremomtaz, Angel Ruiz-Moreno, Alessandra Monti, Zayana Al-Dahmani, Benjamin B. Gyau
Writing – original draft: Ran Zhang
Writing – review & editing: All authors

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Raw data are available on request from the corresponding author.

References

1. Muñoz-Espín D, Cañamero M, Maraver A, *et al*., 2013, Programmed cell senescence during mammalian embryonic development. *Cell*, 155(5): 1104–1118.

   https://doi.org/10.1016/j.cell.2013.10.019

2. Coppé JP, Desprez PY, Krtolica A, *et al*., 2010, The senescence-associated secretory phenotype: The dark side of tumor suppression. *Annu Rev Pathol*, 5: 99–118.

   https://doi.org/10.1146/annurev-pathol-121808-102144

3. Young AR, Narita M, 2009, SASP reflects senescence. *EMBO Rep*, 10(3): 228–230.

   https://doi.org/10.1038/embor.2009.22

4. Birch J, Gil J, 2020, Senescence and the SASP: Many therapeutic avenues. *Genes Dev*, 34(23–24): 1565–1576.

   https://doi.org/10.1101/gad.343129.120

5. Campisi J, 2000, Cancer, aging and cellular senescence. *In Vivo*, 14(1): 183–188.

6. Marei HE, Althani A, Afifi N, *et al*., 2021, p53 signaling in cancer progression and therapy. *Cancer Cell Int*, 21(1): 703.

   https://doi.org/10.1186/s12935-021-02396-8

7. Joerger AC, Fersht AR, 2010, The tumor suppressor p53: From structures to drug discovery. *Cold Spring Harb Perspect Biol*, 2(6): a000919.

   https://doi.org/10.1101/cshperspect.a000919

8. Aubrey BJ, Kelly GL, Janic A, *et al*., 2018, How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? *Cell Death Differ*, 25: 104–113.

   https://doi.org/10.1038/cdd.2017.169

9. Fischer M, 2017, Census and evaluation of p53 target genes. *Oncogene*, 36(28): 3943–3956.

   https://doi.org/10.1038/onc.2016.502

10. Harris CC, 1996, p53 tumor suppressor gene: At the

crossroads of molecular carcinogenesis, molecular epidemiology, and cancer risk assessment. *Environ Health Perspect*, 104 Suppl 3(Suppl 3): 435–439.

https://doi.org/10.1289/ehp.96104s3435

11. Baar MP, Brandt RM, Putavet DA, *et al*., 2017, Targeted apoptosis of senescent cells restores tissue homeostasis in response to chemotoxicity and aging. *Cell*, 169(1): 132–147.e16.

https://doi.org/10.1016/j.cell.2017.02.031

12. Bourgeois B, Madl T, 2018, Regulation of cellular senescence via the FOXO4-p53 axis. *FEBS Lett*, 592(12): 2083–2097.

https://doi.org/10.1002/1873-3468.13057

13. Rodier F, Muñoz DP, Teachenor R, *et al*., 2011, DNA-SCARS: Distinct nuclear structures that sustain damage-induced senescence growth arrest and inflammatory cytokine secretion. *J Cell Sci*, 124(Pt 1): 68–81.

https://doi.org/10.1242/jcs.071340

14. Wang L, Wang N, Zhang W, *et al*., 2022, Therapeutic peptides: Current applications and future directions. *Signal Transduct Target Ther*, 7(1): 48.

https://doi.org/10.1038/s41392-022-00904-4

15. Mandal R, Kohoutova K, Petrvalska O, *et al*., 2022, FOXO4 interacts with p53 TAD and CRD and inhibits its binding to DNA. *Protein Sci*, 31(5): e4287.

https://doi.org/10.1002/pro.4287

16. Le HH, Cinaroglu SS, Manalo EC, *et al*., 2021, Molecular modelling of the FOXO4-TP53 interaction to design senolytic peptides for the elimination of senescent cancer cells. *EBioMedicine*, 73: 103646.

https://doi.org/10.1016/j.ebiom.2021.103646

17. Kim J, Ahn D, Park CJ, 2022, Biophysical investigation of the dual binding surfaces of human transcription factors FOXO4 and p53. *FEBS J*, 289(11): 3163–3182.

https://doi.org/10.1111/febs.16333

18. Dümmler A, Lawrence AM, de Marco A, 2005, Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb Cell Fact*, 4: 34.

https://doi.org/10.1186/1475-2859-4-34

19. Caporale A, Doti N, Sandomenico A, *et al*., 2017, Evaluation of combined use of Oxyma and HATU in aggregating peptide sequences. *J Pept Sci*, 23(4): 272–281.

https://doi.org/10.1002/psc.2977

20. Caporale A, Doti N, Monti A, *et al*., 2018, Automatic procedures for the synthesis of difficult peptides using oxyma as activating reagent: A comparative study on the use of bases and on different deprotection and agitation conditions. *Peptides*, 102: 38–46.

https://doi.org/10.1016/j.peptides.2018.02.006

21. Nanotempertech, 2021, Protein Labeling-improved Quantitation of Biomolecular Interactions by MST Using the His-tag Labeling Kit RED-tris-NTA 2nd Generation. Available from: https://resources.nanotempertech.com/technical-notes [Last accessed on 2023 Jun 20].

22. Van Zundert GC, Rodrigues JP, Trellet M, *et al*., 2016, The HADDOCK2.2 web server: User-friendly integrative modeling of biomolecular complexes. *J Mol Biol*, 428: 720–725.

https://doi.org/10.1016/j.jmb.2015.09.014

23. Petty TJ, Emamzadah S, Costantino L, *et al*., 2011, An induced fit mechanism regulates p53 DNA binding kinetics to confer sequence specificity. *EMBO J*, 30(11): 2167–2176.

https://doi.org/10.1038/emboj.2011.127

24. Boura E, Rezabkova L, Brynda J, *et al*., 2010, Structure of the human FOXO4-DBD-DNA complex at 1.9 Å resolution reveals new details of FOXO binding to the DNA. *Acta Crystallogr D Biol Crystallogr*, 66(Pt 12): 1351–1357.

https://doi.org/10.1107/S0907444910042228

25. Kuriata A, Gierut AM, Oleniecki T, *et al*., 2018, CABS-flex 2.0: A web server for fast simulations of flexibility of protein structures. *Nucleic Acids Res*, 46(W1): W338–W343.

https://doi.org/10.1093/nar/gky356

26. Pettersen EF, Goddard TD, Huang CC, *et al*., 2021, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*, 30(1): 70–82.

https://doi.org/10.1002/pro.3943

27. Sela M, Zisman E, 1997, Different roles of D-amino acids in immune phenomena. *FASEB J*, 11(6): 449–456.

https://doi.org/10.1096/fasebj.11.6.9194525

ACCSCIENCE PUBLISHING

Gene & Protein in Disease

CASE REPORT

Clinical debate on papillary thyroid microcarcinoma-could genetic testing change the decision of papillarythyroid microcarcinoma (mPTC) treatment: A case report

Haitao Peng[1,2†], Chang cai[1†], Zifeng Luo[3], Chong Wang[1], Yuanwei Luo[1], and Song Wang[1*]

[1]Department of Thyroid and Breast Surgery, Key Laboratory of Biological Targeting Diagnosis Therapy and Rehabilitation of Guangdong Higher Education Institutes, The Fifth Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong, China

[2]Department of Thyroid and Breast Surgery, Shantou Third Peoples Hospital, Shantou, Guangdong, China

[3]School of International Studies, Hunan Institute of Technology, Hengyang, Hunan, China

Abstract

The objective of this case report is to explore whether genetic testing is appropriate to guide the treatment of papillary thyroid microcarcinoma (PTMC). In this case report, we describe a 27-year-old female who had no significant medical history and no tumor detected in the thyroid gland by ultrasound or computed tomography (CT), but had multiple enlarged lymph nodes at the right cervical lymph node levels II, III, IV, and V. The intraoperative frozen section pathology results showed that seven out of 23 right lymph nodes were metastatic from thyroid papillary carcinoma, but no primary lesion was found in the thyroid gland. After multiple post-operative sampling and micro section, only a 1 mm papillary carcinoma in the left thyroid gland was identified. The genetic testing result showed that the patient was positive for CCDC6-exon 1-RET-exon 12 fusion. The final diagnosis of the patient was left thyroid micro-papillary carcinoma (1 mm) with the right lateral cervical lymph node metastasis. She received iodine-131 treatment with nuclide and lifelong levothyroxine therapy in the 3rd month after surgery. Continuous follow-up showed no elevation of thyroglobulin, no recurrence signs on positron emission tomography-CT re-examination, and good self-reported health status. In conclusion, the identification of CCDC6-RET fusion by genetic testing in the PTMC case aided diagnosis and treatment.

Keywords: Papillary thyroid microcarcinoma; Thyroid cancer; Active surveillance; CCDC6-RET fusion; Genetic testing

†These authors contribute equally to this work.

***Corresponding author:**
Song Wang
(gzwangmomo1983@sina.com)

Citation: Peng H, Cai C, Luo Z, *et al*., 2023, Clinical debate on papillary thyroid microcarcinoma could genetic testing change the decision of papillary thyroid microcarcinoma (mPTC) treatment: A case report. *Gene Protein Dis,* 2(3): 0371.
https://doi.org/10.36922/gpd.0371

Received: March 28, 2023
Accepted: August 9, 2023
Published Online: August 30, 2023

1. Introduction

Papillary thyroid carcinoma (PTC), which accounts for more than half of the thyroid microcarcinoma (PTMC) cases, has a diameter of ≤1 cm[1]. Considering the inert nature of all PTMCs, PTMC is usually slow-growing, with uncomplicated clinical presentation, and rarely endangers the patient's life. In recent years, PTMC metastasis to lymph

nodes, uterus, chest wall, or skeleton has been frequently reported[2-4]. We need to pay attention to these tumors, but there is no good way to distinguish and treat them. Active surveillance (AS) is recommended as a treatment of low-risk PTMC, as it is deemed comparable to surgical intervention. PTMCs with lymph node or distant metastasis, thyroid extrusion, and cytological indications are all amenable to the AS approach[5]. The case investigated in this study was a left papillary thyroid microcarcinoma with multiple right cervical lymph node metastases. We also aimed to explore the value of genetic testing in deciding the appropriate treatment, that is, surgery and AS, for PTMC.

2. Case presentation

We present the case of a 27-year-old female patient with no previous medical history who was referred to the Department of Thyroid and Breast Surgery, The Fifth Affiliated Hospital of Guangzhou Medical University (Guangzhou, Guangdong, China). This patient had a large mass in area II of her right neck. A fine-needle aspiration biopsy confirmed that it was a metastatic lesion of PTC. However, no tumor was detected in the thyroid gland by routine ultrasound and computed tomography (CT) examinations. Multiple enlarged lymph nodes were found and distributed in the right cervical lymph node levels (Figure 1A-C). Intraoperative pathological diagnosis revealed that seven out of 23 right lymph nodes were metastatic from PTC, but no obvious cancer lesion was seen in the thyroid gland (Figure 1D-F). After multiple postoperative sampling and micro section analysis, a 1 mm papillary carcinoma in the left thyroid gland was discovered (Figure 1G and H). To further investigate, a modern genetic testing approach was applied, which showed that the patient was positive for CCDC6-exon 1-RET-exon 12 fusion (Figure 2), a driver of thyroid cancer. Other genes tested, such as *BRAFV600E*, *KRAS*, *NTRK1*, *NTRK2*, *NTRK3*, *EGFR*, *PIK3CA*, *ALK*, *NRAS*, *BRCA1*, *BRCA2*, *ROS1*, *PDGFRA*, *MET*, *HER2*, *NRAS*, and *KIT*, were negative. The patient underwent post-operative radioactive iodine treatment and thyroid-stimulating hormone suppression.

3. Discussion

Due to the advanced ultrasound used in diagnosis, the incidence of PTC has been rapidly increasing in many countries since the last decade of the 20th century[6]. According to the American Thyroid Association Guidelines, AS is the best clinical management for patients with PTMC, which shows no evidence of extrathyroidal extension[7]. Previously, Sugitani *et al.* reported prospective clinical trials of AS for low-risk PTMC conducted at two Japanese centers since the 1990s[8]. Factors such as age, family history, vascular distribution, imaging features, desire to have children, and pregnancy were assessed. The results showed that 8% of the patients had a tumor increase of ≥3 mm, while 3.8% of the patients had lymph node metastasis during the 10-year AS. The incidence of adverse events, such as temporary vocal cord paralysis and temporary or permanent parathyroidism in the immediate surgery group, was significantly higher than in the AS group. Therefore, the authors strongly recommend AS as the best option of clinical management for low-risk PTMC[8].

The concept of AS in the management of malignancies has stimulated intense clinical debate. AS was regarded as a long-term management strategy for patients, if clinical support systems are in place, and compliance with observation indications[9].The aggressiveness of PTMC and the deteriorating clinical status of patients should be considered potential challenges in AS, even if thyroidectomy could help reduce the risks[10]. Al-Qurayshi *et al.*[11] retrospectively evaluated the risk of PTMC in 30,180 patients with 19% of advanced features who underwent surgeries in the United States from 2010 to 2014. The association of pathological features, epidemiologic factors, and pathologic features with overall survival was analyzed. The authors suggested that lobectomy should be recommended as a diagnostic and therapeutic intervention for PTMC patients even without extrathyroid, lymph node, or distant metastasis to avoid further risk of delayed treatment[11].

Choosing AS or surgery as the treatment has always been a debatable topic among the PTMC patients, mainly because there is a potential bias in the treatment selection process. In this case, the patient had a 1 mm tumor in the left thyroid with skip lymph node metastasis. However, a microcarcinoma <3 mm in size may be undetectable on routine pathological examination, even if the entire thyroid gland is examined. Moreover, some PTCs may present with intratumoral fibrosis, which is associated with increased incidences of both extrathyroidal invasion and lymph node metastasis[12]. In many cancers, a single microscopic tumor cell is adequate to cause the loss of life. Therefore, treatment decisions based only on tumor size, age, family history, or imaging of low-risk PTMC are insufficient. To address this issue, genetic testing offers an approach to identify whether the PTMC has a metastatic potential and to provide details to aid in tumor classification and individualized management planning. The genetic testing showed that patient was positive for CCDC6-exon 1-RET-exon 12 fusion (RET/PTC1). At present, 19 possible RET/PTC rearrangements of oncogenic genes have been reported, which are the result of the fusion of domains with

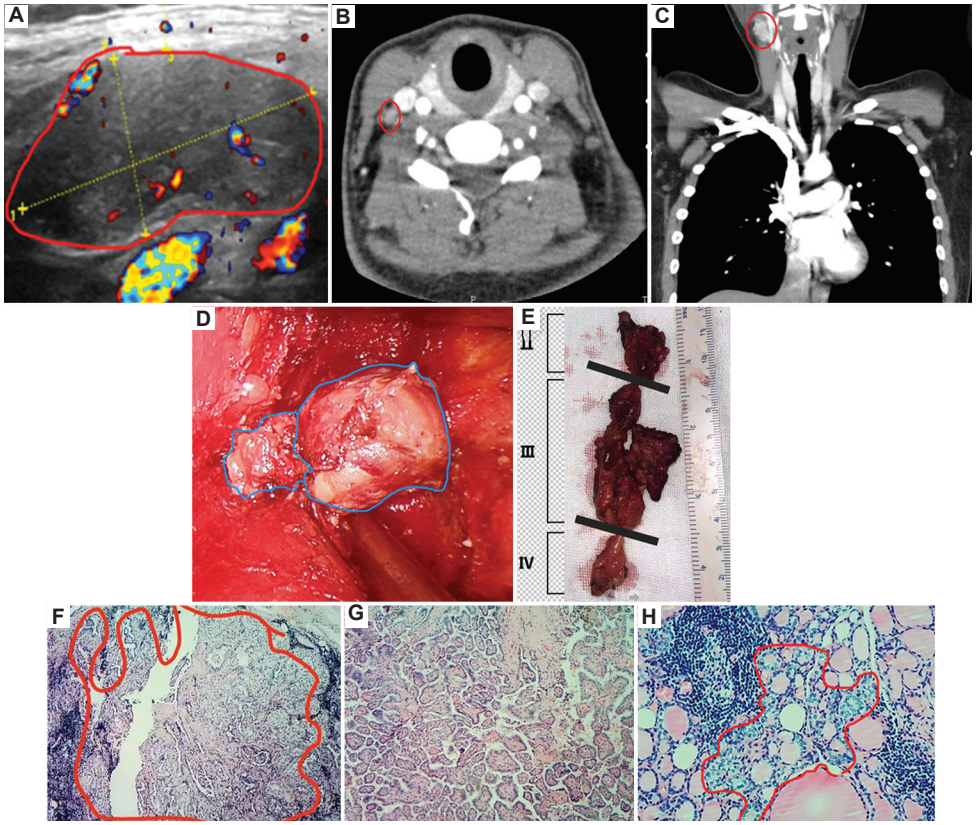


Figure 1. (A-C) Neck-chest computed tomography showed that there was no thyroid mass and massive lymph node enlargement; (D-E) enlarged right cervical lymph nodes before and after surgical resection; (F) Enlarged right cervical lymph node under hematoxylin and eosin (HE) staining; 5× magnification; (G) Lymphocytic thyroiditis in the right thyroid lobe and isthmus under HE staining; 5× magnification; (H) Left papillary thyroid microcarcinoma (1 mm) under HE staining; 5× magnification.

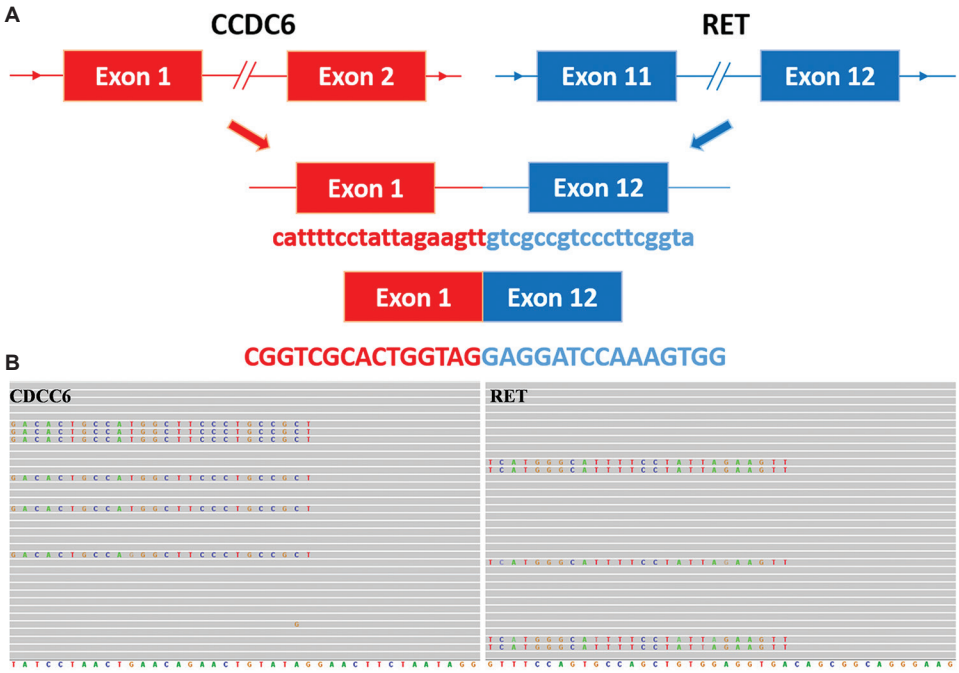


Figure 2. Integrative Genomics Viewer visualization of CCDC6-RET fusion in whole-genome sequencing data.

different genes[13,14]. They are characterized by the presence of protein-coding nucleotide sequences with a high probability of helix-helix domain formation, thus allowing the constitutive dimerization of the RET-TK domain[15]. This dimerization in microcarcinoma plays a driving role in the tumoral transformation[16]. Recent studies have shown that patients with such rearrangements have an 8-fold increased risk of developing lymph node lesions compared with those patients without rearrangements[17]. The result of our patient also supports this hypothesis. Genetic testing of RET/PTC fusions in PTMC may influence the decision on the extent of thyroidectomy and lymph node dissection[18], but this theory needs to supported by longer follow-up studies of higher quality.

With the emergence of gene chips and next-generation sequencing technology, genetic testing will play an important role and have an impact on clinical surgery decisions in the management of PTMC. González *et al.* established a diagnostic model consisting of 10 genes. The diagnostic sensitivity and specificity of the model for benign and malignant indeterminate nodules were 93% and 81%, respectively. The positive predictive value of the model was 78%, and the negative predictive value was 98%, as shown in a multi-center study[19]. Three molecular diagnostic tools, namely, Thygenx/Thyramir, Positive GSC, and Thyroseq V3, have been successfully translated into clinical applications, demonstrating high sensitivity and corresponding high negative predictive values for Bethesda Categories III and IV indeterminate nodules[20-22].

In summary, with technological innovation, individualized diagnosis and prognosis will be the direction of basic research on PTMC. The patient is currently on conventional treatment regimens, after the genetic testing results on RET/PTC rearrangement were known. LOXO-292 treatment would be an optimal option if the disease progresses. An ongoing Phase I/II clinical trial demonstrated that LOXO-292, a highly selective RET inhibitor, has a strong antitumor activity in RET-fusion-positive patients[23].

4. Conclusion

The selection of the most appropriate treatment for PTMC patients, after considering the trade-off between their benefits and risks, has become the focus of attention. The obscure mechanism underlying the metastasis of thyroid carcinoma poses difficulty in diagnosis and treatment development. Nevertheless, thyroid oncogene microarray and next-generation sequencing technology have significantly improved the accuracy of the pre-operative diagnosis of thyroid cancer. Therefore, to avoid overtreatment and under treatment, we suggest that genetic testing should be performed for PTMC patients because the test results can assist in deciding the appropriate treatments, such as surgery or AS.

Acknowledgments

We thank Dr. Dong Liu (Guangzhou KingMed Diagnostics Group Co., Ltd.) and his colleagues for the help with pathological diagnoses and guidance.

Funding

This study is supported by Clinical key specialty construction project funding of Guangdong Province (Guangdong Health Medical Letter [2022] No. 39); the Key Laboratory of Guangdong Higher Education Institutes (2021KSYS009); National College Student Innovation and Entrepreneurship Training Project (No.202310570015); and Guangzhou Medical University Student Innovation Ability Enhancement Program Project (Guangzhou Medical University Letter [2022] No. 66).

Conflict of interest

All authors have completed the ICMJE uniform disclosure form. The authors have no conflicts of interest to declare.

Author contributions

Conceptualization: Song Wang
Investigation: Haitao Peng, Song Wang
Methodology: Chong Wang, Yuanwei Luo
Formal analysis: Zifeng Luo, Song Wang
Writing – original draft: Chang Cai
Writing – review & editing: Haitao Peng

Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee(s) and with the Helsinki Declaration (as revised in 2013).

Consent for publication

Written informed consent was obtained from the patient for publication of this case report and accompanying images.

Availability of data

Supporting data can be obtained from corresponding author following formal request.

References

1. Bai Y, Kakudo K, Jung CK, 2020, Updates in the pathologic classification of thyroid neoplasms: A review of the world

health organization classification. *Endocrinol Metab* (*Seoul*), 35(4): 696–715.

https://doi.org/10.3803/enm.2020.807

2. Saijo H, Kitamura Y, Takenaka H, *et al.*, 2017, Occult thyroid follicular carcinoma diagnosed as metastasis to the chest wall. *Intern Med*, 56(15): 2033–2037.

https://doi.org/10.2169/internalmedicine.56.8340

3. Zheng W, Tan J, Zhang G, 2015, Extensive bone metastases as the initial symptom of papillary thyroid microcarcinoma: A case report. *Exp Ther Med*, 9(6): 2104–2108.

https://doi.org/10.3892/etm.2015.2423

4. Schwaiger K, Koeninger F, Wimbauer J, *et al.*, 2019, Occult papillary thyroid cancer presenting as cystic metastasis of the lateral neck: A case report. *Medicine* (*Baltimore*), 98(30): e16659.

https://doi.org/10.1097/MD.0000000000016659

5. Pollack R, Mazeh H, 2021, Active surveillance of thyroid microcarcinoma-can this approach be safely implemented worldwide? *J Surg Res*, 258: 145–152.

https://doi.org/10.1016/j.jss.2020.08.061

6. Vaccarella S, Franceschi S, Bray F, *et al.*, 2016, Worldwide thyroid-cancer epidemic? The Increasing impact of overdiagnosis. N Engl J Med, 375(7): 614–617.

https://doi.org/10.1056/NEJMp1604412

7. Haugen BR, Alexander EK, Bible KC, *et al.*, 2016, 2015 American thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid*, 26(1): 1–133.

https://doi.org/10.1089/thy.2015.0020

8. Sugitani I, Ito Y, Takeuchi D, *et al.*, 2021, Indications and strategy for active surveillance of adult low-risk papillary thyroid microcarcinoma: Consensus statements from the Japan association of endocrine surgery task force on management for papillary thyroid microcarcinoma. *Thyroid*, 31(2): 183–192.

https://doi.org/10.1089/thy.2020.0330

9. Haser GC, Tuttle RM, Su HK, *et al.*, 2016, Active surveillance for papillary thyroid microcarcinoma: New challenges and opportunities for the health care system. *Endocr Pract*, 22(5): 602–611.

https://doi.org/10.4158/ep151065.ra

10. Orlando G, Scerrino G, Corigliano A, *et al.*, 2022, Papillary thyroid microcarcinoma: Active surveillance against surgery. Considerations of an Italian working group from a systematic review. *Front Oncol*, 12: 859461.

https://doi.org/10.3389/fonc.2022.859461

11. Al-Qurayshi Z, Nilubol N, Tufano RP, *et al.*, 2020, Wolf in sheep's clothing: Papillary thyroid microcarcinoma in the US. *J Am Coll Surg*, 230(4): 484–491.

https://doi.org/10.1016/j.jamcollsurg.2019.12.036

12. Xu B, Scognamiglio T, Cohen PR, *et al.*, 2017, Metastatic thyroid carcinoma without identifiable primary tumor within the thyroid gland: A retrospective study of a rare phenomenon. *Hum Pathol*, 65: 133–139.

https://doi.org/10.1016/j.humpath.2017.05.013

13. Yakushina VD, Lerner LV, Lavrov AV, 2018, Gene fusions in thyroid cancer. *Thyroid*, 28(2): 158–167.

https://doi.org/10.1089/thy.2017.0318

14. Hsiao SJ, Nikiforov YE, 2014, Molecular approaches to thyroid cancer diagnosis. *Endocr Relat Cancer*, 21: T301–T313.

https://doi.org/10.1530/erc-14-0166

15. Romei C, Elisei R, 2012, RET/PTC translocations and clinico-pathological features in human papillary thyroid carcinoma. *Front Endocrinol* (*Lausanne*), 3: 54.

https://doi.org/10.3389/fendo.2012.00054

16. Viglietto G, Chiappetta G, Martinez-Tello FJ, *et al.*, 1995, RET/PTC oncogene activation is an early event in thyroid carcinogenesis. *Oncogene*, 11(6): 1207–1210.

17. Lan X, Bao H, Ge X, *et al.*, 2020, Genomic landscape of metastatic papillary thyroid carcinoma and novel biomarkers for predicting distant metastasis. *Cancer Sci*, 111(6): 2163–2173.

https://doi.org/10.1111/cas.14389

18. Musholt TJ, Staubitz JI, Cámara RJ, *et al.*, 2019, Detection of RET rearrangements in papillary thyroid carcinoma using RT-PCR and FISH techniques-a molecular and clinical analysis. *Eur J Surg Oncol*, 45(6): 1018–1024.

https://doi.org/10.1016/j.ejso.2018.11.009

19. González HE, Martínez JR, Vargas-Salas S, *et al.*, 2017, A 10-gene classifier for indeterminate thyroid nodules: Development and multicenter accuracy study. *Thyroid*, 27(8): 1058–1067.

https://doi.org/10.1089/thy.2017.0067

20. Steward DL, Carty SE, Sippel RS, *et al.*, 2019, Performance of a multigene genomic classifier in thyroid nodules with indeterminate cytology: A prospective blinded multicenter study. *JAMA Oncol*, 5(2): 204–212.

https://doi.org/10.1001/jamaoncol.2018.4616

21. Labourier E, Shifrin A, Busseniers AE, *et al.*, 2015, Molecular testing for miRNA, mRNA, and DNA on fine-needle aspiration improves the preoperative diagnosis of thyroid nodules with indeterminate cytology. *J Clin Endocrinol Metab*, 100(7): 2743–2750.

https://doi.org/10.1210/jc.2015-1158

22. Titov S, Demenkov PS, Lukyanov SA, *et al*., 2020, Preoperative detection of malignancy in fine-needle aspiration cytology (FNAC) smears with indeterminate cytology (Bethesda III, IV) by a combined molecular classifier. *J Clin Pathol*, 73(11): 722–727. https://doi.org/10.1136/jclinpath-2020-206445

23. Dias-Santagata D, Lennerz JK, Sadow PM, *et al*., 2020, Response to RET-specific therapy in RET fusion-positive anaplastic thyroid carcinoma. *Thyroid*, 30(9): 1384–1389. https://doi.org/10.1089/thy.2019.0477

ACCSCIENCE PUBLISHING

Gene & Protein in Disease

CASE REPORT

Multisystem inflammatory syndrome in adult with longitudinally extensive transverse myelitis following SARS-CoV-2 vaccination: A case report

Anirban Gupta[1]*, Sudheer Pandey[1], Pawan Dhull[1], Amit Sreen[1], Satish Barki[1], Krishna Kumar[2], and Kompella Kiran Kumar[3]

[1]Department of Neurology, Army Hospital (R&R), New Delhi, India

[2]Department of Anesthesia and Critical Care, Army Hospital (R&R), New Delhi, India

[3]Department of Internal Medicine, Army Hospital (R&R), New Delhi, India

Abstract

Multisystem inflammatory syndrome in adults (MIS-A) is a known complication arising after severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection or vaccination. However, longitudinal extensive transverse myelitis (LETM), which is a neurological manifestation of MIS-A, is rarely reported among individuals after SARS-CoV-2 infection or vaccination. Here, we present a case of a 38-year-old female, with a history of SARS-CoV-2 infection in May 2021 and subsequent SARS-CoV-2 vaccination (on November 1, 2021; COVISHIELD™, ChAdOx1 nCoV-19 Corona Virus Vaccine), who presented with acute flaccid quadriparesis with bowel and bladder involvement since December 26, 2021. While being treated in the hospital, she developed fever, rash, shock, myocarditis, breathlessness, jaundice, and acute kidney injury. Magnetic resonance Imaging revealed LETM. She was negative for SARS-CoV-2 infection and showed raised in SARS-CoV-2 antibody titer. Thus, she was diagnosed as having MIS-A with LETM and managed with Intravenous Immunoglobulin, antibiotics, hemodialysis, and steroids with improvement.

Keywords: Severe acute respiratory syndrome coronavirus 2; Multisystem inflammatory syndrome in adults; Vaccination; Flaccid quadriparesis; Longitudinal extensive transverse myelitis

***Corresponding author:**
Anirban Gupta
(dranirban1974@gmail.com)

Citation: Gupta A, Pandey S, Dhull P, *et al.*, 2023, Multisystem inflammatory syndrome in adult with longitudinally extensive transverse myelitis following SARS-CoV-2 vaccination: A case report. *Gene Protein Dis,* 2(3): 1320.
https://doi.org/10.36922/gpd.1320

Received: July 18, 2023
Accepted: September 5, 2023
Published Online: September 29, 2023

1. Background

Multisystem inflammatory syndrome shares an inexplicable link with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Multisystem inflammatory syndrome in children (MIS-C) following SARS-CoV-2 infection was first described in May 2020[1,2], whereas the first case of multisystem inflammatory syndrome in adults (MIS-A) following SARS-CoV-2 infection was described in October 2020[3].

The invention of a viable SARS-CoV-2 vaccine was first announced in December 2020, marking a monumental milestone in our battle against SARS-CoV-2 infection, and following that, a series of vaccines developed in the USA, the UK, Russia, India, and China were rolled out, whose ability in conferring protective immunity had been clinically tested. COVISHIELD™, ChAdOx1 nCoV-19 Corona Virus Vaccine is one of the principal

vaccines widely used in India. Nonreplicating adenovirus is used in COVISHIELD™ to express SARS-CoV-2 spike protein, which is presented to the antigen-presenting cells of the host[4]. Various hypotheses have been proposed to delineate the exact mechanisms causing immune-mediated neurological complications following SARS-CoV-2 vaccination, and among them, molecular mimicry is recognized as the most plausible mechanism[5]. Components of the adenovector-based SARS-CoV-2 vaccine can activate an unregulated inflammatory cascade, which damages the brain and peripheral nerves[5]. In adenovector-based SARS-CoV-2 vaccines, the SARS-CoV-2 spike protein may also act as an autoantigen, with the adjuvants enhancing the immune reactivity further[5]. MIS-A after SARS-CoV-2 vaccination is an extremely rare complication[6-8]. According to the definition by Center for Disease Control and Prevention (CDC), USA, MIS-A encompasses a spectrum of new-onset neurological manifestations, such as encephalopathy, seizures, signs of meningeal irritation, and peripheral neuropathy like Guillain–Barre syndrome; the presence of these manifestations constitutes the secondary clinical criteria of MIS-A. Longitudinally extensive transverse myelitis (LETM) is regarded as a rare complication after SARS-CoV-2 vaccination[8,9]. Thus far, only one case has been described, establishing LETM as a neurological complication of MIS-C after SARS-CoV-2 infection[10]. Here, we report a case of MIS-A with LETM after SARS-CoV-2 vaccination in an adult woman.

2. Case presentation

A 38-year-old female was presented at our hospital with acute-onset progressive flaccid quadriparesis on December 26, 2021, which had progressed for more than 36 h with sign of sensory paresthesia up to lower abdomen with bladder and bowel involvement, accompanied by urinary retention and bowel incontinence. This patient has a medical history of hypothyroidism in the last 5 – 6 years but has not been on regular medication. She also has a history of mild SARS-CoV-2 infection in May 2021 from which she recovered in 2 weeks, and received the second dose of SARS-CoV-2 vaccine (COVISHIELD™, ChAdOx1 nCoV-19 Corona Virus Vaccine) on November 1, 2021. She had breathing difficulty since December 28, 2021, requiring ventilator support, and faced severe hypotension necessitating ionotropic support in the form of injection noradrenaline and vasopressin. There was no immediately preceding history of fever, headache, photosensitive rashes, or recurrent orogenital ulcers. The patient showed a history of low-grade fever in November after SARS-CoV-2 vaccination for which she took intermittent non-steroidal anti-inflammatory medication. Dietary, family, psychosocial, family, and obstetric history was noncontributary. Physical examination revealed that the patient had pulse rate of 130/min, blood pressure of 120/86 mmHg on injection noradrenaline infusion, and breathing rate of 26 breaths/min. With the aid of an oxygen face mask administering oxygen at 8 L/min, her oxygen saturation was maintained at 98%. Neurologically, Glasgow–Coma scale (GCS) was 15/15. Higher mental functions and cranial nerve findings of the patient were normal. An examination of her motor system revealed that she had flaccid quadriparesis with medical research council (MRC) grade 4+/5 power in upper limbs and MRC grade 0/5 power in lower limbs. All reflexes were absent and plantars were mute. She had reduced sensation to touch, pain, and temperature below umbilicus, which was more pronounced in the lower back and lower limbs. Blood investigations revealed polymorphonuclear leukocytosis, thrombocytopenia, acute kidney injury, acute liver injury, and disseminated intravascular coagulation. Reverse-transcription polymerase chain reaction (RT-PCR) test for SARS-CoV-2 was negative. Table 1 presents the results of preliminary investigations.

Chest X-ray showed features of acute respiratory distress syndrome. 2D echocardiography showed global left ventricular hypokinesia with left ventricular ejection fraction (LVEF) of 40 – 45%. The level of N-terminal pro B type natriuretic peptide (NT Pro-BNP) was found to exceed 70000 pg/mL. Within 1 day of admission, she became febrile but testing of viral markers and tropical fever showed negative results. The patient also had a normal vasculitis profile and serum angiotensin-converting enzyme (ACE) level, but the levels of serum inflammatory markers, such as C-reactive protein (CRP), interleukin-6 (IL-6) and procalcitonin, were raised. Total SARS-CoV-2 antibody titer, measured by ELISA, was 2816 IU/mL (normal <1000 IU/mL). On January 2, 2022, she developed maculopapular rash over the angle of mouth and the face. In view of the presence of cardiac illness, rashes, new-onset neurological signs and symptoms, shock, thrombocytopenia, as well as raised CRP, IL-6, procalcitonin, and SARS-CoV-2 antibody levels, she was diagnosed as having MIS-A with underlying sepsis.

Accordingly, the patient was treated with mechanical ventilation, ionotropic support in the form of noradrenaline injection, intravenous immunoglobulin (IVIG) 2 g/kg over 5 days since January 2, 2022, intravenous broad-spectrum antibiotics, and hemodialysis, along with other supportive management. Contrast-enhanced computed tomography of chest and abdomen conducted on January 6, 2022, revealed diffuse ground-glass opacity and consolidation in the bilateral lung field involving perihilar, peripheral, and bilateral upper lobes. An analysis of the patient's cerebrospinal fluid revealed albumin-cytological

Table 1. Results of preliminary investigation

	December 27, 2021	December 28, 2021; morning	December 28, 2021; evening	December 29, 2021
Hemoglobin (mmol/L)	1.61	1.49	1.63	1.64
Total leukocyte count (/L)	11.01×10^9	13.7×10^9	19.9×10^9	18.9×10^9
Differential leukocyte count	P73L20%	P95L3.5%	P95L5%,	P93L4%
Platelets (/L)	176×10^9	68×10^9	70×10^9	44×10^9
Blood urea nitrogen (mmol/L)	5.71	/	12.5	12.5
Creatinine (μmol/L)	64.5	/	348.3	384.5
Total bilirubin/direct (μmol/L)	/	80.5/34.9	/	122.8/74.9
SGOT/SGPT (μKat/L)	0.4/0.33	2.09/0.9	7.02/3.32	7.6/3.39
Sodium/potassium/calcium (mmol/L)	139/4.05/7.7	136/4.1/7.6	132/3.6/7.5	139/3.9/7.7
PT/APTT/INR (s)	/	/	22.7 (13.6)/70 (33)/1.71	34.3 (13.6)/79.8 (33)/2.64
D dimer/Fib	/	/	/	>Max/-
Peripheral blood smear	/	/	Normocytic macrocytes, schistocytes and no toxic granules	/
Urine routine and culture	/	Normal/no growth	/	/
Blood culture	/	No growth	/	/
Miscellaneous	• Serum porphyrins: negative • CPK: normal • ABG: normal • Thyroid function test: normal • RT-PCR for SARS-CoV-2: negative	• U1-nRNP/Sm: negative • Sm: negative • SS-A: negative • Ro-52: negative • SS-B/La: negative • Scl-70: negative • Pm-Scl: negative • Jo-1: negative • CENP-B: negative • PCNA: negative • dsDNA: negative • ANA (IF) : negative • cANCA/pANCA: negative	• CSF-Cells-02 • Glucose (CSF/serum): 1.14/2.56 mmol/L • Protein: 1050 mg/L, • Stains: negative • COVID antibody (total): 2816 IU/mL (normal<1000 IU/mL)	• Spot urine protein: 495 mg/dl • ABG: respiratory alkalosis • CPK: 316 U/L • CPK-MB: 0.43 μKat/L • Dengue/malaria/typhoid/ leptospira/scrub typhus/ lyme: negative • NT-ProBNP: >8260 pmol/L • Procalcitonin: >200 ng/mL • CRP: 1240 mg/L • HBsAg/Anti HCV/HIV/ VDRL: negative • USG (W/A): normal • 2D echocardiography: global left ventricular hypokinesia • LVEF: 45%

Note: /: Not done. Abbreviations: ABG: Arterial blood gas; ANA: Antinuclear antibody; ANCA: Antineutrophilic cytoplasmic antibody; APTT: Partial thromboplastin time–activated; cANCA: ANCA that targets proteinase 3; CRP: C-reactive protein; CSF: Cerebrospinal fluid; CPK: Creatine phosphokinase; dsDNA: Double-stranded DNA; HBsAg: Hepatitis B surface antigen; HCV: Hepatitis C virus; HIV: Human immunodeficiency virus; INR: International normalized ratio; LVEF: Left ventricular ejection fraction; pANCA: ANCA that targets myeloperoxidase; PCNA: Proliferating cell nuclear antigen; PT: Prothrombin time; RT-PCR: Reverse-transcription polymerase chain reaction; SGOT: Serum glutamic-oxaloacetic transaminase; SGPT: Serum glutamic-pyruvic transaminase; USG (W/A): Ultrasonography of whole abdomen; VDRL: Venereal disease research laboratory test.

dissociation. On the other hand, nerve conduction study revealed pure motor involvement in both upper and lower limbs, and magnetic resonance imaging (MRI) of the brain and spine performed on January 11, 2022, revealed minimal sub-arachnoid hemorrhage in left Sylvian fissure and long segment altered signal intensity from C6 to conus medullaris, which was suggestive of LETM. Figure 1 shows an MRI image displaying LETM in cervico-dorsal cord. Further tests showed that the patient was negative for serum neuromyelitis optica and myelin oligodendrocyte glycoprotein antibody.

In view of less-than-satisfactory neurological recovery, she was further treated with pulse intravenous methylprednisolone 1 g for 5 days during January 15–19, 2022, followed by oral prednisolone taper. Hemodialysis treatment was indicated for her for 1 month since she had anuric acute kidney injury, and her renal functions improved gradually over 1 month.

Table 2. Results of the follow-up investigations

	January 13, 2022	January 21, 2022	January 31, 2022	February 4, 2022	February 8, 2022
Hemoglobin (mmol/L)	1.1	1.66	1.64	1.30	1.33
Total leukocyte count (/L)	10.16×10^9	11.0×10^9	16.6×10^9	16.9×10^9	8.54×10^9
Differential leukocyte count	P73L20%	P76L21%	P94L21%	P90L8%	P80L16%
Platelets (/L)	422×10^9	323×10^9	220×10^9	168×10^9	248×10^9
Blood urea nitrogen (mmol/L)	76.0	38.9	/	36.05	28.9
Creatinine (μmol/L)	619.7	381.9	/	263.4	127.3
Total bilirubin (μmol/L)	14.9	13.0	9.40	10.9	9.91
SGOT/SGPT (μKat/L)	1.26/0.78	0.81/0.63	0.34/0.36	0.53/0.73	0.35/0.46
Sodium/potassium (mmol/L)	133/5.5	138/5.2	128/5.5	133/3.9	131/4.3
PT/APTT/INR (s)	19.2/180/1.5	/	13.3/-/0.98	16.1/<min/1.2	/
D dimer/Fib	/	/	/	12.90 mg/L	/
Peripheral blood smear	/	/	No toxic granules	/	/
2D echocardiography	/	/	/	/	• Normal left ventricular function • No regional wall motion abnormality • LVEF: 60%

Note: /: Not done. Abbreviations: APTT: Partial thromboplastin time–activated; INR: International normalized ratio; LVEF: Left ventricular ejection fraction; PT: Prothrombin time; SGOT: Serum glutamic-oxaloacetic transaminase; SGPT: Serum glutamic-pyruvic transaminase.

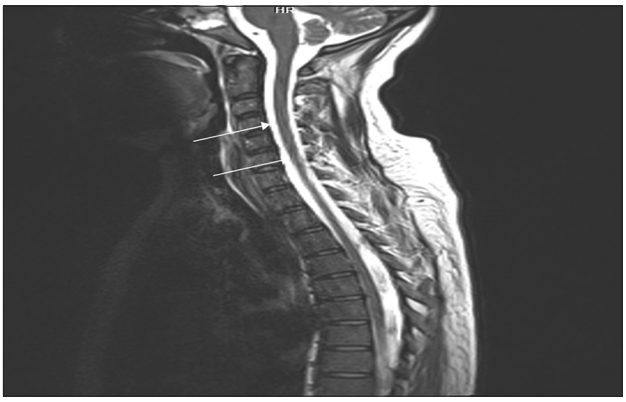


Figure 1. Longitudinally extensive transverse myelitis from the C6 segment to conus medullaris in the cervico-dorsal cord.

Overall, the patient started recovering after IVIG was applied. Inotropes were tapered off on January 5, 2022, and she was extubated on the same day. The patient regained MRC 5/5 power in both upper limbs and MRC 2/5 power in both lower limbs and was able to sit with support after 1 month. In addition, her bowel function and bladder sensation improved after 1 month. Table 2 presents the results of the follow-up investigations. Informed consent was taken from the patient before submitting this case report.

3. Discussion

To the best of our knowledge, LETM has never been reported as a neurological manifestation of MIS-A after SARS-CoV-2 infection or vaccination. Here we present a case of LETM as a neurological manifestation of MIS-A after SARS-CoV-2 vaccination. The case satisfied the diagnostic criterion of MIS-A and also had LETM as a neurological manifestation. This patient was initially diagnosed as having sepsis with septic shock and multi-organ dysfunction. However, the diagnosis of sepsis was instantly ruled out given the constellation of symptoms and rapid response to immunotherapy. For differential diagnosis purposes, we also examined other causes of LETM, such as vasculitis, neuromyelitis optica, and myelin oligodendrocyte glycoprotein-related demyelination, which were however found to be negative in this patient. After ruling out all types of infections, including SARS-CoV-2, as the contributory factors of LETM, we considered vaccination a plausible cause given the elevated SARS-CoV-2 antibody levels. As the patient caught the SARS-CoV-2 infection approximately 7 months before consulting us in the hospital, we believe it is more convincing to attribute SARS-CoV-2 vaccination to the development of LETM and MIS-A, taking into account the recent vaccination 8 weeks before her visit to our hospital. However, despite this speculation, a recent yet undiagnosed SARS-CoV-2 infection cannot be ruled out completely.

4. Conclusion

At present, there are no other cases reported in literature corroborating LETM as a neurological manifestation of MIS-A after SARS-CoV-2 infection or vaccination; therefore, it is challenging to reach a definitive conclusion on this topic. As the SARS-CoV-2 pandemic waxes and

wanes with majority of population having access to vaccines, vaccine-related adverse events are gradually emerging. Various reports of Guillain-Barre syndrome, transverse myelitis, encephalitis, meningitis, Bell's palsy, and hemorrhagic strokes, CNS demyelination, myasthenic syndromes, and MIS-A have emerged after SARS-CoV-2 vaccination[11-14]. Our case report provides insights to aid further investigation on this topic. Based on the case presented here, the onset of MIS-A with LETM occurs following SARS-CoV-2 vaccination. However, the potential association between vaccination and these symptoms should not dissuade us from vaccination, which regardless has proven to be a protective approach against SARS-CoV-2 infection.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Anirban Gupta, Sudheer Pandey
Investigation: All authors
Writing – original draft: Anirban Gupta, Sudheer Pandey
Writing – review & editing: Anirban Gupta

Ethics approval and consent to participate

Ethical approval was waived due to retrospective nature of report. Informed consent was taken from of patients.

Consent for publication

Consent for publication was taken from the patient verbally and in writing.

Availability of data

Not applicable.

Further disclosure

This case was presented in Indian Academy of Neurology Conference in New Delhi in 2022 and the paper has been uploaded to Research Square (https://doi.org/10.21203/rs.3.rs-1564946/v1).

References

1. Riphagen S, Gomez X, Gonzalez-Martinez C, *et al*., 2020, Hyperinflammatory shock in children during COVID-19 pandemic. *Lancet*, 395(10237): 1607–1608.

https://doi.org/10.1016/S0140-6736(20)31094-1

2. Verdoni L, Mazza A, Gervasoni A, *et al*., 2020, An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: An observational cohort study. *Lancet*, 395(10239): 1771–1778.

https://doi.org/10.1016/S0140-6736(20)31103-X

3. Morris SB, Schwartz NG, Patel P, *et al*., 2020, Case series of multisystem inflammatory syndrome in adults associated with SARS-CoV-2 infection-United Kingdom and United States, March-August 2020. *MMWR Morb Mortal Wkly Rep*, 69(40): 1450–1456.

https://doi.org/10.15585/mmwr.mm6940e1

4. Das S, Kar SS, Samanta S, *et al*., 2022, Immunogenic and reactogenic efficacy of Covaxin and Covishield: A comparative review. *Immunol Res*, 70(3): 289–315.

https://doi.org/10.1007/s12026-022-09265-0

5. Garg RK, Paliwal V, Malhotra HS, *et al*., 2023, Spectrum of serious neurological and psychiatric adverse events in Indian COVID-19 vaccine recipients: A systematic review of case reports and case series. *Neurol India*, 71: 209–227.

https://doi.org/10.4103/0028-3886.375420

6. Uwaydah AK, Hassan NM, Abu Ghoush MS, *et al*., 2021, Adult multisystem inflammatory syndrome in a patient who recovered from COVID-19 postvaccination. *BMJ Case Rep*, 14(4): e242060.

https://doi.org/10.1136/bcr-2021-242060

7. Salzman MB, Huang CW, O'Brien CM, *et al*., 2021, Multisystem inflammatory syndrome after SARS-CoV-2 infection and COVID-19 vaccination. *Emerg Infect Dis*, 27(7): 1944–1948.

https://doi.org/10.3201/eid2707.210594

8. Gao JJ, Tseng HP, Lin CL, *et al*., 2021, Acute transverse myelitis following Covid-19 vaccination. *Vaccines* (*Basel*), 9(9): 1008.

https://doi.org/10.3390/vaccines9091008

9. Pagenkopf C, Südmeyer M, 2021, A case of longitudinally extensive transverse myelitis following vaccination against Covid-19. *J Neuroimmunol*, 358: 577606.

https://doi.org/10.1016/j.jneuroim.2021.577606

10. Pourmoghaddas Z, Sadeghizadeh A, Tara SZ, *et al*., 2021, Longitudinally extensive transverse myelitis as a sign of multisystem inflammatory syndrome following COVID-19 infection: A pediatric case report. *J Neuroimmunol*, 360: 577704.

https://doi.org/10.1016/j.jneuroim.2021.577704

11. Ismail II, Salama S, 2022, A systematic review of cases of CNS demyelination following COVID-19 vaccination.

J Neuroimmunol, 362: 577765.

https://doi.org/10.1016/j.jneuroim.2021.577765

12. Tan WY, Khan AH, Yaakob MN, *et al*., 2021, Longitudinal extensive transverse myelitis following ChAdOx1 nCOV-19 vaccine: A case report. *BMC Neurol*, 21(1): 395.

https://doi.org/10.1186/s12883-021-02427-x

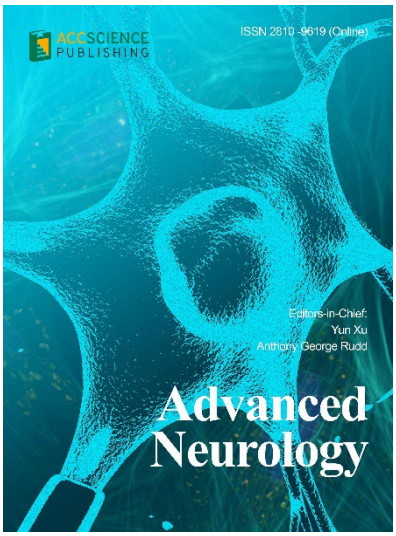
13. Notghi AA, Atley J, Silva M, 2021, Lessons of the month 1: Longitudinal extensive transverse myelitis following AstraZeneca COVID-19 vaccination. *Clin Med* (*London*), 21(5): e535–e538.

https://doi.org/10.7861/clinmed.2021-0470

14. Grome HN, Threlkeld M, Threlkeld S, *et al*., 2021, Fatal multisystem inflammatory syndrome in adult after SARS-COV-2 natural infection and Covid-19 vaccination. *Emerg Infect Dis*, 27(11): 2914–2918.

https://doi.org/10.3201/eid2711.211612

Our Journals



Advanced Neurology is a peer-reviewed and open-access journal that aims to publish and disseminate novel research in the breadth of neurology and neuroscience. The journal aims to advance our understanding in the nervous system and provide a platform to neuroscientists and physicians to showcase their findings in original fundamental and clinical research as well as to present new ideas that highlight the changes in the neurological clinical practice.

Advanced Neurology covers subject areas, including but not limited to the following:

- Neurological disorders
- Neurodegenerative disease
- Cerebrovascular disease
- Epilepsy and movement disorders
- Neuroimmune disease
- Neurological infections
- Muscle disease
- Molecular and cellular neuroscience
- Systems neuroscience
- Cognitive neuroscience
- Computational modeling of nervous system

Global Translational Medicine is a quarterly journal that focuses on medicine, biological sciences, and biomaterials engineering. The goal of *Global Translational Medicine* is to provide a platform to researchers for showcasing their latest research works in translational medicine so as to advance the field towards the betterment of human health. Despite the advancement of omics and new technologies, the process of transforming these technologies and scientific research results into effective therapies and putting them into clinical use still has a long way to go. *Global Translational Medicine* provides a platform to fill the gaps in preclinical and inter-disciplinary research, to promote clinical translation of scientific research results, and to contribute to the conception of new and improved preventive measures as well as diagnostic and therapeutic techniques of diseases.

Global Translational Medicine covers the following themes: cardiovascular disease, metabolism/diabetes/obesity, neuroscience/neurology, cancer, biomaterials and their applications in medicine, proteomics/metabolomics, pharmacogenomics, biomarkers, bioinformatics and data mining, animal and clinical research, and medical methods arising from interdisciplinary crossover.

Start a new journal

Write to us via email if you are interested to start a new journal with AccScience Publishing. Please attach your CV, professional profile page and a brief pitch proposal in your email. We shall inform you of our decision whether we are interested to collaborate in starting a new journal.

Contact: info@accscience.com

https://accscience.com/journal/GPD

Contact
www.accscience.com
8 Burn Road, #15-03 Trivex, Singapore 369977
Email: editorial@accscience.com
Phone: +65 8182 1586