JNRS

JOURNAL OF NEW RESULTS IN SCIENCE

http://dergipark.gov.tr/jnrs

ISSN: 1304 - 7981

VOLUME : 13
NUMBER : 2
YEAR : 2024

jnrs@gop.edu.tr

Journal of New Results in Science (abbreviated as J. New Results Sci. or JNRS) is an international, peer-reviewed, open-access journal focusing on new science theories or their applications. JNRS publishes research and review articles in Mathematical and Statistical sciences and Biological, Chemical, and Physical sciences.

J. New Results Sci. was founded in 2011, and its first issue was published on 01 January 2012.

Language: As of 2023, JNRS accepts contributions in **American English** only.

Frequency: 3 Issues Per Year.

Publication Dates: April, August, and December.

ISSN: 1304-7981

Editor-in-Chief: Ali YAKAR

Email: togu.jnrs@gmail.com

Publisher: Tokat Gaziosmanpaşa University

APC: JNRS incurs no article processing charges.

Review Process: Single-Blind Peer Review.

DOI Numbers: The published papers are assigned DOI numbers.

Editor-in-Chief

Editors

Section Editors

Statistics

Those who contributed 2012-2023

Statistics Editor

Language Editors

Layout Editors

Production Editor

CONTENTS

Journal of New Results in Science

https://dergipark.org.tr/en/pub/jnrs

Research Article

Open Access

E-ISSN:1304-7981 https://doi.org/10.54187/jnrs.1504722

Goodness-of-fit tests based on Kullback-Leibler divergence for bladder cancer survival analysis: Applications to exponentiated exponential distribution

Gülcan Gencer[1]

Keywords:
Censoring, Survival Analysis, Kullback-Leibler Divergence, Goodness of Fit Test, Cancer

Abstract – Bladder cancer is among the ten most common types of cancer worldwide, with approximately 550,000 new cases occurring each year. It accounts for comprehensively compared to 3% of all newly diagnosed cancer cases and contributes to 2.1% of cancer-related deaths globally. This article introduces goodness-of-fit tests that aim to fit the exponentialized exponential distribution. These tests are based on the Kullback-Leibler difference and have been applied to censored and complete samples of Bladder Cancer Patients. We calculated critical values and statistical power measurements, considering the best and worst bandwidth scenarios. We then comprehensively compared essential values and power across various parameters, accounting for optimal and suboptimal bandwidth choices derived from the Kullback–Leibler difference. In the final phase of our study, we used a dataset of individuals diagnosed with bladder cancer to demonstrate the practical applicability of our proposed research. Finally, this modeling type can benefit researchers and healthcare professionals through time-to-event analysis (survival analysis), investigation of events, medical decision-making, and risk prediction.

Subject Classification (2020): 62N02, 62N03

1. Introduction

Bladder cancer is one of the ten most prevalent types of cancer worldwide, representing a significant global health concern. Each year, approximately 550,000 new cases are diagnosed, making bladder cancer a substantial contributor to the worldwide cancer burden. This type of cancer accounts for about 3% of all newly diagnosed cancer cases, reflecting its widespread impact. Furthermore, bladder cancer is responsible for approximately 2.1% of all cancer-related deaths, highlighting its severity and the challenges associated with its treatment and management. Various risk factors, including smoking, exposure to certain industrial chemicals, chronic bladder inflammation, and a history of schistosomiasis in certain regions, influence the incidence of bladder cancer. Despite advances in medical research and treatment options, bladder cancer remains a formidable disease, with significant implications for patient quality of life and overall public health. Ongoing research and clinical efforts focus on improving early detection, treatment outcomes, and survival rates for individuals diagnosed with bladder cancer [1].

An exponentiated exponential distribution is a statistical model describing a measured variable's distribution. For example, a patient's survival times. This model determines the probability of an event, e.g.,

[1]gulcan.gencer@afsu.edu.tr (Corresponding Author)
[1]Department of Biostatistics and Medical Informatics, Faculty of Medicine, Afyonkarahisar Health Sciences University, Afyonkarahisar, Türkiye
Article History: Received: 25 Jul 2024 — Accepted: 27 Aug 2024 — Published: 31 Aug 2024

death, occurring within a given time and calculates this probability over the distribution of the observed variable, e.g., survival time of patients. The exponential distribution is primarily used to model how events occur over time, and this distribution is particularly suitable for rare or irregularly occurring events. Thus, using this distribution allows us to understand specific patterns and trends in the data better, such as the survival time of patients. Modeling the mortality rates of bladder cancer patients with an exponentiated exponential distribution can be used to examine the probability of a death event over a given time.

This type of modeling can provide researchers and healthcare professionals with the following benefits:

i. Event Duration Analysis (Survival Analysis): Exponentialized exponential distribution is a frequently used model for event duration analysis. This analysis evaluates how long patients survive until a specific event (e.g., death) by examining the time for a particular event.
ii. Examining Events: Modeling mortality rates of bladder cancer patients can be used to understand what factors influence the risk of death in a particular population. These factors may include genetic characteristics, treatment methods, age, and gender.
iii. Medical Decision Making: Modeling results can help determine treatment processes and strategies for patients with bladder cancer. For example, it can be used to evaluate how a particular treatment method or medication affects patients' survival.
iv. Risk Forecasting: Exponentialized exponential distribution can be used to estimate the probability of an event occurring within a given time. This could be useful for predicting patients' future survival and adapting treatment plans accordingly.

Moreover, states, public institutions, organizations, and businesses use decision-making tools when planning for the future. The most crucial step in this observation-based decision-making process is modeling the population from which the observations are obtained. Naturally, the model's accuracy will be a critical consideration for the decision-making process. Inaccurate findings from a model that isn't determined appropriately could be irreversible. A probability distribution is a model representing this population, and goodness of fit tests are performed to determine whether a given probability distribution is appropriate for that population. Since the beginning, statisticians have begun their analysis by distributing the observed data. Then, they verified that their chosen distribution was suitable for the observed data. As a result, various test processes have been developed over time, and the study of these procedures is known as goodness of fit [2]. Pearson [3] invented the chi-square test in 1900, which helped to pioneer appropriateness tests. Since then, numerous more tests have been devised, each reflecting the subject's relevance and demands.

Furthermore, it is impossible to dispute the significance of statistical or probabilistic modeling in the modern world. High-speed computers have enabled the development and using complicated models for crucial operations. These models facilitate Effective decision-making and associated statistical analyses in various domains, including marketing, medicine, management, politics, military systems, and food science. Evaluating the validity of models with statistical distributions is known as "goodness of fit."

It is a fundamental and occasionally overlooked part of modeling work. Nonetheless, a wide range of statistical or probabilistic distribution models have found widespread use in engineering, science, economics, and medicine. It's critical to assess these models' applicability or determine how well the data fits the suggested distribution model. To ensure that the selected model accurately represents the underlying population, various distribution families have been proposed for goodness of fit tests, particularly in complete and censored samples. These families are chosen because they offer a flexible framework for capturing the characteristics of different types of data, and the Kullback-Leibler divergence measure is utilized to quantify the fit between the data and the proposed distribution. This approach is critical as it allows for a more precise assessment of the model's suitability across various applications. Here are a few of these studies. Arizona and Ohta [4] presented a normal distribution appropriateness test for a complete

sample based on the Kullback-Leibler divergence measure, the extended version of entropy. To determine which test was the strongest compared to other tests using normal distributions, they compared it to the Durbin version of the K-S test, Cramer Von Mises, weighted Cramer Von Mises, and Chi-Square tests. Şenoğlu and Sürücü [5] used the Kullback-Leibler divergence measure, Shapiro Wilk, Tiku's test, and sample correlation test for Normal, Exponential, and Uniform distributions to compare tests for various distributions (skewed, long, and short-tailed symmetric). Based on the test results, it was observed that the test with the Kullback-Leibler divergence measure generally has more power in distributions with short tails than in distributions with long tails. Choi et al. [6] put out an exponential test that relied on the divergence metric of Kullback-Leibler. Van Es and Correa's entropy was utilized as an estimator of Shannon entropy. Its more significant power than other tests has been determined by comparison with different goodness of fit tests. Park [7] proposed an exponentiality test specifically tailored for type-2 censored data, leveraging Kullback-Leibler insights. When scrutinizing the test's statistical power across alternative distributions such as Gamma, Weibull, and Chi-square, it was observed that the suggested statistical metric exhibited greater sensitivity when applied to distributions with hazard functions that show a consistent upward trend.

In a related study, Lim and Park [8] compared statistical power comprehensively, focusing on partial Kullback-Leibler divergence within the context of type-2 censored samples. This comparison encompassed distributions characterized by monotone decreasing, increasing, and non-monotone hazard functions, particularly for the Exponential and Normal distributions. Notably, the Tukey test exhibited superior power in the normality test. At the same time, distributions marked by monotonically increasing hazard functions displayed greater power compared to other tests in the context of the exponentiality test.

Expanding on the theme of goodness-of-fit tests, Balakrishnan et al. [9] proposed an exponentiality goodness-of-fit test grounded in Kullback-Leibler principles, specifically for progressive type-2 censored data. This test demonstrated strength, particularly in scenarios involving alternatives with non-monotonic hazard functions, as revealed through comparisons with various options.

Lim and Park [10] aim to develop a Kullback-Leibler Divergence-based information measure and goodness of fit test for working with censored datasets. In particular, they focus on how this test can be applied to cases of Type II censoring. Rad et al. [11] proposed a goodness-of-fit test for progressive type-2 censored data based on Kullback-Leibler information. For model parameters related to Pareto, Log-normal, and Weibull distributions, their analysis considered both maximum and approximate maximum likelihood estimators, evaluating the test's effectiveness over various choices and sample sizes. Gurevich and Davidson [12] show how statistical tests based on the Kullback-Leibler Divergence can be standardized to test their suitability for particular distributions.

Furthermore, Park and Pakyari [13] presented Kullback-Leibler data and conducted a comparative analysis of goodness-of-fit test results, focusing on progressive type-2 censored data. Meanwhile, Elsherpieny et al. [14] delved into the challenge of discriminating between gamma and log-logistic distributions in the context of progressive type-censored samples. They employed the minimized Kullback-Leibler divergence ratio method and the maximum likelihood ratio approach to differentiate between these two distributions. Simulation experiments were conducted to identify optimal choices, especially in cases with limited sample sizes. Additionally, asymptotic findings and selection probabilities were estimated to determine the minimal sample size required for effective discrimination.

Lastly, Bitaraf et al. [15] proposed a novel Kullback-Leibler distance test based on Verma's entropy, adding to the statistical methods used in similar research. The results regarding mean square error, critical values, and powers were examined against a few alternatives for conformity with the normal and exponential distributions. The differential entropy $H(f)$ of the random variable X with distribution function $F$and continuous density function $f$is defined as follows:

$$H(f) = -\int_{-\infty}^{\infty} f(x)\log f(x)\,dx$$

Using a novel Kullback-Leibler knowledge under the type 2 censored sample that advanced in his paper, Noughabi [16] created a general goodness of fit test. He contrasted the test's robustness under various censorship models for the exponential distribution. While Kullback-Leibler divergence has been used in the literature to build goodness of fit tests for numerous distribution families in complete sampling, this number is notably inadequate when considering both complete and censored samples. The progressive type of censored samples is the most widely used among the censored samples. This study is crucial to calculate critical values, obtain power comparisons under various alternatives, and handle goodness of fit tests for various continuous distribution families based on Kullback-Leibler divergence under complete and progressive type censored samples. Among the estimators used, the Vasicek estimator, Van Es's estimator, and Correa estimator are given by

$$HV_{mn} = \frac{1}{n}\sum_{i=1}^{m}\log\left\{\frac{n}{2m}X_{(i+m)} - X_{(i-m)}\right\}$$

$$HVE_{mn} = \frac{1}{n-m}\sum_{i=1}^{n-m}\log\left(\frac{n+1}{m}\left(X_{(i+m)} - X_{(i)}\right)\right) + \sum_{k=m}^{n}\frac{1}{k} + \log(m) - \log(n+1)$$

and

$$HC_{mn} = -\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\sum_{j=i-m}^{i+m}\left(X_{(j)} - \bar{X}_{(i)}\right)(j-i)}{n\sum_{j=i-m}^{i+m}\left(X_{(j)} - \bar{X}_{(i)}\right)^2}\right)$$

Where the window size m is a positive integer smaller than $n/2$, $X_{(i)} = X_{(1)}$ if $i < 1$, $X_{(i)} = X_{(n)}$, $i > n$, and $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ the order statistics are based on a random sample of size n. The main topic of this article is Kullback-Leibler information-based appropriateness tests for exponential and exponentiated exponential Poisson distributions. Some new goodness-of-fit tests are offered for given distributions using various entropy estimates. Next, the critical values of the suggested test statistics for different sample sizes were found using a Monte Carlo simulation. Under the best and worst bandwidth, essential values and powers are produced. Furthermore, power values have been contrasted with other options.

2. The Proposed Tests

2.1 Goodness of Fit Test for Exponentiated Exponential Distribution

The Exponentiated Exponential distribution introduced by Gupta and Kundu [17] has attracted much attention with the generalization of the Exponential distribution. Here, the exponential distribution is obtained when $\lambda = 1$ in Figure 1 and scale parameter distribution. With the exponential distribution $x > 0$, $\beta > 0$, and $\lambda > 0$, the probability density and distribution functions are as follows:

$$f(x) = \lambda\beta e^{-x\beta}\left(1 - e^{-x\beta}\right)^{\lambda-1}$$

and

$$F(x) = \left(1 - e^{-x\beta}\right)^{\lambda}$$

For a random variable with an Exponentiated Exponential distribution with λ and β parameters, its representation will be used (X∼Exponentiated Exponential(λ, β)). A random variable's expected value and variance expressed in this way are as follows:

$$E(X) = \{\psi(\lambda + 1) + C\}/\beta$$

and

$$Var(X) = \{\pi^2 - 6\psi'(\lambda + 1)/6\beta^2\}$$

2.1.1. Complete Sample Status

Let $X_1, X_2, \dots, X_n$ be independent random variables having Exponentiated Exponential(λ, β) distribution with λ and β parameters. Then, the log-likelihood function is provided by,

$$\ln L\,(\lambda, \beta) = \sum_{i=1}^{n} \ln f\,(x_i) \cong n \ln \lambda + n \ln \beta - \beta \sum_{i=1}^{n} x_i + (\lambda - 1) \sum_{i=1}^{n} \ln\left(1 - e^{-x_i \beta}\right) \qquad (2.1)$$

The hypothesis to be tested here,

H_0: The population probability distribution is Exponentiated Exponential
H_1: is not

or is provided by

$$H_0: F_0(x) = \left(1 - e^{-x\beta}\right)^{\lambda}$$

$$H_1: F_0(x) \neq \left(1 - e^{-x\beta}\right)^{\lambda}$$

The statistics to be utilized for testing the mentioned hypothesis, based on the log-likelihood function provided in (2.1), are as follows:

i. Vasicek's test (TV)

ii. VanEs' test (TVE)

iii. Correa's test (TC)

These statistics are commonly employed in the context of the hypothesis being discussed [18-20]. We reject H_0 large values TV_{mn} [1].

$$\begin{aligned} TV_{mn} &= -HV_{mn} - \frac{1}{n}\left(n \ln \lambda + n \ln \beta - \beta \sum_{i=1}^{n} x_i + (\lambda - 1) \sum_{i=1}^{n} ln\left(1 - e^{-x_i \beta}\right)\right) \\ &= -HV_{mn} - ln\,\lambda - ln\,\beta - \frac{\beta}{n} - \frac{(\lambda - 1)}{n} \sum_{i=1}^{n} ln\left(1 - e^{-x_i \beta}\right) \\ TVE_{mn} &= -HVE_{mn} - \frac{1}{n}\left(n \ln \lambda + n \ln \beta - n \sum_{i=1}^{n} x_i + (\lambda - 1) \sum_{i=1}^{n} ln\left(1 - e^{-x_i \beta}\right)\right) \\ &= -HVE_{mn} - ln\,\lambda - ln\,\beta - \frac{\beta}{n} - \frac{(\lambda - 1)}{n} \sum_{i=1}^{n} ln\left(1 - e^{-x_i \beta}\right) \\ TC_{mn} &= -HC_{mn} - \frac{1}{n}\left(n \ln \lambda + n \ln \beta - n \sum_{i=1}^{n} x_i + (\lambda - 1) \sum_{i=1}^{n} ln\left(1 - e^{-x_i \beta}\right)\right) \\ &= -HC_{mn} - ln\,\lambda - ln\,\beta - \frac{\beta}{n} - \frac{(\lambda - 1)}{n} \sum_{i=1}^{n} ln\left(1 - e^{-x_i \beta}\right) \end{aligned}$$

2.1.2 Progressively Type-II Censored Status

Let $X_{1:m:n}^{\boldsymbol{R}}, X_{2:m:n}^{\boldsymbol{R}}, \ldots, X_{m:m:n}^{\boldsymbol{R}}$ be independent random variables having Exponentiated Exponential (λ, β) distribution with λ and β parameters. Then, the log-likelihood function is given by

$$\ln L(\lambda,\beta) \propto m\ln\lambda + m\ln\beta - \beta\sum_{i=1}^{m} x_i + (\lambda-1)\sum_{i=1}^{m}\ln\left(1-e^{-x_i\beta}\right) + \sum_{i=1}^{m} R_i \ln\left(1-\left(1-e^{-x_i\beta}\right)^{\lambda}\right)$$

Maximum likelihood estimators of the λ and β parameters are obtained from the solution of the likelihood equations concerning $\hat{\lambda}$ and $\hat{\beta}$.

$$\frac{\partial \ln L(\lambda,\beta)}{\partial\lambda} = \frac{m}{\lambda} + \sum_{i=1}^{m}\ln\left(1-e^{-x_i\beta}\right) + \sum_{i=1}^{m} R_i \frac{\left(1-e^{-x_i\beta}\right)^{\lambda}\ln\left(1-e^{-x_i\beta}\right)}{\left(1-(1-e^{-x_i\beta})^{\lambda}\right)} = 0$$

$$\frac{\partial \ln L(\lambda,\beta)}{\partial\beta} = \frac{m}{\beta} - \sum_{i=1}^{m} x_i + (\lambda-1)\sum_{i=1}^{m}\frac{x_i e^{-x_i\beta}}{\left(1-e^{-x_i\beta}\right)} - \lambda\sum_{i=1}^{m} R_i \frac{\left(1-e^{-x_i\beta}\right)^{\lambda} x_i e^{-x_i\beta}}{\left(1-(1-e^{-x_i\beta})^{\lambda}\right)} = 0$$

Since these equations have no analytical solutions, they are only solved through numerical techniques and maximum likelihood estimations of their parameters. The following hypotheses are put out in this context to verify conformance to the Exponentiated Exponential distribution.

$$H_0: F_0 = \left(1-e^{-x\beta}\right)^{\lambda}$$

$$H_A: F_0 \neq \left(1-e^{-x\beta}\right)^{\lambda}$$

The suitability of the Exponentiated Exponential distribution for progressive type-censored samples is assessed by calculating the Kullback-Leibler information, which is obtained as follows:

$$I_{1\ldots m:m:n}(f:f^0) = -n\bar{H}_{1\ldots m:m:n} - \left(\begin{matrix} m\,ln\,\lambda + m\,ln\,\beta - \beta\sum_{i=1}^{m} x_i + (\lambda-1)\sum_{i=1}^{m} ln\left(1-e^{-x_i\beta}\right) \\ + \sum_{i=1}^{m} R_i\, ln\left(1-\left(1-e^{-x_i\beta}\right)^{\lambda}\right) \end{matrix}\right)$$

$$= -n\bar{H}_{1\ldots m:m:n} - m\,ln\,\lambda - m\,ln\,\beta + \beta\sum_{i=1}^{m} x_i - (\lambda-1)\sum_{i=1}^{m} ln\left(1-e^{-x_i\beta}\right)$$

$$-\sum_{i=1}^{m} R_i\, ln\left(1-\left(1-e^{-x_i\beta}\right)^{\lambda}\right)$$

The estimators are the above-mentioned by taking derivatives from the following parameters λ and β the Kullback-Leibler the aforesaid data. The test statistic below is generated by substituting the maximum likelihood estimators for the λ and β parameters. The following formula is used to get the Kullback-Leibler information test statistics for the progressive type of censored sample.

$$\mathrm{TA}(w,n,m) = -\frac{1}{n}\sum_{i=1}^{m}\log m\sum\left\{\frac{G\left(X_{(i+w:m:n)};\hat{\theta}\right) - G\left(X_{(i-w:m:n)};\hat{\theta}\right)}{X_{i+w:m:n} - X_{i-w:m:n}}\right\} + \frac{1}{n}\sum_{i=1}^{m} R_i \log\sum\left\{\frac{1-m/n}{1-G\left(X_{(i:m:n)};\hat{\theta}\right)}\right\}$$

3. Simulation Study

This section discusses the goodness of fit tests based on the Kullback-Leibler mismatch under complete and progressively censored samples for the Exponentially Expanded Distribution. TV, TVE, and TC test statistics are given in the complete sample case, and critical values with 0.05 significance levels for these tests were obtained with the simulation study. In addition, the powers of this goodness of fit test for the relevant distribution under different alternatives are obtained and compared with the famous test K-S test. Critical values are obtained for the distributions considered under progressively censored samples, and complete sample cases and power analysis are performed using different alternatives. All simulations mentioned here are performed over 10000 repetitions, considering various sample sizes. Inf values in calculations indicate cases where the critical value is numerically infinite, while NA values indicate cases where calculation is impossible. These cases reflect testing limitations encountered at specific sample sizes or bandwidths. Since the distribution of test statistics is complex for analytical evaluation, critical values are calculated using Monte Carlo simulation. The following steps are followed to determine the critical values of the proposed test statistics.

i. Create a sample of size n from the exponentially exponential distribution,
ii. Calculate the proposed statistics according to the created samples.

3.1. Critical Values for Exponentiated Exponential Distribution

Tables 1-4 present the critical values for the Exponentiated Exponential Distribution obtained through simulation.

Table 1. Critical values for the Exponentiated Exponential test in case n = 10, 20, 30, and 40

	n	**10**			**20**			**30**			**40**		
	Tests	**TV**	**TVE**	**TC**	**TV**	**TVE**	**TC**	**TV**	**TVE**	**TC**	**TV**	**TVE**	**TC**
Bandwidth (w)	**1**	NA	-2.79	-0.83	NA	-8.82	-6.75	NA	-14.90	-12.75	NA	-20.92	-18.62
	2	NA	-2.38	-0.95	NA	-8.23	-6.73	NA	-14.45	-12.82	NA	-20.43	-18.66
	3	-0.34	-2.26	-0.94	NA	-8.00	-6.71	NA	-14.17	-12.76	NA	-20.38	-18.90
	4	-0.61	-2.27	-0.99	NA	-8.09	-6.85	NA	-14.16	-12.87	NA	-20.12	-18.72
	5	0	0	0	Inf	-7.79	-6.63	NA	-14.05	-12.82	NA	-20.02	-18.72
	6	0	0	0	-6.42	-8.01	-6.89	NA	-13.74	-12.57	NA	-20.25	-18.97
	7	0	0	0	-6.55	-8.00	-6.90	NA	-13.78	-12.63	NA	-20.03	-18.84
	8	0	0	0	-6.51	-7.83	-6.73	-12.04	-13.84	-12.77	NA	-19.86	-18.71
	9	0	0	0	-6.52	-7.80	-6.69	-12.38	-13.88	-12.76	NA	-19.99	-18.88
	10	0	0	0	0	0	0	-12.36	-13.69	-12.61	Inf	-19.77	-18.70
	11	0	0	0	0	0	0	-12.36	-13.73	-12.65	-18.18	-19.84	-18.74
	12	0	0	0	0	0	0	-12.60	-13.85	-12.76	-18.47	-19.93	-18.85
	13	0	0	0	0	0	0	-12.44	-13.60	-12.58	-18.46	-19.77	-18.72
	14	0	0	0	0	0	0	-12.65	-13.72	-12.69	-18.51	-19.80	-18.78
	15	0	0	0	0	0	0	0	0	0	-18.76	-20.00	-18.97
	16	0	0	0	0	0	0	0	0	0	-18.70	-19.88	-18.83
	17	0	0	0	0	0	0	0	0	0	-18.66	-19.81	-18.79
	18	0	0	0	0	0	0	0	0	0	-18.47	-19.58	-18.54
	19	0	0	0	0	0	0	0	0	0	-18.78	-19.77	-18.79
	20 and above	0	0	0	0	0	0	0	0	0	0	0	0

Table 2. Critical values for the Exponentiated Exponential test in case n = 50, 60, and 70

	n	50			60			70		
	Tests	**TV**	**TVE**	**TC**	**TV**	**TVE**	**TC**	**TV**	**TVE**	**TC**
Bandwidth (w)	**1**	NA	-26.93	-24.53	NA	-33.52	-30.98	NA	-39.88	-37.32
	2	NA	-26.79	-25.03	NA	-33.12	-31.22	NA	-39.58	-37.56
	3	NA	-26.55	-24.91	NA	-32.89	-31.30	NA	-39.11	-37.38
	4	NA	-26.26	-24.80	NA	-32.57	-31.09	NA	-38.77	-37.18
	5	NA	-26.17	-24.85	NA	-32.61	-31.14	NA	-38.84	-37.35
	6	NA	-26.33	-25.07	NA	-32.49	-31.17	NA	-39.02	-37.64
	7	NA	-26.13	-24.87	NA	-32.51	-31.20	NA	-38.66	-37.34
	8	NA	-25.98	-24.80	NA	-32.44	-31.20	NA	-38.67	-37.37
	9	NA	-25.96	-24.82	NA	-32.47	-31.24	NA	-38.58	-37.33
	10	NA	-26.24	-25.09	NA	-32.26	-31.11	NA	-38.61	-37.40
	11	NA	-25.88	-24.74	NA	-32.41	-31.25	NA	-38.61	-37.40
	12	NA	-26.02	-24.93	NA	-32.48	-31.29	NA	-38.58	-37.42
	13	-24.09	-25.92	-24.86	NA	-32.30	-31.17	NA	-38.52	-37.40
	14	-24.27	-25.79	-24.75	NA	-32.40	-31.31	NA	-38.62	-37.48
	15	-24.67	-26.07	-25.04	Inf	-32.20	-31.15	NA	-38.82	-37.75
	16	-24.57	-25.86	-24.82	-30.66	-32.31	-31.25	NA	-38.66	-37.57
	17	-24.65	-25.91	-24.89	-30.88	-32.36	-31.35	NA	-38.45	-37.39
	18	-24.49	-25.74	-24.72	-30.83	-32.18	-31.14	-36.69	-38.57	-37.56
	19	-24.69	-25.91	-24.89	-30.93	-32.21	-31.19	-36.87	-38.41	-37.38
	20	-24.65	-25.79	-24.78	-31.10	-32.32	-31.28	-37.06	-38.50	-37.48
	21	-24.99	-26.15	-25.10	-31.00	-32.22	-31.24	-36.96	-38.33	-37.30
	22	-24.81	-25.95	-24.88	-31.16	-32.39	-31.38	-37.18	-38.44	-37.45
	23	-24.72	-25.76	-24.75	-30.84	-31.96	-31.02	-37.62	-38.83	-37.83
	24	-25.00	-25.90	-24.92	-30.87	-31.97	-30.99	-37.27	-38.46	-37.52
	26	0	0	0	-30.94	-32.04	-31.06	-37.45	-38.61	-37.66
	27	0	0	0	-31.09	-32.14	-31.15	-37.20	-38.33	-37.42
	28	0	0	0	-31.27	-32.24	-31.27	-37.01	-38.16	-37.15
	29	0	0	0	-30.80	-31.78	-30.77	-37.12	-38.20	-37.21
	30	0	0	0	0	0	0	-37.48	-38.46	-37.52
	31	0	0	0	0	0	0	-37.67	-38.65	-37.71
	32	0	0	0	0	0	0	-37.64	-38.62	-37.63
	33	0	0	0	0	0	0	-37.34	-38.31	-37.32
	34	0	0	0	0	0	0	-37.49	-38.35	-37.37
	35 and above	0	0	0	0	0	0	0	0	0

Table 3. Critical values for the Exponentiated Exponential test in case n = 80, 90, and 100

	n	80			90			100		
	Tests	**TV**	**TVE**	**TC**	**TV**	**TVE**	**TC**	**TV**	**TVE**	**TC**
Bandwidth (w)	**1**	NA	-46.19	-43.54	NA	-52.49	-49.79	NA	-59.27	-56.53
	2	NA	-45.69	-43.69	NA	-52.31	-50.20	NA	-58.67	-56.56
	3	NA	-45.70	-43.89	NA	-51.87	-50.05	NA	-58.41	-56.53
	4	NA	-45.41	-43.74	NA	-51.78	-50.11	NA	-58.07	-56.35
	5	NA	-45.26	-43.68	NA	-51.82	-50.25	NA	-58.32	-56.71
	6	NA	-45.43	-43.94	NA	-51.79	-50.26	NA	-57.81	-56.27
	7	NA	-45.37	-43.99	NA	-51.85	-50.37	NA	-58.22	-56.73
	8	NA	-45.12	-43.77	NA	-51.57	-50.18	NA	-58.25	-56.82
	9	NA	-44.97	-43.68	NA	-51.35	-50.02	NA	-57.80	-56.42
	10	NA	-44.77	-43.51	NA	-51.43	-50.16	NA	-57.74	-56.37
	11	NA	-45.54	-44.28	NA	-51.30	-50.04	NA	-57.79	-56.51
	12	NA	-44.91	-43.69	NA	-51.38	-50.11	NA	-57.89	-56.58
	13	NA	-45.02	-43.86	NA	-51.31	-50.09	NA	-57.92	-56.66
	14	NA	-44.78	-43.61	NA	-51.44	-50.26	NA	-57.88	-56.64
	15	NA	-45.17	-44.00	NA	-51.27	-50.07	NA	-57.73	-56.50
	16	NA	-44.81	-43.69	NA	-51.16	-50.01	NA	-58.01	-56.81
	17	NA	-44.92	-43.84	NA	-51.54	-50.42	NA	-57.99	-56.84
	18	NA	-44.53	-43.47	NA	-51.09	-49.96	NA	-57.62	-56.47
	19	NA	-44.74	-43.71	NA	-51.16	-50.11	NA	-57.85	-56.72
	20	Inf	-45.12	-44.07	NA	-51.31	-50.20	NA	-57.61	-56.49
	21	-43.1	-44.9	-43.8	NA	-51.11	-50.03	NA	-57.65	-56.58
	22	-43.1	-44.7	-43.6	NA	-51.19	-50.12	NA	-57.85	-56.76
	23	-43.4	-44.8	-43.8	-49.31	-51.26	-50.20	NA	-57.67	-56.60
	24	-43.4	-44.8	-43.8	-49.87	-51.52	-50.47	NA	-57.67	-56.60
	26	-43.5	-44.8	-43.8	-49.76	-51.20	-50.17	NA	-57.98	-56.98
	27	-43.8	-45.0	-44.0	-49.69	-51.09	-50.09	-55.77	-57.51	-56.44
	28	-43.5	-44.7	-43.7	-49.84	-51.18	-50.16	-56.12	-57.68	-56.68
	29	-43.2	-44.4	-43.4	-49.49	-50.76	-49.76	-56.48	-57.92	-56.92
	30	-43.5	-44.7	-43.7	-50.13	-51.34	-50.36	-56.29	-57.70	-56.67
	31	-43.5	-44.7	-43.7	-50.26	-51.40	-50.41	-56.63	-57.93	-56.94
	32	-43.5	-44.6	-43.6	-49.85	-51.01	-50.04	-56.63	-57.90	-56.91
	33	-43.6	-44.7	-43.7	-49.97	-51.19	-50.20	-56.18	-57.39	-56.40
	34	-44.0	-45.0	-44.1	-49.94	-51.09	-50.13	-56.27	-57.43	-56.48
	35	-43.7	-44.7	-43.7	-49.70	-50.88	-49.87	-56.19	-57.35	-56.38
	36	-43.7	-44.8	-43.8	-50.16	-51.25	-50.32	-56.72	-57.89	-56.92
	37	-43.7	-44.6	-43.6	-50.14	-51.19	-50.24	-56.36	-57.49	-56.54
	38	-43.9	-44.7	-43.8	-50.04	-51.10	-50.14	-56.34	-57.44	-56.52
	39	-43.9	-44.7	-43.7	-49.91	-50.93	-49.96	-56.31	-57.37	-56.45
	40	0	0	0	-50.16	-51.09	-50.17	-56.54	-57.55	-56.63
	41	0	0	0	-50.49	-51.39	-50.44	-56.46	-57.46	-56.55
	42	0	0	0	-50.08	-51.05	-50.05	-56.62	-57.66	-56.69
	43	0	0	0	-50.29	-51.14	-50.21	-56.67	-57.64	-56.69
	44	0	0	0	-50.29	-51.11	-50.12	-56.71	-57.65	-56.71
	45	0	0	0	0	0	0	-56.63	-57.58	-56.64
	46	0	0	0	0	0	0	-56.57	-57.42	-56.55
	47	0	0	0	0	0	0	-56.84	-57.72	-56.77
	48	0	0	0	0	0	0	-56.78	-57.57	-56.63
	49	0	0	0	0	0	0	-56.85	-57.64	-56.65

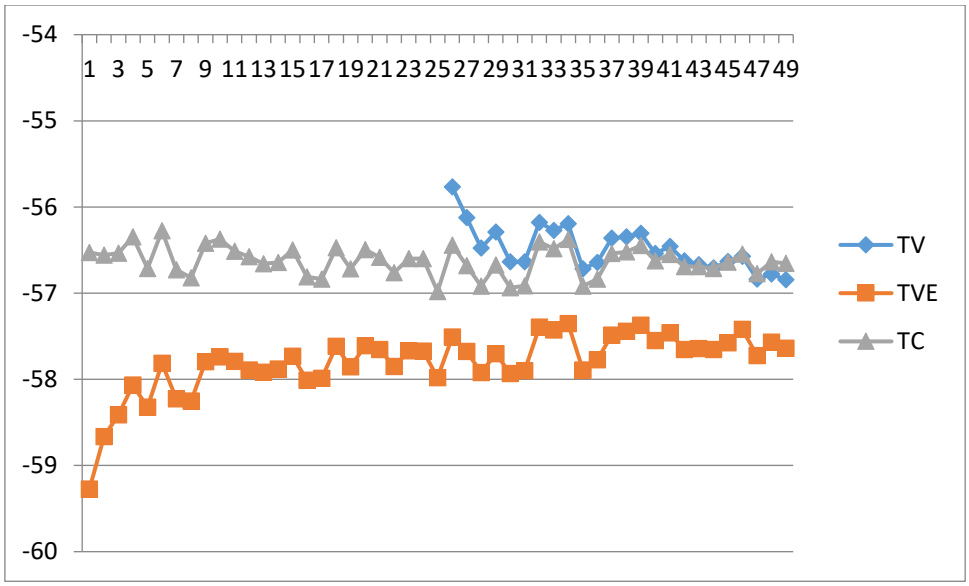


Figure 1. Critical values with 0.05 significance level for tests for goodness to Exponentiated Exponential distribution versus bandwidth for n = 100

Observing the graphs of critical values for n = 100 in Figure 1, it becomes evident that as the bandwidth w increases, the critical values for the TVE test remain relatively constant. On the other hand, the critical values for the TV test tend to stabilize as they decrease, and the critical values for the TC test show an increasing trend in stability.

3.2. Exponentiated Exponential under Complete Sample

This section analyzes the best and worst powers concerning the choice of bandwidth w under various distributions and parameter scenarios, assuming that the true distribution is the Exponentiated Exponential distribution.

Table 5. If the true distribution is $Weibull(5,3)$, the best bandwidth w and corresponding powers (Complete Sample)

n	K-S	TV			TVE			TC		
	Power	w	Critical Values (%95)	Power	w	Critical Values (%95)	Power	w	Critical Values (%95)	Power
10	0	4	-0.53472	0.0057	1	-2.80384	0.0064	1	-0.87421	0.0063
20	0.0025	6	-6.5173	0.0015	6	-8.0926	0.0015	6	-6.98912	0.0016
30	0.01	8	-12.1032	4.00E-04	4	-14.0371	6.00E-04	4	-12.7877	6.00E-04
40	0.0193	14	-18.4543	2.00E-04	4	-20.2122	2.00E-04	4	-18.8493	2.00E-04
50	0.0338	20	-24.7359	1.00E-04	3	-26.3433	1.00E-04	10	-24.9431	1.00E-04
60	0.0563	1	NA	0	1	-33.5308	0	1	-30.9908	0
70	0.0777	1	NA	0	1	-39.903	0	1	-37.2755	0
80	0.1141	1	NA	0	18	-44.9686	1.00E-04	18	-43.9219	1.00E-04
90	0.1491	1	NA	0	1	-52.6165	0	1	-49.8152	0
100	0.1678	1	NA	0	1	-59.2637	0	1	-56.5043	0

Table 6. If the true distribution is $Weibull(5,3)$, worst bandwidth w and corresponding powers (Complete Sample)

n	TV			TVE			TC		
	w	Critical Values (%95)	Power	w	Critical Values (%95)	Power	w	Critical Values (%95)	Power
10	2	NA	0	4	-2.16067	0.0047	4	-0.87718	0.0048
20	5	Inf	0	9	-7.74306	0.0012	1	-6.56822	0.0011
30	7	NA	0	14	-13.4926	4.00E-04	14	-12.4562	4.00E-04
40	19	-18.7924	0	19	-19.7692	0	19	-18.7544	0
50	24	-25.0013	0	24	-25.9611	0	24	-24.9456	0
60	29	-31.2907	0	29	-32.1708	0	29	-31.2009	0
70	34	-37.6988	0	34	-38.5237	0	34	-37.5905	0
80	39	-43.9757	0	39	-44.7696	0	39	-43.8098	0
90	44	-50.1323	0	44	-50.9229	0	44	-49.9419	0
100	49	-56.5206	0	49	-57.2733	0	49	-56.3515	0

It was understood that TV, TVE, and TC tests could not be performed against the worst w, and they did not have power; this situation is somewhat important for selecting w for TV, TVE, and TC tests. When Tables 5 and 6 are examined for testing the suitability of $Weibull(5,3)$ the population with A distribution to the Exponentiated Exponential distribution, the power of TV, TVE, and TC tests is almost equal to the best bandwidth w, and the power of K-S test is higher than TV, TVE, and TC tests.

Table 7. If the true distribution is $EG(0.4,2)$, worst bandwidth w and corresponding powers (Complete Sample)

n	K-S	TV			TVE			TC		
	Power	w	Critical Values (%95)	Power	w	Critical Values (%95)	Power	w	Critical Values (%95)	Power
10	4.00E-04	3	-0.33979	0.839	4	-2.26843	0.8384	1	-0.83	0.8561
20	0.001	7	-6.54705	0.987	7	-8.00128	0.9852	1	-6.75347	0.9888
30	0.0034	9	-12.3847	0.999	5	-14.0542	0.9989	3	-12.7581	0.9995
40	0.0042	11	-18.1826	1	7	-20.0335	1	2	-18.6572	1
50	0.0036	13	-24.0919	1	1	-26.932	1	1	-24.5333	1
60	0.006	16	-30.6575	1	1	-33.5186	1	1	-30.9845	1
70	0.0059	18	-36.6944	1	1	-39.8808	1	1	-37.3195	1
80	0.0096	21	-43.1355	1	1	-46.1896	1	1	-43.5357	1
90	0.0115	23	-49.3148	1	1	-52.494	1	1	-49.7867	1
100	0.0114	26	-55.7665	1	1	-59.2729	1	1	-56.5291	1

Table 8. If the true distribution is $EG(0.4,2)$, the best bandwidth w and corresponding powers (Complete Sample)

n	TV			TVE			TC		
	w	Critical Values (%95)	Power	w	Critical Values (%95)	Power	w	Critical Values (%95)	Power
10	2	NA	0	1	-2.84598	0.8183	4	-0.9794	0.8428
20	5	Inf	0	1	-8.64008	0.9788	9	-6.64341	0.985
30	7	NA	0	1	-14.6716	0.9984	14	-12.5628	0.9991
40	10	Inf	0	4	-20.3324	0.9995	19	-18.9264	0.9997
50	12	NA	0	24	-25.8575	1	24	-24.8451	1
60	15	Inf	0	29	-32.1584	1	29	-31.2042	1
70	17	NA	0	34	-38.7469	1	34	-37.7665	1
80	20	Inf	0	39	-44.4957	1	39	-43.535	1
90	22	NA	0	44	-51.0945	1	44	-50.1238	1
100	25	Inf	0	49	-57.6001	1	49	-56.6904	1

Upon examining Tables 7 and 8 for assessing the suitability of a population with an actual distribution to the Exponentiated Exponential Distribution, the following observations can be made:

The powers of the TV, TVE, and TC tests when using the best bandwidth are nearly equal, and they all surpass the power of the K-S (Kolmogorov-Smirnov) test. Among these, the TC test exhibits slightly superior performance.

It is worth noting that the TV test cannot be executed effectively when employing the worst bandwidth w, emphasizing the significance of carefully selecting w for the TV test.

The power of both the TVE and TC tests experiences a decline as the sample size is reduced to n = 10, 20, 30, and 40, particularly when the worst bandwidth w is chosen. Thus, the choice of bandwidth w for these tests becomes somewhat crucial in such scenarios.

3.3. Exponentiated Exponential under Censored Sample

The Exponentiated Exponential test critical values, derived from simulation, are utilized to investigate the best and worst powers for different distributions and parameter settings, contingent on the bandwidth *w* selection. We used the Noughabi [21] censoring scheme. This is assuming that the real distribution follows an exponential distribution. Tables 9 and 10 present the findings from these analyses.

Table 9. Best bandwidth w and corresponding power obtained for different alternative distributions (Progressive Type Censored Sample)

Censoring Scheme No	Weibull(5.3)			EP(2.2)			EG(0.4.2)		
	w	Critical value (%95)	Power	w	Critical value (%95)	Power	w	Critical value (%95)	Power
1	1	-0.1517	0.1173	3	-0.2720	0.1197	3	-0.2720	0.2820
2	1	0.1370	0.0443	1	0.1370	0.0526	2	0.0703	**0.5159**
3	1	0.0892	0.0439	2	0.0221	0.0549	2	0.0221	0.5543
4	2	-0.2738	0.2276	5	-0.3449	0.1357	5	-0.3449	0.2479
5	1	0.2040	0.0577	2	0.1004	0.0612	2	0.1004	**0.6613**
6	4	-0.1666	0.2915	2	-0.1252	0.0787	5	-0.1584	0.1429
7	3	-0.2442	0.3088	7	-0.3110	0.1413	7	-0.3110	0.1987
8	1	0.2758	0.0905	3	0.0957	0.0850	2	0.1247	**0.6050**
9	7	-0.2472	0.3690	3	-0.1907	0.0998	3	-0.1907	0.0705
10	2	-0.1689	0.1990	4	-0.1947	0.1377	4	-0.1947	0.4651
11	1	0.0925	0.0460	2	0.0478	0.0529	2	0.0478	**0.7259**
12	1	0.0667	0.0490	1	0.0667	0.0555	2	0.0236	0.4914
13	3	-0.2909	**0.4518**	8	-0.3313	**0.1798**	8	-0.3313	0.3833
14	1	0.1828	0.0633	2	0.0884	0.0643	2	0.0884	**0.9376**
15	1	0.0850	0.1099	3	-0.0447	0.1005	3	-0.0447	0.1153
16	4	-0.2862	**0.5980**	13	-0.3383	**0.1832**	13	-0.3383	0.2810
17	1	0.2646	0.1171	7	0.0624	0.0991	3	0.0903	**0.9425**
18	1	0.1618	0.1611	5	-0.0351	0.1152	14	-0.0726	0.8783
19	3	-0.2291	**0.5032**	8	-0.2565	**0.1760**	8	-0.2565	0.5211
20	1	0.1216	0.0648	3	0.0475	0.0599	2	0.0617	**0.9717**
21	1	0.0791	0.0749	3	0.0053	0.0627	2	0.0186	0.6923
22	4	-0.3102	**0.7423**	16	-0.3577	**0.2134**	17	-0.3584	0.3614
23	1	0.2345	0.1196	6	0.0599	0.0864	3	0.0814	**0.9965**
24	4	-0.1055	0.3736	5	-0.1155	0.1557	5	-0.1155	0.1297
25	5	-0.2456	**0.8172**	21	-0.2939	**0.2190**	22	-0.2943	0.2748
26	1	0.2839	0.2222	8	0.0597	0.1315	4	0.0775	**0.9835**
27	2	0.0381	0.3167	8	-0.0449	0.1762	24	-0.0725	0.9340

Table 10. Worst bandwidth w and corresponding power obtained for different alternative distributions (Progressive Type Censored Sample)

Censoring Scheme No	Weibull(5.3)			EP(2.2)			EG(0.4.2)		
	w	Critical value (%95)	Power	w	Critical value (%95)	Power	w	Critical value (%95)	Power
1	3	-0.2720	0.0551	1	-0.1517	0.0593	1	-0.1517	0.1298
2	3	0.0343	0.0260	3	0.0343	0.0416	1	0.1370	**0.4659**
3	3	-0.0111	0.0194	3	-0.0111	0.0508	1	0.0892	0.4853
4	5	-0.3449	0.0832	1	-0.1355	0.0515	1	-0.1355	0.0734
5	4	0.0523	0.0137	5	0.0326	0.0527	1	0.2040	**0.5474**
6	1	0.0138	0.1423	5	-0.1584	0.0310	2	-0.1252	0.0697
7	7	-0.3110	0.2055	1	-0.0255	0.0520	1	-0.0255	0.0603
8	6	0.0474	0.0044	1	0.2758	0.0591	1	0.2758	**0.4321**
9	1	0.0216	0.2126	7	-0.2472	0.0321	7	-0.2472	0.0489
10	4	-0.1947	0.0391	1	-0.1092	0.0609	1	-0.1092	0.1846
11	3	0.0304	0.0309	4	0.0141	0.0467	1	0.0925	**0.6450**
12	4	-0.0085	0.0202	4	-0.0085	0.0450	1	0.0667	0.4309
13	9	-0.3290	**0.0120**	1	-0.1521	**0.0507**	1	-0.1521	0.0841
14	7	0.0309	0.0129	9	0.0179	0.0364	1	0.1828	**0.8564**
15	9	-0.0938	0.0091	1	0.0850	0.0554	9	-0.0938	0.0501
16	14	-0.3366	**0.0570**	1	-0.0700	**0.0499**	1	-0.0700	0.0605
17	11	0.0384	0.0014	1	0.2646	0.0557	1	0.2646	**0.8201**
18	14	-0.0726	0.0004	1	0.1618	0.0629	1	0.1618	0.6458
19	9	-0.2543	**0.0053**	1	-0.1362	**0.0479**	1	-0.1362	0.1178
20	6	0.0281	0.0174	9	0.0100	0.0286	1	0.1216	**0.9276**
21	8	-0.0274	0.0090	9	-0.0325	0.0416	1	0.0791	0.5432
22	19	-0.3561	**0.0043**	1	-0.1252	**0.0528**	1	-0.1252	0.0660
23	16	0.0267	0.0007	19	0.0229	0.0470	1	0.2345	**0.9719**
24	19	-0.1494	0.0063	19	-0.1494	0.0397	18	-0.1502	0.0550
25	24	-0.2933	**0.1099**	1	-0.0023	**0.0478**	1	-0.0023	0.0527
26	24	0.0282	0.0000	1	0.2839	0.0655	1	0.2839	**0.8965**
27	24	-0.0725	0.0000	1	0.1856	0.0710	1	0.1856	0.6935

Examining Tables 9 and 10 will help determine whether a population with an $Weibull(5,3)$ $EP(2,2)$ or $EG(0.4,2)$ distribution is suitable for the exponential distribution. The censoring systems for the $Weibull(5,3)$ distribution show notable shifts in the authorities. When the initial observation is made, the level of censorship in the $R_m = 12$ and $R_i = 0, i \neq 12$-shaped censorship schemes is higher, outcomes that are comparable to the distribution $Weibull(5,3)$ are also obtained for distribution $EP(2,2)$, the powers for the $EP(2,2)$ dispersion are typically minimal, for the $EG(0.4,2)$ distribution, The censoring schemes indicate substantial shifts in the authorities. It has been established that the choice of bandwidth w is important since there is a difference between the distribution's best and worst powers. The powers are more significant in the $R_m = 20, R_i = 0$,i=1,2, ... ,m-1--shaped censoring schemes where all censorship is made after the last observation is collected.

3.4. Real Data Analysis in Bladder Cancer

This section demonstrates the proposed method with a real data set. Since this real dataset is publicly available, it does not require ethics committee approval. The actual data set is from Abbas et al. [22]. The proposed dataset representing recovery times (in months) of a random sample of 128 bladder cancer patients was used in Abbas et al. [22] (n=128, m=20).

The Real dataset shows the number of months that 128 patients with bladder cancer went into remission.

0.08	0.2	0.4	0.5	0.51	0.81	0.9	1.05	1.19	1.26	1.35	1.4	1.46	1.76	2.02	2.02
2.07	2.09	2.23	2.26	2.46	2.54	2.62	2.64	2.69	2.69	2.75	2.83	2.87	3.02	3.25	3.31
3.36	3.36	3.48	3.52	3.57	3.64	3.7	3.82	3.88	4.18	4.23	4.26	4.33	4.34	4.4	4.5
4.51	4.87	4.98	5.06	5.09	5.17	5.32	5.32	5.34	5.41	5.41	5.49	5.62	5.71	5.85	6.25
6.31	6.54	6.76	6.93	6.94	6.97	7.09	7.26	7.28	7.32	7.39	7.59	7.62	7.63	7.66	7.87
7.93	8.26	8.37	8.53	8.65	8.66	9.02	9.22	9.47	9.74	10.06	10.34	10.66	10.75	11.25	11.64
11.79	11.98	12.02	12.03	12.07	12.63	13.11	13.29	13.8	14.24	14.76	14.77	14.83	15.96	16.62	17.12
17.14	17.36	18.1	19.13	20.28	21.73	22.69	23.63	25.74	25.82	32.15	34.26	36.66	43.01	46.12	79.05

Censored Data-1, m = 20

0.08	0.2	0.4	0.5	0.51	0.81	0.9	1.05	1.19	1.26	1.35	1.4	1.46
1.76	2.02	2.02	2.07	2.09	2.23	2.26						

We test exponentiation using the proposed procedure.

i. The value of the test statistics for w = 2 are calculated as follows: TA(w = 2) = −0.8017, normally 0.0981 are critical values corresponding to 0.05.

ii. The value of the test statistics for w = 3 are calculated as follows: TA(w = 3) = -2.3271, normally 0.0867 are critical values corresponding to 0.05.

iii. The value of the test statistics for w = 6 are calculated as follows: TA(w = 6) = −1.9900, normally 0.0607 are critical values corresponding to 0.05.

iv. The value of the test statistics for w = 9 are calculated as follows: TA(w = 9) = −0.4562, normally 0.0502 are critical values corresponding to 0.05.

Therefore, the hypothesis that Type-II censored people are from an exponentiated exponential distribution is accepted at the 0.05 significance level. Test statistics and corresponding values were calculated on censored datasets via R software.

4. Conclusion

In conclusion, utilizing the best and worst bandwidths, comparing critical values and power was performed under different conditions, based on Kullback-Leibler divergence, for both complete and censored sampling scenarios. It was shown that the test's power rose as the sample size grew in all comparisons. The differences between the best and worst bandwidths in most cases highlight the importance of the bandwidth selection (denoted as 'w') while addressing the goodness-of-fit testing problem.

Furthermore, it became evident that when bandwidth 'w' is generally selected for the TV test, the test cannot be effectively performed, highlighting the pivotal role of 'w' in this particular test. Among the tests with bandwidth comparisons, the TC test consistently emerged as the strongest, followed by the TV test, with the TVE test showing the lowest performance.

As a result of numerous comparisons, it can be concluded that the choice of bandwidth 'w' holds considerable importance, as substantial differences in power were observed between the best and worst bandwidths in nearly all cases.

To demonstrate the usefulness of our suggested research, we used a dataset of people who had been diagnosed with bladder cancer in the last phase of our investigation. Ultimately, models like this can help academics and medical practitioners with tasks like risk prediction, investigation of events, medical decision-making, and time-to-event analysis (survival analysis).

Author Contributions

The author read and approved the final version of the paper.

Conflict of Interest

The author declares no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, A. Jemal, *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA: A Cancer Journal for Clinicians 68 (6) (2018) 394-424.

[2] R. B. D'agostino, M. A. Stephens, Goodness of fit techniques. New York, MDI Press, 1986.

[3] K. Pearson, *On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 50 (302) (1900) 157-175.

[4] I. Arizono, H. Ohta, *A Test for normality based on Kullback-Leibler information*, The American Statistician 43 (1989) 20-22.

[5] B. Şenoğlu, B. Sürücü, *Goodness-of-fit tests based on Kullback-Leibler information*, IEEE Transactions on Reliability 53 (2004) 357-361.

[6] B. Choi, K. Kim, S. Heun Song, *Goodness-of-fit test for exponentiality based on Kullback–Leibler information*, Communications in Statistics - Simulation and Computation 33 (2004) 525-536.

[7] S. Park, *Testing exponentiality based on the Kullback-Leibler information with the type II censored data*, IEEE Transactions on Reliability 54 (2005) 22-26.

[8] J. Lim, S. Park, *A goodness of fit tests based on the partial Kullback-Leibler information with the type II censored data*, Journal of Applied Statistics 34 (2007) 1051-1064.

[9] N. Balakrishnan, A. H. Rad, N. R. Arghami, *Testing exponentiality based on Kullback-Leibler information with progressively type-II censored data*, IEEE Transactions on Reliability 56 (2007) 301-307.

[10] J. Lim, S. Park, *Censored Kullback-Leibler information and goodness-of-fit test with type II censored data*, Quality control and applied statistics 54 (1) (2009) 79-80.

[11] A. H. Rad, F. Yousefzadeh, N. Balakrishnan, *Goodness-of-fit test based on Kullback-Leibler information for progressively type-II censored data*, IEEE Transactions on Reliability 60 (2011) 570-579.

[12] G. Gurevich, A. Davidson, Standardized forms of Kullback-Leibler information based statistics for normality and exponentiality. Computer Modelling and New Technologies 12(1) (2008) 14-25.

[13] S. Park, R. Pakyari, *Cumulative residual Kullback–Leibler information with the progressively Type-II censored data*, Statistics & Probability Letters 106 (2015) 287-294.

[14] E. Elsherpieny, H. Muhammed, N. Radwan, *On Discriminating between gamma and log-logistic distributions in case of progressive type-II censoring*, Pakistan Journal of Statistics and Operation Research 13 (2017) 157.

[15] M. Bitaraf, M. Rezaei, F. Yousefzadeh, *Goodness-of-fit tests based on Verma Kullback–Leibler information*, Communications in Statistics - Theory and Methods 46 (24) (2017) 12152-12164.

[16] H. A. Noughabi, *Testing exponentiality based on Kullback—Leibler information for progressively type II censored data*, Communications in Statistics - Simulation and Computation 46 (2017) 7624-7638.

[17] R. D. Gupta D. Kundu, *Theory & methods: Generalized exponential distributions*, The Australian & New Zealand Journal of Statistics 41 (1999) 173-188.

[18] J. C. Correa, *A new estimator of entropy*, Communications in Statistics - Theory and Methods 24 (1995) 2439-2449.

[19] B. Van Es, *Estimating functionals related to a density by a class of statistics based on spacings*, Scandinavian Journal of Statistics 19 (1992) 61-72.

[20] O. Vasicek, *A test for normality based on sample entropy*, Journal of the Royal Statistical Society. Series B (Methodological) 38 (1976) 54-59.

[21] H. A. Noughabi, *Testing exponentiality based on Kullback—Leibler information for progressively Type II censored data*, Communications in Statistics-Simulation and Computation 46 (10) (2017) 7624-7638.

[22] K. Abbas, Z. Hussain, N. Rashid, A. Ali, M. Taj, S. A. Khan, S. Manzoor, U. Khalil, D. M. Khan, *Bayesian estimation of Gumbel type-II distribution under type-II censoring with medical applications*, Hindawi Computational and Mathematical Methods in Medicine Article ID 1876073 (2020) 11 pages.

Journal of New Results in Science

https://dergipark.org.tr/en/pub/jnrs

Research Article

Open Access

E-ISSN:1304-7981 https://doi.org/10.54187/jnrs.1506171

Pell Leonardo numbers and their matrix representations

Çağla Çelemoğlu[1]

Keywords:
Leonardo numbers, Pell numbers, Binet's formula, Generating function, Matrix representation

Abstract — In this study, we investigate Pell numbers and Leonardo numbers and describe a new third-order number sequence entitled Pell Leonardo numbers. We then construct some identities, including the Binet formula, generating function, exponential generating function, Catalan, Cassini, and d'Ocagne's identities for Pell Leonardo numbers and obtain a relation between Pell Leonardo and Pell numbers. In addition, we present some summation formulas of Pell Leonardo numbers based on Pell numbers. Finally, we create a matrix formula for Pell Leonardo numbers and obtain the determinant of the matrix.

Subject Classification (2020): 11B39, 11Y55

1. Introduction

There are many famous number sequences whose elements are integers. The most well-known of these is the Fibonacci sequence. Different generalizations of the Fibonacci sequence have been made in the literature. One of these is the Pell sequence. In 1975, Bicknell [1] described the primer of the Pell sequence by the following recurrence relation: For $n \geq 2$ ($n \in \mathbb{N}$),

$$P_n = 2P_{n-1} + P_{n-2} \tag{1.1}$$

and the initial conditions $P_0 = 0$ and $P_1 = 1$. The characteristic equation of (1.1) is

$$\vartheta^2 - 2\vartheta - 1 = 0 \tag{1.2}$$

which has roots $\alpha = 1 + \sqrt{2}$ and $\beta = 1 - \sqrt{2}$. Binet's formula for the P_n is

$$P_n = \frac{\alpha^n - \beta^n}{\alpha - \beta} = \frac{\alpha^n - \beta^n}{2\sqrt{2}} \tag{1.3}$$

A lot of research has been done on Pell numbers. Some of these are as follows: In [2], Horadam studied modified Pell numbers and their applications. In [3], Melham obtained new identities of Pell numbers, and in [4], Santana and Diaz-Barrero gave new properties about the summation of Pell numbers. In [5], Mushtaq and Hayat presented a matrix with Pell numbers of entries. In [6], Dasdemir built new matrices based on Pell

[1]cagla.ozyilmaz@omu.edu.tr (Corresponding Author)
[1]Department of Mathematics, Faculty of Science, Ondokuz Mayıs University, Samsun, Türkiye
Article History: Received: 27 Jun 2024 — Accepted: 19 Aug 2024 — Published: 31 Aug 2024

numbers and modified Pell and Pell-Lucas numbers. Çelik et al. [7] obtained new recurrences on Pell and Pell-Lucas numbers.

In this study, we are interested in another well-known sequence closely related to the Fibonacci sequence and the Leonardo sequence. Catarino and Borges [8] presented some properties of the Leonardo sequence that they described by the following recurrence relation: For $n \geq 2$,

$$Le_n = Le_{n-1} + Le_{n-2} + 1$$

and the boundary conditions $Le_0 = Le_1 = 1$. In addition, there is an equation between the Leonardo numbers as follows: For $n \geq 2$,

$$Le_{n+1} = 2Le_n - Le_{n-2} \tag{1.4}$$

The characteristic equation of (1.4) is

$$\mu^3 - 2\mu^2 + 1 = 0 \tag{1.5}$$

The Binet's formula of the Le_n number is

$$Le_n = \frac{2\delta^{n+1} - 2\theta^{n+1} - \delta + \theta}{\delta - \theta}$$

δ and θ are roots of (1.5). There are many more works on Leonardo's numbers in literature. Some of these are as follows: In [9], Shannon found an inhomogeneous extension of Leonardo's numbers. Alp and Koçer [10] obtained new identities of Leonardo numbers, and in [11], they introduced hybrid Leonardo numbers. Kuhapatanakul and Chobsorn [12] defined generalized Leonardo numbers and obtained matrix representations. In [13], Karatas introduced complex Leonardo numbers and gave some properties and summation formulas. Tan and Leung [14] introduced Leonardo p −numbers and incomplete Leonardo p −numbers. In [15], Soykan described generalized Horadam Leonardo numbers and gave Binet's formulas, generating functions, Simson formulas, and the summation formulas.

Inspired by these studies, in this article, we first aim to define a new third-order sequence we named Pell Leonardo numbers. Thus, we introduce a new type of Leonardo sequence. Additionally, we intend to find some of their equations, including the Binet formula, generating function, exponential generating function, Catalan, Cassini, and d'Ocagne's identities for Pell Leonardo numbers. We obtain some summation formulas for Pell Leonardo numbers based on Pell numbers. Finally, we build a matrix representation for Pell Leonardo numbers and get the determinant of the matrix.

2. Pell Leonardo Numbers

In this section, we describe nth Pell Leonardo number by the following recurrence relation for $n \geq 2$,

$$PLe_n = 2PLe_{n-1} + PLe_{n-2} + 1 \tag{2.1}$$

and the initial conditions $PLe_0 = 0$ and $PLe_1 = 1$. The first fifteen Pell Leonardo numbers are

$$0, 1, 3, 8, 20, 49, 119, 288, 696, 1681, 4059, 9800, 23660, 57121, 137903$$

In addition, we can mention the terms with negative subscripts of Pell Leonardo number sequences. By using (2.1), the first five terms with negative subscripts of Pell Leonardo numbers are as follows:

$$PLe_{-1} = 0, PLe_{-2} = -1, PLe_{-3} = 1, PLe_{-4} = -4, \text{and } PLe_{-5} = 8$$

And then, according to (2.1), it is observed that

$$PLe_{n+1} = 3PLe_n - PLe_{n-1} - PLe_{n-2} \tag{2.2}$$

where PLe_n is nth Pell Leonardo number.

The characteristic equation of (2.2) is

$$\gamma^3 - 3\gamma^2 + \gamma + 1 = 0$$

The relation between Pell Leonardo and Pell numbers is expressed in the following proposition.

Proposition 2.1. For $n \geq 0$,

$$PLe_n - PLe_{n-1} = P_n \tag{2.3}$$

Proof. We prove by induction on n. We know that from the definition of Pell Leonardo numbers, $PLe_{-1} = 0$. It is easily seen that (2.3) is held for $n = 0$ and $n = 1$. We suppose that (2.3) is true for all $1 < l \leq n$, and we prove that (2.3) holds for $l = n + 1$. In fact, by using the induction hypothesis and the recurrence relation (2.2), we can write

$$\begin{aligned} PLe_{n+1} &= 3PLe_n - PLe_{n-1} - PLe_{n-2} \\ &= 2PLe_n - PLe_{n-1} - PLe_{n-2} + PLe_n \\ &= 2(PLe_n - PLe_{n-1}) + PLe_{n-1} - PLe_{n-2} + PLe_n \\ &= 2P_n + P_{n-1} + PLe_n \end{aligned}$$

and thus,

$$PLe_{n+1} - PLe_n = P_{n+1}$$

□

Theorem 2.2. The Binet formula for Pell Leonardo numbers is given by

$$PLe_n = \frac{1}{4}(\rho^{n+1} + \sigma^{n+1}) - \frac{1}{2}$$

where $\rho = 1 + \sqrt{2}$ and $\sigma = 1 - \sqrt{2}$.

Proof. Because the roots of (1.2) are equal to ρ and σ, Binet's formula of the Pell Leonardo numbers can be easily calculated by using (1.3) and (2.3). □

Theorem 2.3. The generating function of PLe_n is determined by

$$GF_{PLe}(t) = \frac{t}{(1 - 3t + t^2 + t^3)}$$

Proof. For the generating function of $\{PLe_n\}_{n=0}^{\infty}$. Firstly, we will consider the following power series:

$$GF_{PLe}(t) = \sum_{n=0}^{\infty} PLe_n t^n$$

Therefore,

$$GF_{PLe}(t) = PLe_0 + PLe_1 t + PLe_2 t^2 + \cdots + PLe_k t^k + \cdots$$

Thus,

$$\begin{aligned}
-3tGF_{PLe}(t) &= -3PLe_0t - 3PLe_1t^2 - 3PLe_2t^3 - \cdots - 3PLe_kt^{k+1} + \cdots \\
+t^2GF_{PLe}(t) &= PLe_0t^2 + PLe_1t^3 + PLe_2t^4 + \cdots + PLe_kt^{k+2} + \cdots \\
+t^3GF_{PLe}(t) &= PLe_0t^3 + PLe_1t^4 + PLe_2t^5 + \cdots + PLe_kt^{k+3} + \cdots
\end{aligned}$$

If the three equations above are considered together, we obtain

$$\begin{aligned}
(1 - 3t + t^2 + t^3)GF_{PLe}(t) &= PLe_0 + (PLe_1 - 3PLe_0)t + (PLe_2 - 3PLe_1 + PLe_0)t^2 \\
&\quad +(PLe_3 - 3PLe_2 + PLe_1 + PLe_0)\,t^3 \\
&\quad + \cdots + (PLe_{k+1} - 3PLe_k + PLe_{k-1} + PLe_{k-2})t^{k+1} + \cdots
\end{aligned}$$

By using (2.2), we have

$$\begin{aligned}
GF_{PLe}(t) &= \frac{PLe_0 + (PLe_1 - 3PLe_0)t + (PLe_2 - 3PLe_1 + PLe_0)t^2}{(1 - 3t + t^2 + t^3)} \\
&= \frac{t}{(1 - 3t + t^2 + t^3)}
\end{aligned}$$

□

Theorem 2.4. The exponential generating function of PLe_n is

$$EGF_{PLe}(t) = \frac{1}{4}(e^{\rho t} + e^{\sigma t}) - \frac{1}{2}$$

where $\rho = 1 + \sqrt{2}$ and $\sigma = 1 - \sqrt{2}$.

Proof. For the exponential generating function of $\{PLe_n\}_{n=0}^{\infty}$, we will deal with the following series representation:

$$EGF_{PLe}(t) = \sum_{n=0}^{\infty} PLe_n \frac{t^n}{n!}$$

Thus, using the Binet formula of PLe_n and $e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!}$, it is obtained

$$EGF_{PLe}(t) = \frac{1}{4}(e^{\rho t} + e^{\sigma t}) - \frac{1}{2}$$

□

Theorem 2.5. For $m \geq n$, the Catalan identity for Pell Leonardo numbers is

$$PLe_{m-n}PLe_{m+n} - {PLe_m}^2 = \frac{(-1)^{m-n+1}}{2}{P_n}^2 + PLe_m - \frac{1}{2}(PLe_{m+n} + PLe_{m-n})$$

where m and n are nonnegative numbers, P_n is nth Pell number, $\rho = 1 + \sqrt{2}$, and $\sigma = 1 - \sqrt{2}$.

Proof. To find Catalan identity for Pell Leonardo numbers, we first calculate the following:

$$
\begin{aligned}
PLe_{m-n}PLe_{m+n} - PLe_m{}^2 &= \left(\frac{\rho^{m-n+1}+\sigma^{m-n+1}}{4}-\frac{1}{2}\right)\left(\frac{\rho^{m+n+1}+\sigma^{m+n+1}}{4}-\frac{1}{2}\right)\\
&-\left(\frac{\rho^{m+1}+\sigma^{m+1}}{4}-\frac{1}{2}\right)\left(\frac{\rho^{m+1}+\sigma^{m+1}}{4}-\frac{1}{2}\right)\\
&=\left(\frac{\rho^{2m+2}+\sigma^{2m+2}}{16}+\frac{(\rho\sigma)^{m-n+1}\,(\sigma^{2n}+\rho^{2n})}{16}\right)\\
&-\left(\frac{\rho^{m-n+1}+\sigma^{m-n+1}}{8}\right)-\left(\frac{\rho^{m+n+1}+\sigma^{m+n+1}}{8}\right)+\frac{1}{4}\\
&-\left(\frac{\rho^{2m+2}+\sigma^{2m+2}+2(\rho\sigma)^{m+1}}{16}\right)+\left(\frac{\rho^{m+1}+\sigma^{m+1}}{8}\right)\\
&+\left(\frac{\rho^{m+1}+\sigma^{m+1}}{8}\right)-\frac{1}{4}
\end{aligned}
$$

By using $\rho\sigma=-1$,

$$
\begin{aligned}
PLe_{m-n}PLe_{m+n} - PLe_m{}^2 &= \frac{(-1)^{m-n+1}}{2}\left(\frac{\sigma^n-\rho^n}{2\sqrt{2}}\right)^2-\frac{\rho^{m-n+1}}{4}\left(\frac{\rho^n-1}{\rho-1}\right)^2-\frac{\sigma^{m-n+1}}{4}\left(\frac{\sigma^n-1}{\sigma-1}\right)^2\\
&=\frac{(-1)^{m-n+1}}{2}P_n{}^2-\frac{\rho^{m-n+1}}{4}\left(\frac{\rho^n-1}{\rho-1}\right)^2-\frac{\sigma^{m-n+1}}{4}\left(\frac{\sigma^n-1}{\sigma-1}\right)^2\\
&=\frac{(-1)^{m-n+1}}{2}P_n{}^2+PLe_m-\frac{1}{2}(PLe_{m+n}+PLe_{m-n})
\end{aligned}
$$

□

If $n=1$ in Catalan identity, Cassini identity is obtained as follows:

Corollary 2.6. For $m\geq 1$,

$$PLe_{m-1}PLe_{m+1} - PLe_m{}^2 = \frac{(-1)^m}{2}-PLe_{m-1}-\frac{1}{2}$$

Theorem 2.7. For $m>n, n\geq 1$, and $m-n>1$, d'Ocagne's identity for Pell Leonardo numbers is

$$PLe_mPLe_{n+1} - PLe_{m+1}PLe_n = \tfrac{(-1)^{n+2}}{4}(PLe_{m-n-2}+PLe_{m-n}+1)+\tfrac{1}{2}(P_{m+1}-P_{n+1})$$

where m and n are nonnegative numbers and P_n is nth Pell number.

Proof. To find d'Ocagne's identity for Pell Leonardo numbers, we first calculate the following:

$$
\begin{aligned}
PLe_mPLe_{n+1} - PLe_{m+1}PLe_n &= \left(\frac{\rho^{m+1}+\sigma^{m+1}}{4}-\frac{1}{2}\right)\left(\frac{\rho^{n+2}+\sigma^{n+2}}{4}-\frac{1}{2}\right)\\
&-\left(\frac{\rho^{m+2}+\sigma^{m+2}}{4}-\frac{1}{2}\right)\left(\frac{\rho^{n+1}+\sigma^{n+1}}{4}-\frac{1}{2}\right)\\
&=\left(\frac{\rho^{m+n+3}+\sigma^{m+n+3}}{16}+\frac{(\rho\sigma)^{n+2}\,(\sigma^{m-n-1}+\rho^{m-n-1})}{16}\right)
\end{aligned}
$$

$$-\left(\frac{\rho^{m+1}+\sigma^{m+1}}{8}\right)-\left(\frac{\rho^{n+2}+\sigma^{n+2}}{8}\right)+\frac{1}{4}$$
$$-\left(\frac{\rho^{m+n+3}+\sigma^{m+n+3}}{16}+\frac{(\rho\sigma)^{n+1}\,(\sigma^{m-n+1}+\rho^{m-n+1})}{16}\right)$$
$$+\left(\frac{\rho^{m+2}+\sigma^{m+2}}{8}\right)+\left(\frac{\rho^{n+1}+\sigma^{n+1}}{8}\right)-\frac{1}{4}$$

By using $\rho\sigma=-1$,

$$PLe_mPLe_{n+1}-PLe_{m+1}PLe_n = \frac{(-1)^{n+2}}{16}\rho^{m-n-1}(1+\rho^2)+\sigma^{m-n-1}(1+\sigma^2) + \frac{\rho^{m+1}\,(\rho-1)-\rho^{n+1}\,(\rho-1)+\sigma^{m+1}\,(\sigma-1)-\sigma^{n+1}\,(\sigma-1)}{8}$$

By using $\rho=1+\sqrt{2}$ and $\sigma=1-\sqrt{2}$,

$$PLe_mPLe_{n+1}-PLe_{m+1}PLe_n = \frac{(-1)^{n+2}}{4}(PLe_{m-n-2}+PLe_{m-n}+1) + \frac{1}{2}\left(\frac{\rho^{m+1}-\sigma^{m+1}}{2\sqrt{2}}\right)-\frac{1}{2}\left(\frac{\rho^{n+1}-\sigma^{n+1}}{2\sqrt{2}}\right)$$
$$= \frac{(-1)^{n+2}}{4}(PLe_{m-n-2}+PLe_{m-n}+1)+\frac{1}{2}(P_{m+1}-P_{n+1})$$

□

Theorem 2.8. Some summation formulas of Pell Leonardo numbers based on Pell numbers are as follows:

i. $\sum_{k=0}^{n}PLe_k=4(\sum_{k=1}^{n-1}P_k)$

ii. $\sum_{k=0}^{n}PLe_{3k}=(\sum_{k=3n-2}^{3n}P_k)+2(\sum_{k=3n-5}^{3n-3}P_k)+\cdots+(n-1)(\sum_{k=4}^{6}P_k)+n(\sum_{k=1}^{3}P_k)$

where $k\geq 0$ and P_k is kth Pell number.

Proof. By using (2.3), it is easily obtained that the above equations are satisfied. □

3. Matrix Formula of Pell Leonardo Numbers

This section introduces a new matrix formulation for Pell Leonardo numbers. By using (2.2), we have

$$\begin{bmatrix}PLe_{n+2}\\PLe_{n+1}\\PLe_n\end{bmatrix}=\begin{bmatrix}3&-1&-1\\1&0&0\\0&1&0\end{bmatrix}\begin{bmatrix}PLe_{n+1}\\PLe_n\\PLe_{n-1}\end{bmatrix}$$

Then, the new matrix formulation for Pell Leonardo numbers is as follows:

$$\begin{bmatrix}PLe_{n+2}\\PLe_{n+1}\\PLe_n\end{bmatrix}=\begin{bmatrix}3&-1&-1\\1&0&0\\0&1&0\end{bmatrix}^n\begin{bmatrix}PLe_2\\PLe_1\\PLe_0\end{bmatrix}$$

Here, we provide a matrix formulation for Pell Leonardo numbers, called the *PL*-matrix, and denote the *PL*-matrix by $PLeM$. If $PLeM$ is

$$\begin{bmatrix}3&-1&-1\\1&0&0\\0&1&0\end{bmatrix}$$

then $\det PLeM=-1$.

Theorem 3.1. The *PL*-matrix for nth Pell Leonardo number is given as

$$PLeM_n = \begin{bmatrix} PLe_{n+1} & -PLe_n - PLe_{n-1} & -PLe_n \\ PLe_n & -PLe_{n-1} - PLe_{n-2} & -PLe_{n-1} \\ PLe_{n-1} & -PLe_{n-2} - PLe_{n-3} & -PLe_{n-2} \end{bmatrix}$$

Proof. To find the *PL*-matrix, we calculate the following:

$$PLeM_n = (PL)^n = \begin{bmatrix} 3 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}^n$$

By using (2.2), $PLeM_n$ is obtained as follows:

$$PLeM_n = \begin{bmatrix} PLe_{n+1} & -PLe_n - PLe_{n-1} & -PLe_n \\ PLe_n & -PLe_{n-1} - PLe_{n-2} & -PLe_{n-1} \\ PLe_{n-1} & -PLe_{n-2} - PLe_{n-3} & -PLe_{n-2} \end{bmatrix}$$

□

Corollary 3.2. For $n \geq 0$, the determinant of the *PL*-matrix for nth Pell Leonardo number can be found as follows:

$$|PLeM_n| = \begin{vmatrix} PLe_{n+1} & -PLe_n - PLe_{n-1} & -PLe_n \\ PLe_n & -PLe_{n-1} - PLe_{n-2} & -PLe_{n-1} \\ PLe_{n-1} & -PLe_{n-2} - PLe_{n-3} & -PLe_{n-2} \end{vmatrix} = (-1)^n \tag{3.1}$$

Proof. To ensure (3.1), we use induction on n. For $n = 0$, $PLeM_0 = I_3$, where I is the identity matrix. Therefore, it is easily seen that (3.1) is held. We suppose that (3.1) is held for $n = u$, that is,

$$|PLeM_u| = \begin{vmatrix} PLe_{u+2} & -PLe_{u+1} - PLe_u & -PLe_{u+1} \\ PLe_{u+1} & -PLe_u - PLe_{u-1} & -PLe_u \\ PLe_u & -PLe_{u-1} - PLe_{u-2} & -PLe_{u-1} \end{vmatrix} = |PL^u| = |PL|^u = (-1)^u$$

Then, by induction, for $n = u + 1$,

$$|PLeM_{u+1}| = \begin{vmatrix} PLe_{u+2} & -PLe_{u+1} - PLe_u & -PLe_{u+1} \\ PLe_{u+1} & -PLe_u - PLe_{u-1} & -PLe_u \\ PLe_u & -PLe_{u-1} - PLe_{u-2} & -PLe_{u-1} \end{vmatrix} = |PL^{u+1}| = |PL|^u|PL| = (-1)^{u+1}$$

Thus, (3.1) holds for all $n \geq 0$. □

4. Conclusion

In this study, we first introduce a new third-order Leonardo sequence and thus add a new third-order number sequence to the literature. We named that sequence Pell Leonardo numbers. Moreover, we obtain some equations for that sequence, including the Binet formula, generating function, exponential generating function, Catalan, Cassini, and d'Ocagne's identities, and some summation formulas based on the Pell sequence. Finally, we describe a *PL*-matrix for Pell Leonardo numbers and obtain the determinant of *PL*-matrix. Different sequences such as complex, bicomplex, gaussian, polynomial, and quaternion sequences can be defined using this sequence. Again, this number sequence and matrix representation can be used in coding theory, cryptography, and other engineering and physics applications.

Author Contributions

The author read and approved the final version of the paper.

Conflict of Interest

The author declares no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] M. Bicknell, *A primer of the Pell sequence and related sequences*, The Fibonacci Quarterly 13 (4) (1975) 345–349.

[2] A. F. Horadam, A*pplications of modified Pell numbers to representations*, Ulam Quarterly 3 (1) (1994) 34-53.

[3] R. Melham, *Sums involving Fibonacci and Pell numbers*, Portugaliae Mathematica 56 (3) (1999) 309-318.

[4] S. F. Santana, J. L. Díaz-Barrero, *Some properties of sums involving Pell numbers*, Missouri Journal of Mathematical Sciences 18 (1) (2006) 33-40.

[5] Q. Mushtaq, U. Hayat, *Pell numbers, Pell–Lucas numbers and modular group*, In Algebra Colloquium 14 (1) (2007) 97-102.

[6] A. Dasdemir, *On the Pell, Pell-Lucas and modified Pell numbers by matrix method*, Applied Mathematical Sciences 5 (64) (2011) 3173-3181.

[7] S. Çelik, İ. Durukan, E. Özkan, *New recurrences on Pell numbers, Pell-Lucas numbers, Jacobsthal numbers* a*nd Jacobsthal-Lucas numbers*, Chaos, Solitons & Fractals 150 (2021), Article Number 111173 8 pages.

[8] P. M. Catarino and A. Borges, *On Leonardo numbers*, Acta Mathematica Universitatis Comenianae 89 (1) (2019) 75-86.

[9] A. G. Shannon, *A note on generalized Leonardo numbers*, Notes on Number Theory and Discrete Mathematics 25 (3) (2019) 97-101.

[10] Y. Alp, E. G. Koçer, *Some properties of Leonardo numbers*, Konuralp Journal of Mathematics 9 (1) (2021) 183-189.

[11] Y. Alp, E. G. Kocer, *Hybrid Leonardo numbers*, Chaos, Solitons & Fractals 150 (2021) Article Number 111128 5 pages.

[12] K. Kuhapatanakul, J. Chobsorn, *On the generalized Leonardo Numbers*, Integers: Electronic Journal of Combinatorial Number Theory 22 (2022) Article Number A48 7 pages.

[13] A. Karatas, *On complex Leonardo numbers*, Notes on Number Theory and Discrete Mathematics 28 (3) (2022) 458-465.

[14] E. Tan, H. H. Leung, *On Leonardo p-numbers*, Integers: Electronic Journal of Combinatorial Number Theory 23 (2023) 1-11.

[15] Y. Soykan, *Generalized Horadam-Leonardo numbers and polynomials*, Asian Journal of Advanced Research and Reports 17 (8) (2023) 128-169.

Journal of New Results in Science
https://dergipark.org.tr/en/pub/jnrs
Research Article
Open Access

E-ISSN: 1304-7981 https://doi.org/10.54187/jnrs.1517251

Compact operators on the Motzkin sequence space $c_0(\mathcal{M})$

Sezer Erdem[1]

Keywords
Motzkin numbers, Sequence spaces, Hausdorff measure of non-compactness, Compact operators

Abstract — The concept of non-compactness measure is extremely beneficial for functional analysis in theories, such as fixed point and operator equations. Apart from these, the Hausdorff measure of non-compactness also has some applications in the theory of sequence spaces which is an interesting topic of functional analysis. One of these applications is to obtain necessary and sufficient conditions for the matrix operators between Banach coordinate (BK) spaces to be compact. In line with these explanations, in this study, the necessary and sufficient conditions for a matrix operator to be compact from the Motzkin sequence space $c_0(\mathcal{M})$ to the sequence space $\mu \in \{\ell_\infty, c, c_0, \ell_1\}$ are presented by using Hausdorff measure of non-compactness.

Subject Classification (2020): 11B83, 46A45

1. Introduction

The linear space containing all sequences of real or complex numbers is symbolized by ω. Each linear subspace Γ of ω is referred to as a sequence space. Some prominent instances of sequence spaces are c (convergent sequences' space), c_0 (null sequences' space), ℓ_∞ (bounded sequences' space) and ℓ_p (p-summable sequences' space). The aforementioned spaces are Banach spaces due to the norms $\|u\|_{\ell_\infty} = \|u\|_c = \|u\|_{c_0} = \sup_{s\in\mathbb{N}} |u_s|$ and $\|u\|_{\ell_p} = (\sum_s |u_s|^p)^{1/p}$ for $1 \leq p < \infty$, where the notion $\sum_s$ purports the summation $\sum_{s=0}^{\infty}$ and $\mathbb{N} = \{0, 1, 2, 3, ...\}$. Moreover, the acronym cs denotes the spaces of all convergent series. A Banach space wherein each coordinate functional f_s, defined by $f_s(u) = u_s$, exhibits continuity and is named a Banach coordinate (BK) space. Given spaces $\Gamma, \Psi \subseteq \omega$, the set $M(\Gamma * \Psi)$ is defined as follows:

$$M(\Gamma * \Psi) = \left\{ \mu = (\mu_r) \in \omega : \mu u = (\mu_r u_r) \in \Psi, \text{ for all } u \in \Gamma \right\}$$

In that case, the beta dual of the set Γ is given by $\Gamma^\beta = M(\Gamma * cs)$. For an infinite matrix $D = (d_{rs})$ possessing entries from the real or complex field, D_r denotes the r^{th} row. The D-transform of $u = (u_s) \in \omega$, denoted by $(Du)_r$, is described as $\sum_{s=0}^{\infty} d_{rs} u_s$ such that the series converges for all $r \in \mathbb{N}$.

[1]sezer.erdem@ozal.edu.tr (Corresponding Author)
[1]Department of Basic Engineering Sciences, Faculty of Engineering and Natural Sciences, Malatya Turgut Özal University, Malatya, Türkiye
Article History: Received: 16 Jul 2024 - Accepted: 19 Aug 2024 - Published: 31 Aug 2024

Consider the sequence spaces Γ and Ψ. A matrix D is called a matrix mapping from Γ to Ψ if, for all $u \in \Gamma$, the image Du belongs to Ψ. The class of all such matrices that effectuate a mapping from Γ to Ψ is denoted by $(\Gamma : \Psi)$. Additionally, the notation Γ_D is employed to represent the set of all sequences for which the D-transform is contained in Γ, as expressed by:

$$\Gamma_D = \{u \in \omega : Du \in \Gamma\} \tag{1.1}$$

It is acknowledged that the matrix domain Γ_D constitutes a sequence space. Furthermore, if Γ is identified as a BK-space and D is triangular, then Γ_D also forms a BK-space, with norm defined by $\|u\|_{\Gamma_D} = \|Du\|_\Gamma$, as elucidated in the literature. In light of this principle, a plethora of intriguing BK-spaces have been the subject of scholarly investigation recently.

Obtaining new normed sequence spaces by using the special matrix and addressing some topics in these spaces, such as completeness, inclusion relations, Schauder basis, duals, matrix transformations, compact operators, and core theorems, have been seen as an important field of study since past years and many valuable researches have been carried out in this subject. When the researchers want to reach basic and advanced concepts on the subjects mentioned above, the sources to be consulted can be referred to as [1–10], and monographs [11–13].

2. Preliminaries

Constructing new sequence spaces as domains of special infinite matrices, as the application of summability theory to sequence spaces, has emerged as an important research area in recent years. With this in mind, Başarır and Kara [14] first obtained an infinite matrix using the Fibonacci number sequence and then constructed new sequence spaces with the help of this matrix and comprehensively examined their various properties. Inspired by the mentioned work, various researchers later obtained sequence spaces with similar ideas using number sequences Lucas, Padovan, Pell, Leanardo, Catalan, Bell, Schröder, Motzkin, and Mersenne. The Schröder matrix $\widetilde{S} = (\widetilde{S}_{rs})$ [15] is defined by

$$\widetilde{S}_{rs} = \begin{cases} \dfrac{S_s S_{r-s}}{S_{r+1} - S_r}, & 0 \le s \le r \\ 0, & s > r \end{cases}$$

Recently, the domains $c_0(\widetilde{S}), c(\widetilde{S}), \ell_p(\widetilde{S})$, and $\ell_\infty(\widetilde{S})$ of $\widetilde{S}$ in c_0, c, ℓ_p, and ℓ_∞, respectively, are studied by Dağlı [15, 16]. Quite recently, the construction of sequence spaces using Catalan and Motzkin numbers has been investigated by Karakaş and Dağlı [17]. They studied $\ell_p(\widetilde{C})$ and $\ell_\infty(\widetilde{C})$ where $\widetilde{C} = (\widetilde{c}_{rs})$ is described by

$$\widetilde{c}_{rs} = \begin{cases} \dbinom{r}{s} \dfrac{M_s}{C_{r+1}}, & 0 \le s \le r \\ 0, & s > r \end{cases}$$

For relevant literature, see [18–28].

2.1. Motzkin Numbers and Associated Sequence Spaces

The first basic information about the Motzkin number sequence, one of the most interesting number sequences, is obtained from Motzkin's study [29]. The r^{th} Motzkin number represents the number of different situations in which n distinct points on a circle can be joined by non-intersecting chords in mathematics. To point out a detail here, the chords do not need to touch all points on the circle in each case. The first few terms of the Motzkin number sequence $(M_r)_{r\in\mathbb{N}}$, which has an important

place in combinatorics, number theory, and geometry, are given as follows:

$$1, 1, 2, 4, 9, 21, 51, 127, 323, 835, 2188, 5798, \cdots$$

The recurrence relations of M_r are given the following way:

$$M_r = M_{r-1} + \sum_{s=0}^{r-2} M_s M_{r-s-2} = \frac{2r+1}{r+1} M_{r-1} + \frac{3r-3}{r+2} M_{r-2}$$

Another relation provided by the Motzkin numbers is given below:

$$M_{r+2} - M_{r+1} = \sum_{s=0}^{r} M_s M_{r-s}, \text{ for } r \geq 0$$

The generating function $m(u) = \sum_{r=0}^{\infty} M_r u^r$ of the Motzkin numbers satisfies

$$u^2 + [m(u)]^2 + (u-1)m(u) + 1 = 0$$

and is described by

$$m(u) = \frac{1 - u - \sqrt{1 - 2u - 3u^2}}{2u^2}$$

The expression on Motzkin numbers with the help of integral function is as follows:

$$M_r = \frac{2}{\pi} \int_0^{\pi} \sin^2 u \, (2cosu + 1)^r \, du$$

They have the asymptotic behavior

$$M_r \sim \frac{1}{2\sqrt{\pi}} \left(\frac{3}{r}\right)^{\frac{3}{2}} 3^r, \quad r \to \infty$$

Furthermore, it is known that

$$\lim_{r \to \infty} \frac{M_r}{M_{r-1}} = 3$$

In addition to the basic information stated above, readers can benefit from the studies Aigner [30], Barrucci et al. [31], and Donaghey and Shapiro [32] for more comprehensive content about Motzkin numbers and related subjects.

The remainder of this subsection will provide information about the study conducted by Erdem et al. [27]. It is given the Motzkin matrix $\mathcal{M} = (\mathfrak{m}_{rs})_{r,s \in \mathbb{N}}$ constructed with the help of Motzkin numbers as follows:

$$\mathfrak{m}_{rs} := \begin{cases} \dfrac{M_s M_{r-s}}{M_{r+2} - M_{r+1}}, & 0 \leq s \leq r \\ 0, & s > r \end{cases} \tag{2.1}$$

Furthermore, the Motzkin matrix $\mathcal{M}$ is conservative, that is $\mathcal{M} \in (c : c)$ and it is given the inverse $\mathcal{M}^{-1} = (\mathfrak{m}_{rs}^{-1})$ of the Motzkin matrix $\mathcal{M}$ as

$$\mathfrak{m}_{rs}^{-1} := \begin{cases} (-1)^{r-s} \dfrac{M_{s+2} - M_{s+1}}{M_r} \pi_{r-s}, & 0 \leq s \leq r \\ 0, & s > r \end{cases} \tag{2.2}$$

where $\pi_0 = 0$ and

$$\pi_r = \begin{vmatrix} M_1 & M_0 & 0 & 0 & \cdots & 0 \\ M_2 & M_1 & M_0 & 0 & \cdots & 0 \\ M_3 & M_2 & M_1 & M_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ M_r & M_{r-1} & M_{r-2} & M_{r-3} & \cdots & M_1 \end{vmatrix}$$

for all $r \in \mathbb{N}\backslash\{0\}$. From its definition, it is clear that $\mathcal{M}$ is a triangle. Furthermore, $\mathcal{M}$-transform of $u = (u_s) \in \omega$ is expressed with

$$\nu_r := (\mathcal{M}u)_r = \frac{1}{M_{r+2} - M_{r+1}} \sum_{s=0}^{r} M_s M_{r-s} u_s, r \in \mathbb{N} \tag{2.3}$$

The Motzkin sequence spaces $c(\mathcal{M})$ and $c_0(\mathcal{M})$, which are BK-spaces constructed as the domain of the Motzkin matrix, are given by

$$c(\mathcal{M}) = \left\{ u = (u_s) \in \omega : \lim_{r\to\infty} \frac{1}{M_{r+2} - M_{r+1}} \sum_{s=0}^{r} M_s M_{r-s} u_s \text{ exists } \right\}$$

and

$$c_0(\mathcal{M}) = \left\{ u = (u_s) \in \omega : \lim_{r\to\infty} \frac{1}{M_{r+2} - M_{r+1}} \sum_{s=0}^{r} M_s M_{r-s} u_s = 0 \right\}$$

and in [27], the authors studied some algebraic and topological properties of newly described spaces.

Consider the sets $\aleph_1$ and $\aleph_2$ as

$$\aleph_1 = \left\{ t = (t_s) \in \omega : \sum_{r=s}^{\infty} (-1)^{r-s} \frac{M_{s+2} - M_{s+1}}{M_r} \pi_{r-s} t_r \text{ exists for each } s \in \mathbb{N} \right\}$$

$$\aleph_2 = \left\{ t = (t_s) \in \omega : \sup_{r\in\mathbb{N}} \sum_{s=0}^{r} \left| \sum_{i=s}^{r} (-1)^{i-s} \frac{M_{s+2} - M_{s+1}}{M_i} \pi_{i-s} t_i \right| < \infty \right\}$$

Then, the β-dual of $c_0(\mathcal{M})$ is determined as $\{c_0(\mathcal{M})\}^\beta = \aleph_1 \cap \aleph_2$.

In this article, the necessary and sufficient conditions are presented for a matrix operator to be compact from the Motzkin sequence space $c_0(\mathcal{M})$ to the sequence space $\mu \in \{\ell_\infty, c, c_0, \ell_1\}$ by using Hausdorff measure of non-compactness.

3. Compact Operators on the Motzkin Sequence Space $c_0(\mathcal{M})$

The current section intends to determine the compactness conditions of an operator from $c_0(\mathcal{M})$ to Ψ by using the Hausdorff measure of non-compactness, where $\Psi \in \{c_0, c, \ell_\infty, \ell_1\}$. We may start by reminding the basic concepts and results in this section.

Let $\mathcal{U}_\Gamma$ represents the unit sphere of normed space Γ. The acronym $\mathfrak{B}(\Gamma : \Psi)$ denotes every bounded (continuous) linear operators' family from the Banach space Γ to Banach space Ψ. In that case, $\mathfrak{B}(\Gamma : \Psi)$ is Banach with $\|\mathcal{L}\| = \sup_{u\in\mathcal{U}_\Gamma} \|\mathcal{L}u\|$. We express the notation

$$\|u\|_\Gamma^\diamond = \sup_{x\in\mathcal{U}_\Gamma} \left| \sum_s u_s x_s \right|$$

for $u = (u_s) \in \omega$, with the assumption of the series is convergent for BK-space $\Gamma \supset \Omega$ and $u \in \Gamma^\beta$, where Ω represents all finite sequences' space. Furthermore, a linear operator $\mathcal{L} : \Gamma \to \Psi$ is called as compact operator if the domain of $\mathcal{L}$ is all of Γ and the sequence $(\mathcal{L}(u))$ possesses a convergent subsequence in Ψ for the spaces Γ and Ψ and every bounded sequence $u = (u_s) \in \Gamma$.

Consider the metric space Γ. Then, the Hausdorff measure of non-compactness of a bounded set $A \subseteq \Gamma$ is symbolized by $\chi(A)$ and stated in the following way:

$$\chi(A) = \inf \left\{ \epsilon > 0 : A \subseteq \cup_{j=0}^{n} A(u_j, m_j), u_j \in \Gamma, m_j < \epsilon, n \in \mathbb{N} \right\}$$

where $A(u_j, m_j)$ represents the open ball centered at u_j with radius m_j for all $j \in \{0, 1, 2, \ldots, n\}$. More descriptive information about the Hausdorff measure of non-compactness can be found in [33].

The Hausdorff measure of non-compactness for $\mathcal{L}$, symbolized as $\|\mathcal{L}\|_\chi$, is characterized as $\|\mathcal{L}\|_\chi = \chi(\mathcal{L}(\mathcal{U}_\Gamma))$. There exists a crucial relationship between the concepts of the Hausdorff measure of non-compactness and compact operators, specifically, "A linear operator $\mathcal{L}$ is compact iff $\|\mathcal{L}\|_\chi = 0$".

For further investigation of sequence spaces and the application of the Hausdorff measure of non-compactness in characterizing compact operators between BK-spaces, readers are encouraged to consult the literature [33–38].

Lemma 3.1. [33] $\ell_\infty^\beta = c^\beta = c_0^\beta = \ell_1$ and $\|u\|_\Gamma^\diamond = \|u\|_{\ell_1}$ for $u \in \ell_1$ and $\Gamma \in \{\ell_\infty, c, c_0\}$.

Lemma 3.2. [36] Consider the BK-spaces Γ and Ψ. Then, $(\Gamma : \Psi) \subseteq \mathfrak{B}(\Gamma : \Psi)$. More clearly, for any $D \in (\Gamma : \Psi)$, there corresponds a linear operator $\mathcal{L}_D \in \mathfrak{B}(\Gamma : \Psi)$ such that $\mathcal{L}_D(u) = Du$, for all $u \in \Gamma$.

Lemma 3.3. [33] Let $\Omega \subseteq \Gamma$ be any BK-space and $D \in (\Gamma : \Psi)$. Then,

$$\|\mathcal{L}_D\| = \|D\|_{(\Gamma:\Psi)} = \sup_{r\in\mathbb{N}} \|D_r\|_\Gamma^\diamond < \infty$$

Theorem 3.4. [33] For $u = (u_m) \in c_0$, consider that $A \subseteq c_0$ is bounded and the operator $\lambda_m : c_0 \to c_0$ is described with $\lambda_m(u) = (u_0, u_1, u_2, u_3, ..., u_m, 0, 0, ...)$. Then, for the identity operator I on c_0, we have

$$\chi(A) = \lim_{m\to\infty} \left(\sup_{u\in A} \|(I - \lambda_m)(u)\| \right)$$

We can give the following results for $x = (x_s)$,$y = (y_s) \in \omega$ connected to each other by the relation

$$y_s = \sum_{j=s}^{\infty} (-1)^{j-s} \frac{M_{s+2} - M_{s+1}}{M_j} \pi_{j-s} x_j$$

for all $s \in \mathbb{N}$.

Lemma 3.5. Let us consider the sequence $x = (x_s) \in (c_0(\mathcal{M}))^\beta$. In that case, $y = (y_s) \in \ell_1$ and

$$\sum_s x_s u_s = \sum_s y_s \nu_s \tag{3.1}$$

for all $u = (u_s) \in c_0(\mathcal{M})$.

Lemma 3.6. For all $x = (x_s) \in (c_0(\mathcal{M}))^\beta$, the following statement is held.

$$\|x\|_{c_0(\mathcal{M})}^\diamond = \sum_s |y_s| < \infty$$

Proof. It is achieved from the Lemma 3.5 that $y = (y_s) \in \ell_1$ and (3.1) holds for $x = (x_s) \in (c_0(\mathcal{M}))^\beta$. Since, $\|u\|_{c_0(\mathcal{M})} = \|\nu\|_{c_0}$, we reach that "$u \in \mathcal{U}_{c_0(\mathcal{M})}$ if and only if $\nu \in \mathcal{U}_{c_0}$". Thus, we can write the equality

$$\|x\|_{c_0(\mathcal{M})}^\diamond = \sup_{u\in\mathcal{U}_{c_0(\mathcal{M})}} \left| \sum_s x_s u_s \right| = \sup_{\nu\in\mathcal{U}_{c_0}} \left| \sum_s y_s \nu_s \right| = \|y\|_{c_0}^\diamond$$

By the aid of the Lemma 3.1, we obtain that

$$\|x\|_{c_0(\mathcal{M})}^\diamond = \|y\|_{c_0}^\diamond = \|y\|_{\ell_1} = \sum_s |y_s| < \infty$$

□

Consider the matrices $H = (h_{rs})$ and $D = (d_{rs})$ as

$$h_{rs} = \sum_{j=s}^{\infty} (-1)^{j-s} \frac{M_{s+2} - M_{s+1}}{M_j} \pi_{j-s} d_{rj}$$

for all $r, s \in \mathbb{N}$ whenever the series is convergent.

Lemma 3.7. For $\Psi \subseteq \omega$ and infinite matrix $D = (d_{rs})$, if $D \in (c_0(\mathcal{M}) : \Psi)$, in that case $H \in (c_0 : \Psi)$ and $Du = H\nu$ for all $u \in c_0(\mathcal{M})$.

Proof. It is obvious by Lemma 3.5. □

Lemma 3.8. If $D \in (c_0(\mathcal{M}) : \Psi)$, then it is achieved that

$$\|\mathcal{L}_D\| = \|D\|_{(c_0(\mathcal{M}):\Psi)} = \sup_{r\in\mathbb{N}} \left(\sum_s |h_{rs}| \right) < \infty$$

where $\Psi \in \{c_0, c, \ell_\infty\}$.

Lemma 3.9. [34] Consider the BK-space $\Gamma \supset \Omega$. Each of the following results are well known:

i. Let $D \in (\Gamma : c_0)$. Then, $\|\mathcal{L}_D\|_\chi = \limsup_r \|D_r\|_\Gamma^\diamond$ and $\mathcal{L}_D$ is compact iff $\lim_r \|D_r\|_\Gamma^\diamond = 0$.

ii. Let Γ possesses AK property or $\Gamma = \ell_\infty$ and $D \in (\Gamma : c)$. In that case,

$$\frac{1}{2} \limsup_r \|D_r - \sigma\|_\Gamma^\diamond \leq \|\mathcal{L}_D\|_\chi \leq \limsup_r \|D_r - \sigma\|_\Gamma^\diamond$$

and $\mathcal{L}_D$ is compact iff $\lim_r \|D_r - \sigma\|_\Gamma^\diamond = 0$ for $\sigma = (\sigma_s)$ and $\sigma_s = \lim_r d_{rs}$.

iii. Let $D \in (\Gamma : \ell_\infty)$. In that case, $0 \leq \|\mathcal{L}_D\|_\chi \leq \limsup_r \|D_r\|_\Gamma^\diamond$ and $\mathcal{L}_D$ is compact if $\lim_r \|D_r\|_\Gamma^\diamond = 0$.

iv. Let $D \in (\Gamma : \ell_1)$. In that case,

$$\lim_j \left(\sup_{E\in\mathcal{F}_j} \left\| \sum_{r\in E} D_r \right\|_\Gamma^\diamond \right) \leq \|\mathcal{L}_D\|_\chi \leq 4. \lim_j \left(\sup_{E\in\mathcal{F}_j} \left\| \sum_{r\in E} D_r \right\|_\Gamma^\diamond \right)$$

and $\mathcal{L}_D$ is compact iff $\lim_j \left(\sup_{E\in\mathcal{F}_j} \|\sum_{r\in E} D_r\|_\Gamma^\diamond \right) = 0$, where $\mathcal{F}$ symbolizes all finite subsets' family of $\mathbb{N}$ and $\mathcal{F}_j$ is the subfamily of $\mathcal{F}$ occuring of subsets of $\mathbb{N}$ with elements that are greater than j.

Theorem 3.10. *i.* For $D \in (c_0(\mathcal{M}) : \ell_\infty)$,

$$0 \leq \|\mathcal{L}_D\|_\chi \leq \limsup_r \sum_s |h_{rs}|$$

holds.

ii. For $D \in (c_0(\mathcal{M}) : c)$,

$$\frac{1}{2} \limsup_r \sum_s |h_{rs} - h_s| \leq \|\mathcal{L}_D\|_\chi \leq \limsup_r \sum_s |h_{rs} - h_s|$$

holds where $h_s = \lim_{r\to\infty} h_{rs}$ for each $s \in \mathbb{N}$.

iii. For $D \in (c_0(\mathcal{M}) : c_0)$,

$$\|\mathcal{L}_D\|_\chi = \limsup_r \sum_s |h_{rs}|$$

holds.

iv. For $D \in (c_0(\mathcal{M}) : \ell_1)$,

$$\lim_j \|D\|_{(c_0(\mathcal{M}):\ell_1)}^{(j)} \leq \|\mathcal{L}_D\|_\chi \leq 4. \lim_j \|D\|_{(c_0(\mathcal{M}):\ell_1)}^{(j)}$$

holds where $\|D\|_{(c_0(\mathcal{M}):\ell_1)}^{(j)} = \sup_{E\in\mathcal{F}_j} \left(\sum_s |\sum_{r\in E} h_{rs}| \right)$ for all $j \in \mathbb{N}$.

Proof. *i.* Let $D \in (c_0(\mathcal{M}) : \ell_\infty)$. From the convergence of $\sum_{s=0}^\infty d_{rs}u_s$ for all $r \in \mathbb{N}$, it is observed

that $D_r \in (c_0(\mathcal{M}))^\beta$. From Lemma 3.6, we reach that

$$\|D_r\|^\diamond_{c_0(\mathcal{M})} = \|H_r\|^\diamond_{c_0} = \|H_r\|_{\ell_1} = \left(\sum_s |h_{rs}|\right)$$

From Lemma 3.9-*iii*, it is obtain

$$0 \leq \|\mathcal{L}_D\|_\chi \leq \limsup_r \sum_s |h_{rs}|$$

ii. If $D \in (c_0(\mathcal{M}) : c)$, in that case $H \in (c_0 : c)$ from Lemma 3.7. Thus, from Lemma 3.9-*ii*, we obtain that

$$\frac{1}{2}\limsup_r \|H_r - h\|^\diamond_{c_0} \leq \|\mathcal{L}_D\|_\chi \leq \limsup_r \|H_r - h\|^\diamond_{c_0}$$

where $h = (h_s)$ and $h_s = \lim_{r\to\infty} h_{rs}$ for all $s \in \mathbb{N}$. Hence, from Lemma 3.1, $\|H_r - h\|^\diamond_{c_0} = \|H_r - h\|_{\ell_1} = \sum_s |h_{rs} - h_s|$, for all $r \in \mathbb{N}$.

iii. Consider that $D \in (c_0(\mathcal{M}) : c_0)$. From the relation $\|D_r\|^\diamond_{c_0(\mathcal{M})} = \|H_r\|^\diamond_{c_0} = \|H_r\|_{\ell_1} = (\sum_s |h_{rs}|)$ for each $r \in \mathbb{N}$ and from Lemma 3.9-*i*, we see $\|\mathcal{L}_D\|_\chi = \limsup_r \sum_s |h_{rs}|$.

iv. Let $D \in (c_0(\mathcal{M}) : \ell_1)$. By Lemma 3.7, we reach that $H \in (c_0 : \ell_1)$. It follows from Lemma 3.9 that

$$\lim_j \left(\sup_{E\in\mathcal{F}_j} \left\|\sum_{r\in E} H_r\right\|^\diamond_{c_0}\right) \leq \|\mathcal{L}_D\|_\chi \leq 4.\lim_j \left(\sup_{E\in\mathcal{F}_j} \left\|\sum_{r\in E} H_r\right\|^\diamond_{c_0}\right)$$

Furthermore, Lemma 3.1 implies that

$$\left\|\sum_{r\in E} H_r\right\|^\diamond_{c_0} = \left\|\sum_{r\in E} H_r\right\|_{\ell_1} = \left(\sum_s \left|\sum_{r\in E} h_{rs}\right|\right)$$

□

Thus, using the theorem given above, we can give the following result.

Corollary 3.11. *i.* For $D \in (c_0(\mathcal{M}) : \ell_\infty)$, $\mathcal{L}_D$ is compact if

$$\lim_r \sum_s |h_{rs}| = 0$$

ii. For $D \in (c_0(\mathcal{M}) : c)$, $\mathcal{L}_D$ is compact iff

$$\lim_r \sum_s |h_{rs} - h_s| = 0$$

iii. For $D \in (c_0(\mathcal{M}) : c_0)$, $\mathcal{L}_D$ is compact iff

$$\lim_r \sum_s |h_{rs}| = 0$$

iv. For $D \in (c_0(\mathcal{M}) : \ell_1)$, $\mathcal{L}_D$ is compact iff

$$\lim_j \|D\|^{(j)}_{(c_0(\mathcal{M}):\ell_1)} = 0$$

where $\|D\|^{(j)}_{(c_0(\mathcal{M}):\ell_1)} = \sup_{E\in\mathcal{F}_j} (\sum_s |\sum_{r\in E} h_{rs}|)$, for all $j \in \mathbb{N}$.

4. Conclusion

Obtaining new normed sequence spaces using special matrices and addressing some intriguing topics such as completeness, inclusion relations, Schauder basis, α-, β- and γ-duals, matrix transformations,

compact operators, core theorems and geometric properties in these spaces has been considered an important field of study in recent years and many valuable researches have been done on this subject. Furthermore, the idea of using special number sequences to obtain sequence spaces has begun to be used by authors. In this context, as an application of matrix summability methods to Banach spaces theory, in this study, it is presented the necessary and sufficient conditions for a matrix operator to be compact from the Motzkin sequence space $c_0(\mathcal{M})$ constructed by the aid of Motzkin number sequence to the sequence space $\mu \in \{\ell_\infty, c, c_0, \ell_1\}$ by using Hausdorff measure of non-compactness. It is noted here that the characterization of compact operators on sequence spaces by using Hausdorff's measure of non-compactness will constitute the focus of our future research endeavors. In future work, researchers can investigate the compactness of operators on different sequence spaces, taking into account those that have not been studied before.

Author Contributions

The author read and approved the final version of the paper.

Conflicts of Interest

The author declares no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] B. Altay, F. Başar, *On some Euler sequence spaces of nonabsolute type*, Ukrainian Mathematical Journal 57 (1) (2005) 1–17.

[2] F. Başar, M. Kirişci, *Almost convergence and generalized difference matrix*, Computers & Mathematics with Applications 61 (3) (2011) 602–611.

[3] T. Yaying, B. Hazarika, M. İlkhan, M.Mursaleen, *Poisson-like matrix operator and its application in $p-$summable space*, Mathematica Slovaca 71 (5) (2021) 1189–1210.

[4] G. C. Hazar, M. A. Sarıgöl, *Absolute Cesàro series spaces and matrix operators*, Acta Applicandae Mathematicae 154 (1) (2018) 153–165.

[5] M. Candan, *Domain of the double sequential band matrix in the classical sequence spaces*, Journal of Inequalities and Applications 2012 (2012) Article Number 281 15 pages.

[6] S. Demiriz, A. Şahin, *q-Cesàro sequence spaces derived by q-analogue*, Advances in Mathematics 5 (2) (2016) 97–110.

[7] S. Erdem, S. Demiriz, *On the new generalized block difference sequence space*, Applications and Applied Mathematics: An International Journal 14 (5) (2019) 68–83.

[8] H. B. Ellidokuzoğlu, S. Demiriz, $[\ell_p]_{e.r}$ *Euler-Riesz difference sequence spaces*, Analysis in Theory and Applications 37 (4) (2021) 557–571.

[9] M. İlkhan, P. Z. Alp, E. E. Kara, *On the spaces of linear operators acting between asymmetric cone normed spaces*, Mediterranean Journal of Mathematics 15 (2018) Article Number 136 12 pages.

[10] M. Stieglitz, H. Tietz, *Matrix transformationen von folgenraumen eine ergebnisbersicht*, Mathematische Zeitschrift 154 (1977) 1–16.

[11] F. Başar, Summability theory and its applications, 2nd Edition, CRC Press/Taylor Frencis Group, Boca Raton-London-New York, 2022.

[12] J. Boos, Classical and modern methods in summability, Oxford Science Publications, Oxford University Press, 2000.

[13] M. Mursaleen, F. Başar, Sequence spaces: Topic in modern summability theory, CRC Press, Taylor Francis Group, Series: Mathematics and its applications, Boca Raton, London, New York, 2020.

[14] E. E. Kara, M. Başarır, *An application of Fibonacci numbers into infinite Toeplitz matrices*, Caspian Journal of Mathematical Sciences 1 (1) (2012) 43–47.

[15] M. C. Dağlı, *A novel conservative matrix arising from Schröder numbers and its properties*, Linear and Multilinear Algebra 71 (8) (2023) 1338–1351.

[16] M. C. Dağlı, *Matrix mappings and compact operators for Schröder sequence spaces*, Turkish Journal of Mathematics 46 (6) (2022) 2304–2320.

[17] M. Karakaş, M. C. Dağlı, *Some topologic and geometric properties of new Catalan sequence spaces*, Advances in Operator Theory 8 (2023) Article Number 14 15 pages.

[18] E. E. Kara, *Some topological and geometrical properties of new Banach sequence spaces*, Journal of Inequalities and Applications 2013 (2013) Article Number 38 15 pages.

[19] M. Karakaş, H. Karabudak, *An application on the Lucas numbers and infinite Toeplitz matrices*, Cumhuriyet Science Journal 38 (3) (2017) 557–562.

[20] M. Karakaş, A. M. Karakaş, *New Banach sequence spaces that is defined by the aid of Lucas numbers*, Iğdır University Journal of the Institute of Science and Technology 7 (4) (2017) 103–111.

[21] T. Yaying, B. Hazarika, S. A. Mohiuddine, *Domain of Padovan $q-$difference matrix in sequence spaces ℓ_p and ℓ_∞*, Filomat 36 (3) (2022) 905–919.

[22] T. Yaying, B. Hazarika, O. M. Kalthum S. K. Mohamed, A. A. Bakery, *On new Banach sequence spaces involving Leonardo numbers and the associated mapping ideal*, Journal of Function Spaces 2022 (2022) Article ID 8269000, 21 Pages.

[23] M. İlkhan, *A new conservative matrix derived by Catalan numbers and its matrix domain in the spaces c and c_0*, Linear and Multilinear Algebra 68 (2) (2020) 417–434.

[24] M. İlkhan, E. E. Kara, *Matrix transformations and compact operators on Catalan sequence spaces*, Journal of Mathematical Analysis and Applications 498 (1) (2021) Article Number 124925 17 pages.

[25] M. Karakaş, M. C. Dağlı, *Some topological and geometrical properties of new Catalan sequence spaces*, Advances in Operator Theory 8 14 (2023) 1–15.

[26] M. Karakaş, *On the sequence spaces involving Bell numbers*, Linear and Multilinear Algebra 71 (14) (2022) 2298–2309.

[27] S. Erdem, S. Demiriz, A. Şahin, *Motzkin sequence spaces and Motzkin core*, Numerical Functional Analysis and Optimization 45 (4) (2024) 1–21.

[28] S. Demiriz, S. Erdem, *Mersenne matrix operator and its application in p-summable sequence space*, Communications in Advanced Mathematical Sciences 7 (1) (2024) 42–55.

[29] T. Motzkin, *Relations between hypersurface cross ratios, and a combinatorial formula for partitions of a polygon, for permanent preponderance, and for non-associative products*, Bulletin of the American Mathematical Society 54 (4) (1948) 352–360.

[30] M. Aigner, *Motzkin numbers*, European Journal of Combinatorics 19 (1998) 663–675.

[31] E. Barrucci, R. Pinzani, R. Sprugnoli, *The Motzkin family*, PUMA Series A 2 (3-4) (1991) 249–279.

[32] R. Donaghey, L.W. Shapiro, *Motzkin numbers*, Journal of Combinatorial Theory, Series A 23 (3) (1977) 291–301.

[33] E. Malkowsky, V. Rakocevic, *An introduction into the theory of sequence spaces and measure of noncompactness*, Zbornik Radova 17 (2000) 143–234.

[34] M. Mursaleen, A. K. Noman, *Compactness by the Hausdorff measure of noncompactness*, Nonlinear Analysis 73 (8) (2010) 2541–2557.

[35] M. Mursaleen, A. K. Noman, *Applications of the Hausdorff measure of noncompactness in some sequence spaces of weighted means*, Computers & Mathematics with Applications 60 (5) (2010) 1245–1258.

[36] A. Wilansky, Summability through functional analysis, Amsterdam-New York-Oxford: North-Holland Mathematics Studies, 1984.

[37] F. Başar, E. Malkowsky, *The characterization of compact operators on spaces of strongly summable and bounded sequences*, Applied Mathematics and Computation 217 (2011) 5199–5207.

[38] M. Başarır, E. E. Kara, *On the B-difference sequence space derived by generalized weighted mean and compact operators*, Journal of Mathematical Analysis and Applications 391 (2012) 67–81.

Journal of New Results in Science
https://dergipark.org.tr/en/pub/jnrs
Research Article
Open Access

E-ISSN: 1304-7981 https://doi.org/10.54187/jnrs.1517322

Barrow holographic dark energy models in Lyra and general relativity theories

Arzu Aktaş[1], **Sezgin Aygün**[2]

Keywords
Barrow holographic dark energy, Lyra theory, Hubble parameter, General relativity

Abstract — This study investigates the Barrow holographic dark energy (BHDE) matter distribution in the Bianchi I universe model in Lyra and General Relativity Theories. To this end, it obtains exact solutions by Hubble parameter, conservation equation, and BHDE energy density equation and supports them with graphics. The results show that the solutions are in harmony with the functioning of the universe and the nature of dark energy. It finally discusses the need for further research.

Subject Classification (2020): 83C05, 83C15

1. Introduction

The accelerating expansion of the Universe is the expansion of the Universe observed from the fact that a galaxy at a certain distance is constantly moving away from the observer at an increasing rate over time. The idea of accelerating the universe's expansion was first proposed independently in 1998 by two research teams, the Supernova Cosmology Project and the High-Z Supernova Search Team [1, 2]. By studying a type of supernova called Type Ia supernovae, these teams discovered that supernovae in distant galaxies appear fainter than those in nearby galaxies. This meant that the light from distant galaxies took longer to reach us. Therefore, the galaxies were moving away faster.

The universe's accelerating expansion is one of the most important discoveries of modern cosmology and has radically changed our understanding of how the universe works. This discovery also revealed the existence of a mysterious form of energy called dark energy. Dark energy is estimated to account for about %70 of the universe's total energy and is responsible for accelerating the universe's expansion [3]. Holographic dark energy models have also been frequently used in recent years to explain the universe's acceleration. Holographic dark energy is one of the theoretical models proposed to explain the universe's accelerating expansion and is related to the holographic principle. The holographic principle suggests that all system information can be expressed in a two-dimensional surface carried by its boundaries [4]. Based on this principle, the holographic dark energy model proposed by Li et

[1]aktas_arzu@hotmail.com (Corresponding Author); [2]saygun@comu.edu.tr
[1]Department of Computer Engineering, Engineering Faculty, İstanbul Galata University, İstanbul, Türkiye
[2]Department of Physics, Faculty of Science, Çanakkale Onsekiz Mart University, Çanakkale, Türkiye
Article History: Received: 16 Jul 2024 - Accepted: 19 Aug 2024 - Published: 31 Aug 2024

al. [5] assumes that the dark energy density depends on a length scale determined by the universe's horizon.

Dark energy models and alternative theories of gravitation are frequently used to explain the universe's evolution. Various gravitation theories such as Lyra, Brans-Dicke, f(R), and f(R,T) are counted among alternative gravitation theories. In this study, we have considered the Barrow holographic dark energy (BHDE) model for the Bianchi I metric in the Lyra alternative gravitation theory. Similar studies are summarised below.

Chanda et al. [6] have analyzed the Barrow holographic dark energy matter distribution in the framework of the Brane world theory. Pankaj and Sharma [7] have studied the cosmological evolution of the Barrow holographic dark energy in a flat Friedmann-Lemaitre-Robertson-Walker (FLRW) universe using the Granda-Oliveros cutoff and a time-dependent scale factor. Devi and Kumar [8] have investigated the behavior of the Barrow holographic dark energy model in Brans-Dicke cosmology for the flat FRLW universe. Sobhanbabu et al. have researched the BHDE for the Kantowski-Sachs universe in the framework of the Saez-Ballester theory. In order to find the exact solution of the equations, they [9] assumed that the shear scalar is proportional to the expansion scalar. The cosmological properties of the Generalized BHDE were studied in the Finsler-Randers universe using a dark energy-density parameter, equation of state parameter, and deceleration parameter by Feng et al. [10]. The cosmological evolution of the BHDE model for a non-flat universe has been analyzed using the FRW metric with open and closed geometries. They studied. [11]. Sharma et al. [12] investigated the Barrow holographic dark energy universe model for the flat FRW metric in $f(R,T)$ theory. Saha et al. [13] Barrow presented an approach to explore the holographic dark energy inflationary cosmology in the FRW universe. An interactive scenario of Chaplygin gas and modified generalized Chaplygin gas were analyzed by Sultana et al. [14]. Shukla et al. [15] studied the flat FLRW cosmological model in $f(Q,T)$ theory. Moreover, Shukla et al. considered the field equations as an anisotropy parameter that can indicate the anisotropic behavior of the universe in the past and its approach to isotropy in the present and studied the Bianchi I universe model within the framework of $f(R,T)$ theory [16]. Devi et al. [17] have investigated the Barrow holographic dark energy universe model for FRW metric in $f(R,T)$ theory. Sing and Meitei [18] have studied dark energy models in general relativity theory for the Bianchi V universe using the linear deceleration parameter. Solutions for the FRW universe in fractal cosmology using special Hubble parameters have been made by Pawar et al. [19]. Ditta et al. [20] have compared the dark-energy compact stars in $f(T)$ and $f(T,\mathcal{T})$ alternative gravitational theories.

In the present paper, section 2 derives the equations of the Lyra gravitation theory for the Barrow holographic dark energy model and obtains the solutions with graphical support. Section 3 provides the results and a discussion. Section 4 discusses the need for further research.

2. Lyra Gravitation Theory and Solutions

The Einstein field equations in the Lyra theory, one of the alternative theories of gravitation, are as follows:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \frac{3}{2}\left(\Phi_\mu\Phi_\nu - \frac{1}{2}\,\mathrm{g}_{\mu\nu}\Phi_\xi\Phi^\xi\right) = -T_{\mu\nu} \tag{2.1}$$

Here, Φ_μ is the displacement vector. This study takes the displacement vector as $\Phi_\mu = (0,0,0,0,\beta(t))$. The homogeneous anisotropic Bianchi I metric in (x,y,z,t) coordinates is defined as follows.

$$ds^2 = A^2dx^2 + B^2dy^2 + C^2dz^2 - dt^2 \tag{2.2}$$

here A, B, and C are functions of time.

In this study, the universe is assumed to be filled with matter and a hypothetical fluid known as holographic dark energy. ρ_B is the energy density of the Barrow holographic dark energy, ρ_m is the energy density of matter, P is the pressure, and the total energy-momentum tensor is the sum of the energy-momentum tensors of matter, and dark energy

$$T_{\mu\nu} = (\rho_m + \rho_B + P)u_\mu u_\nu + P g_{\mu\nu} \tag{2.3}$$

In the case of holographic dark energy, dark matter and dark energy are considered as a dual fluid [21]. The most common formulation of holographic dark energy takes the Hubble horizon as an IR cutcutoffd the Bekenstein-Hawking field law for the horizon degree of freedom of the universe [21]. Holographic dark energy based on Barrow entropy has attracted a great deal of attention in recent years [9]. The energy density of Barrow holographic dark energy using Barrow entropy :

$$\rho_B = LH^{2-\xi} \tag{2.4}$$

here L and ξ are constants [9, 21].

From (2.1)-(2.3), we obtain Einstein field equations in Lyra theory is

$$\frac{\ddot{C}}{C} + \frac{\ddot{B}}{B} + \frac{\dot{C}\dot{B}}{BC} + \frac{3\beta^2}{4} = -P \tag{2.5}$$

$$\frac{\ddot{A}}{A} + \frac{\ddot{B}}{B} + \frac{\dot{A}\dot{B}}{AB} + \frac{3\beta^2}{4} = -P \tag{2.6}$$

$$\frac{\ddot{C}}{C} + \frac{\ddot{A}}{A} + \frac{\dot{A}\dot{C}}{AC} + \frac{3\beta^2}{4} = -P \tag{2.7}$$

$$\frac{\dot{B}\dot{C}}{BC} + \frac{\dot{A}\dot{C}}{AC} + \frac{\dot{A}\dot{B}}{AB} - \frac{3\beta^2}{4} = \rho_m + \rho_B \tag{2.8}$$

The above system of equations contains 7 unknowns. In order to solve this system of equations, we need 3 more equations.

i. As is known, the Hubble parameter is an important time-varying parameter that plays a critical role in calculating the age of the universe and analyzing the expansion of the universe. For this reason, it can be used to solve field equations. The Hubble parameter is taken in the following form.

$$H = \frac{1}{3}\left(\frac{\dot{A}}{A} + \frac{\dot{B}}{B} + \frac{\dot{C}}{C}\right) = \frac{1}{\sqrt{2\alpha t + \gamma}} \tag{2.9}$$

here α and γ are constants.

ii. Secondly, in order to describe the nature of Barrow holographic dark energy, we can use (2.4), which gives the relation between the energy density of the BHDE and the Hubble parameter.

iii. One of the methods frequently used in Lyra alternative gravitation theory solutions is including the conservation equation. In Lyra gravitation theory, the conservation equation is provided as follows:

$$\left(R^\nu_\mu - \frac{1}{2}Rg^\nu_\mu\right)_{;\nu} + \frac{3}{2}\left(\Phi_\mu\Phi^\nu\right)_{;\nu} - \frac{3}{4}\left(\Phi_\eta\Phi^\eta g^\nu_\mu\right)_{;\nu} = 0 \tag{2.10}$$

From (2.2) and (2.10), the conservation equation of Lyra gravitation theory is obtained in terms of metric potentials as follows.

$$\frac{3\beta^2}{2}\left(\frac{\dot{A}}{A} + \frac{\dot{B}}{B} + \frac{\dot{C}}{C} + \frac{\dot{\beta}}{\beta}\right) = 0 \tag{2.11}$$

As shown from (2.11), two solutions exist for $\beta = 0$ and $\beta \neq 0$. If we use the solution $\beta = 0$, Lyra's theory of gravity reduces to Einstein's relativity theory. Since solutions are sought in Lyra gravity theory, we will first investigate solutions for the $\beta \neq 0$ state.

Using (2.5)-(2.9) and (2.11), the energy density (ρ_B) of the BHDE for the Bianchi I metric in the Lyra theory of gravitation is obtained as follows:

$$\rho_B = L\,(2\alpha t+\gamma)^{-2+\xi} \tag{2.12}$$

The density of matter (ρ_m) and pressure (P) are also obtained.

$$\rho_m = \frac{3}{2\alpha t+\gamma} - \left(c_2^2 - c_2c_4 + c_4^2 + \frac{3}{4}c_6^2\right) e^{-\frac{6\sqrt{2\alpha t+\gamma}}{\alpha}} - L\,(2\alpha t+\gamma)^{-2+\xi} \tag{2.13}$$

and

$$P = \frac{2\alpha}{(2\alpha t+\gamma)^{\frac{3}{2}}} - \left(c_2^2 - c_2c_4 + c_4^2 + \frac{3}{4}c_6^2\right) e^{-\frac{6\sqrt{2\alpha t+\gamma}}{\alpha}} - \frac{3}{2\alpha t+\gamma} \tag{2.14}$$

The metric potentials A, B, and C are obtained as follows.

$$A = \frac{c_1}{c_3c_5e^{\frac{2\left(\alpha t+\frac{\alpha\sqrt{2\alpha t+\gamma}}{6}+\frac{\gamma}{2}\right)\left(-3+\alpha(c_2-c_4)e^{-\frac{3\sqrt{2\alpha t+\gamma}}{\alpha}}\right)}{3\sqrt{2\alpha t+\gamma}\,\alpha}}} \tag{2.15}$$

$$B = c_5e^{\frac{-c_4\alpha\left(\alpha+3\sqrt{2\alpha t+\gamma}\right)e^{-\frac{3\sqrt{2\alpha t+\gamma}}{\alpha}}+9\sqrt{2\alpha t+\gamma}}{9\alpha}} \tag{2.16}$$

and

$$C = c_3e^{\frac{2\left(c_2\alpha\left(\alpha t+\frac{\alpha\sqrt{2\alpha t+\gamma}}{6}+\frac{\gamma}{2}\right)e^{-\frac{3\sqrt{2\alpha t+\gamma}}{\alpha}}+3\alpha t-\frac{\alpha\sqrt{2\alpha t+\gamma}}{2}+\frac{3\gamma}{2}\right)}{3\sqrt{2\alpha t+\gamma}\,\alpha}} \tag{2.17}$$

Finally, the displacement vector β^2 is as follows.

$$\beta^2 = \frac{c_6^2}{e^{\frac{6\sqrt{2\alpha t+\gamma}}{\alpha}}} \tag{2.18}$$

Here, c_1-c_6 are constants. From (2.15)-(2.17), the time variation of the metric potentials A, B, and C is shown in Figure 1.

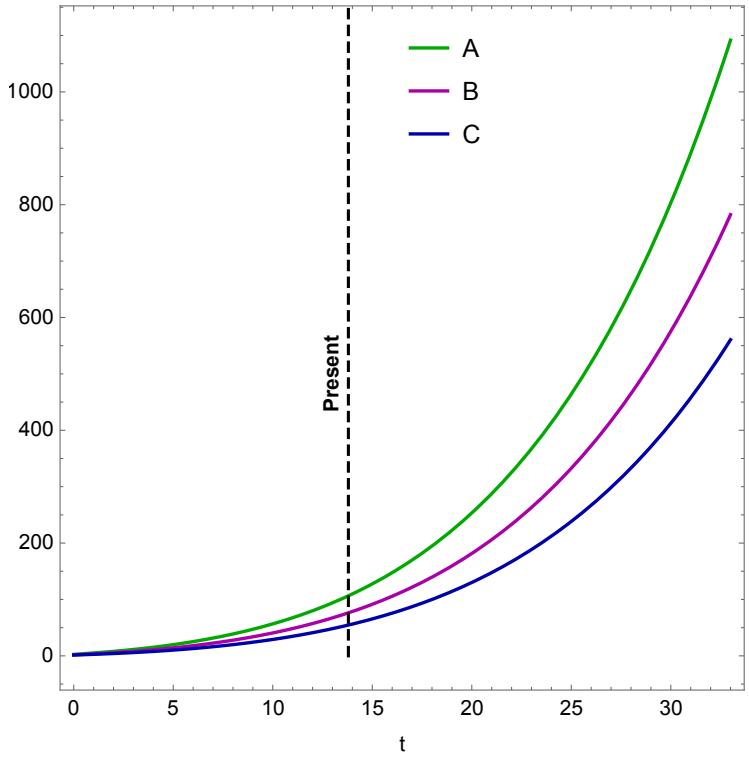


Figure 1. Variation of metric potential versus cosmic time t

The change of BHDE with time is shown in Figure 2.

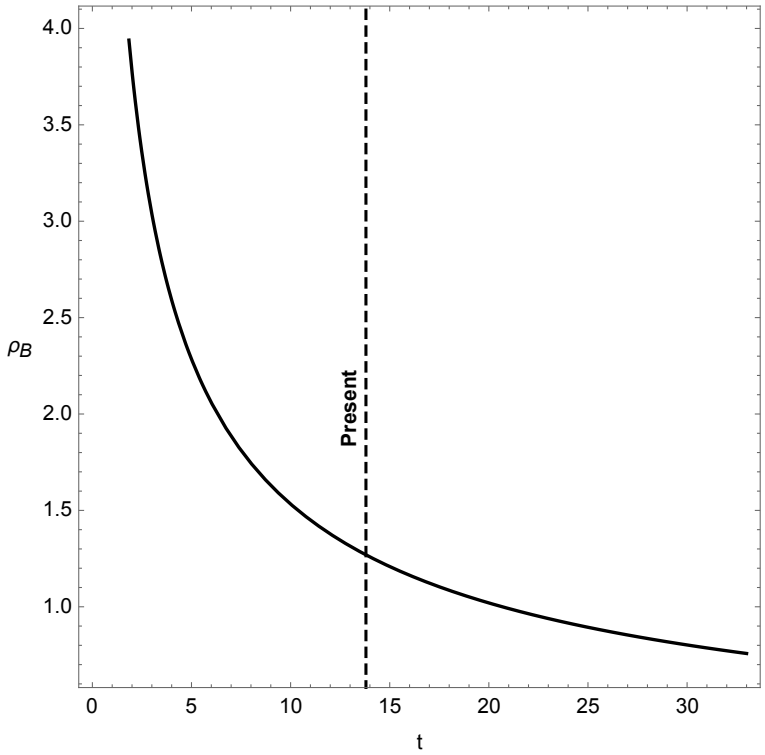


Figure 2. Variation of BHDE density versus cosmic time t

The graph of the change of (2.14) concerning time is shown in Figure 3.

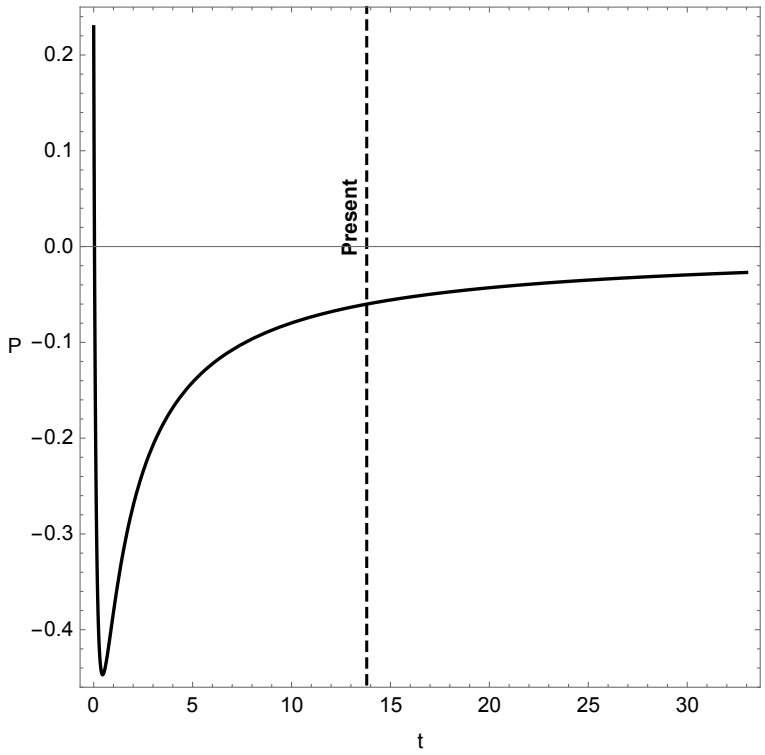


Figure 3. Variation of pressure versus cosmic time t

Figure 4 shows the change of the displacement vector concerning time.

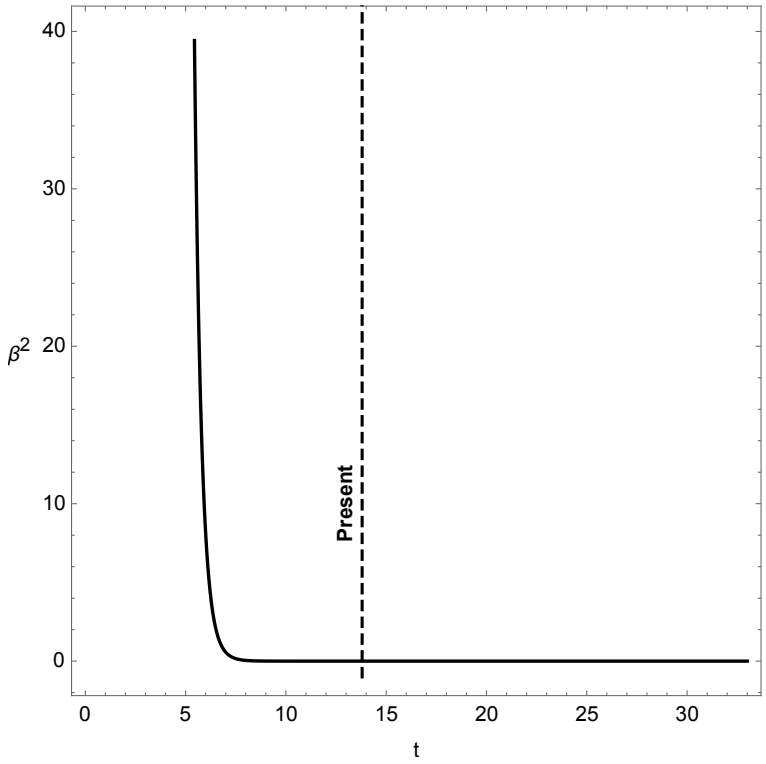


Figure 4. Variation of displacement vector versus cosmic time t

3. Results and Discussions

The Bianchi I universe model, known for its homogeneous and anisotropic properties, was studied in Lyra gravitation theory, accompanied by Barrow holographic dark energy. Hubble parameter, position equation, and density equation of BHDE were used to obtain exact solutions. As can be seen from (2.18), the displacement vector β^2 is obtained depending on time and decreases over time (see figure 4). The fact that the β^2 function is non-zero indicates the existence of Lyra gravitation theory. If $c_6 = 0$ is taken in (2.18), $\beta^2 = 0$. As a result, our solutions are reduced to the General Relativity Theory in the BHDE matter distribution for the Bianchi I metric. From (2.15)-(2.17), we observe that the metric potentials of A, B and C increase over time (see figure 1). Again, the same equations show that the constants c_1, c_3, and c_5 must differ from zero. We can say that the increase in metric potentials over time may indicate the universe's expansion. Also, the present value in the graphs shows $t = 13.8$ Gyr, which is the universe's current age.

As is known, the BHDE model was discussed in this study. In dark energy models, pressure takes negative values. In our research, when we look at figure 3, it is seen that the pressure behaves by the dark energy model. As can be seen from (2.14), there is singularity for $t = -\frac{\gamma}{2\alpha}$ in our model. As seen from (2.12) and (2.13), Barrow's dark energy density and matter density decrease over time. This situation can be seen in figure 2.

As it is known, $p = w\rho_B$ is called the equation of state (EoS) that gives the relationship between pressure and density. In dark energy models, w is a parameter that takes negative values. Using (2.12) and (2.14), the w value and graph for BHDE are as follows:

$$w = \frac{\frac{2\alpha}{(2\alpha t+\gamma)^{\frac{3}{2}}} - \left(c_2^2 - c_2c_4 + c_4^2 + \frac{3}{4}c_6^2\right) e^{-\frac{6\sqrt{2\alpha t+\gamma}}{\alpha}} - \frac{3}{2\alpha t+\gamma}}{L\left(2\alpha t+\gamma\right)^{-2+\xi}} \tag{3.1}$$

Figure 5 shows the change of the equation of state concerning time.

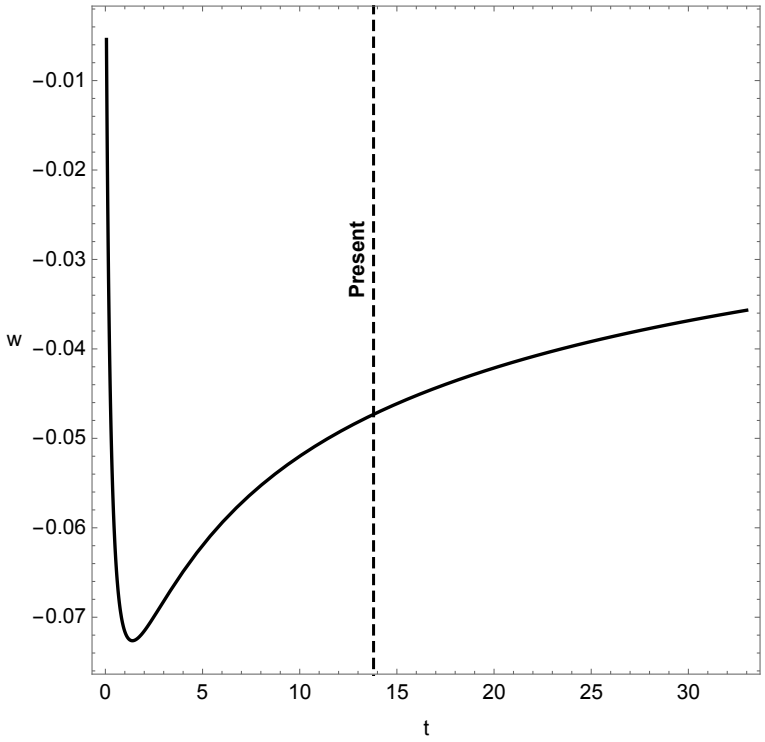


Figure 5. Variation of equation of state versus cosmic time t

As seen in Figure 5, the w value was obtained as negative in our solutions by the nature of dark energy.

As mentioned above, for the case $c_6 = 0$, $\beta^2 = 0$. In this case, the dynamics that are different in the Bianchi I universe model for BHDE in the General Relativity Theory are as follows.

$$\rho_m = \frac{3}{2\alpha t+\gamma} - \left(c_2^2 - c_2c_4 + c_4^2\right) e^{-\frac{6\sqrt{2\alpha t+\gamma}}{\alpha}} - L\left(2\alpha t+\gamma\right)^{-2+\xi} \tag{3.2}$$

$$P = \frac{2\alpha}{(2\alpha t+\gamma)^{\frac{3}{2}}} - \left(c_2^2 - c_2c_4 + c_4^2\right) e^{-\frac{6\sqrt{2\alpha t+\gamma}}{\alpha}} - \frac{3}{2\alpha t+\gamma} \tag{3.3}$$

Other dynamics are as in the Lyra theory. No change occurs.

4. Conclusion

This study investigates the homogeneous and anisotropic Bianchi I universe model in Lyra gravitation theory with Barrow holographic dark energy matter distribution. To obtain the exact solutions of the field equations in Lyra gravitation theory, Hubble parameter, conservation equation, and variation of the energy density of the BHDE were used. In future studies, it will be worthwhile to investigate universe models with BHDE matter distribution within the framework of other alternative gravity theories such as $f(Q)$, $f(R,T)$, $f(Q,T)$, and $f(R)$.

Author Contributions

All the authors equally contributed to this work. They all read and approved the final version of the paper.

Conflicts of Interest

All the authors declare no conflict of interest

Ethical Review and Approval

No approval from the Board of Ethics is required.

Acknowledgement

This work was supported by Çanakkale Onsekiz Mart University The Scientific Research Coordination Unit, Project number FHD-2024-4612.

References

[1] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, J. Tonry, *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, The Astronomical Journal 116 (3) (1998) 1009.

[2] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, T. S. C. Project, *Measurements of* Ω *and* Λ *from 42 high-redshift supernovae*, The Astrophysical Journal 517 (2) (1999) 565–586.

[3] P. J. E. Peebles, B. Ratra, *The cosmological constant and dark energy*, Reviews of Modern Physics 75 (2003) 559–606.

[4] L. Susskind, *The world as a hologram*, Journal of Mathematical Physics 36 (11) (1995) 6377–6396.

[5] M. Li, *A model of holographic dark energy*, Physics Letters B 603 (1-2) (2004) 1–5.

[6] A. Chanda, A. K. Mitra, S. Dey, S. Ghose, B. C. Paul, *Barrow holographic dark energy in Brane world cosmology*, Classical and Quantum Gravity 41 (3) (2024) 035004.

[7] Pankaj, A. Remya, U. K. Sharma, *Cosmological model in Barrow holographic dark energy with Granda-Oliveros cut-off*, International Journal of Modern Physics A 38 (2023) 2350185–6.

[8] K. Devi, P. Kumar, *Recent acceleration and future deceleration in Brans-Dicke theory*, General Relativity and Gravitation 56 (1) (2023) 4.

[9] Y. Sobhanbabu, R. S. Rao, Y. J. Prasuna, G. Satyanarayana, *Kantowski-Sachs Barrow holographic dark energy model in Saez-Ballester theory of gravitation*, New Astronomy 104 (2023) 102066.

[10] W. Feng, W. Yang, B. Jiang, Y. Wang, T. Han, Y. Wu, *Theoretical analysis on the Barrow holographic dark energy in the Finsler-Randers cosmology*, International Journal of Modern Physics D 32 (5) (2023) 2350029.

[11] A. Remya, U. K. Pankaj, Sharma, N. M. Ali, *Cosmological evolution and stability analysis in non-flat universe and Barrow holographic model of dark energy*, Astrophysics and Space Science 368 (3) (2023) 15.

[12] U. K. Sharma, M. Kumar, G. Varshney, *Scalar field models of Barrow holographic dark energy in f(R,T) Gravity*, Universe 8 (12) (2022) 642.

[13] S. Saha, S. Chattopadhyay, E. Güdekli, *A coupled-fluid approach to explore bounce and inflationary cosmology with Barrow holographic as the driving dark fluid*, European Physical Journal C 84 (3) (2024) 314.

[14] S. Sultana, E. Güdekli, S. Chattopadhyay, *Some versions of Chaplygin gas model in modified gravity framework and validity of generalized second law of thermodynamics*, Zeitschrift Naturforschung Teil A 79 (1) (2024) 51–70.

[15] B. K. Shukla, D. Sofuoğlu, H. Chaudhary, F. Atamurotov, G. Mustafa, *Cosmic evolution in f(Q,T) gravity with observational constraints: A comparative analysis with* ΛCDM, Journal of High Energy Astrophysics 43 (2024) 1–14.

[16] B. K. Shukla, R. K. Tiwari, A. Beesham, D. Sofuoğlu, *Cosmology with the late time isotropy of the Bianchi type I model in f(R,T) theory*, International Journal of Geometric Methods in Modern Physics 21 (3) (2024) 2450060–134.

[17] K. Devi, A. Kumar, P. Kumar, *Barrow holographic dark energy model in f(R,T) theory*, Astrophysics and Space Science 369 (7) (2024) 73.

[18] K. P. Singh, A. J. Meitei, *Late time acceleration in Bianchi type-V dark energy cosmological models with linear deceleration parameter*, Physica Scripta 99 (7) (2024) 075046.

[19] D. D. Pawar, D. K. Raut, A. P. Kale, *Two forms of dark energy in fractal cosmological model using specific Hubble parameter*, New Astronomy 110 (2024) 102214.

[20] A. Ditta, G. Mustafa, S. K. Maurya, D. Sofuoğlu, A. Mahmood, *Comparative analysis of dark energy compact stars in* $f(T,\mathcal{T})$ *and* $f(T)$ *gravity theories via conformally flat condition*, Classical and Quantum Gravity 41 (15) (2024) 155004.

[21] M. Srivastava, M. Kumar, S. Srivastava, *Barrow holographic dark energy with hybrid expansion law*, Gravitation and Cosmology 28 (1) (2022) 70–80.

Journal of New Results in Science

https://dergipark.org.tr/en/pub/jnrs

Research Article

Open Access

E-ISSN:1304-7981 https://doi.org/10.54187/jnrs.1520708

The invasive Nearctic wasp *Isodontia mexicana* (Hymenoptera, Sphecidae) now established in Türkiye

İlyas Can[1]

Keywords:
Türkiye,
Hymenoptera,
Invasive,
Isodontia mexicana,
New record

Abstract — The genus *Isodontia* Patton, 1880, consisting of grass-carrying wasps belonging to the family Sphecidae, has 62 described species distributed worldwide. Two are native to the Western Palaearctic and occur in Türkiye: *Isodontia paludosa* (Rossi, 1790) and *Isodontia splendidula* (A. Costa, 1858). *Isodontia mexicana* (de Saussure, 1867) and *Isodontia nigella* (F. Smith, 1856) are non-native species recorded in the Western Palaearctic. Former is the North American species of the genus, accidentally introduced into Europe and spread to many countries. This study reports for the first-time presence of *Isodontia mexicana* in Türkiye. The present record from near the western edge of Asia Minor constitutes the second locality where the species was detected in the Asian continent after Iran. Differential morphologic characters for the species are provided and illustrated. This new species record brings the number of Sphecidae fauna of Türkiye to 80.

Subject Classification (2020):

1. Introduction

The genus *Isodontia* Patton, 1880 contains 62 described species distributed worldwide [1]. A characteristic feature of most species belonging to this genus is that females use pieces of grass stem to make partitions between their brood cells and plug the entrance holes with dry grass leaves; hence the name “grass-carrying wasps” [2]. Unlike other fossorial species of the subfamily Sphecinae, *Isodontia* spp. builds its nests in various cavities such as hollow plant branches and stems, abandoned bee holes, crevices in building walls, and metal shutters around houses. They choose numerous species of Orthoptera as their prey, especially from the families Gryllidae and Tettigoniidae [3,4]. In Western Palaearctic fauna, this genus is represented by two native species, *Isodontia paludosa* (Rossi, 1790) and *Isodontia splendidula* (A. Costa, 1858), and two invasive species, *Isodontia mexicana* (de Saussure, 1867) and *Isodontia nigella* (F. Smith, 1856), both native species present in Türkiye [5].

Isodontia mexicana, commonly known as the Nearctic grass-carrying wasp, is native to Central and North America. It was accidentally introduced to Europe and some Pacific islands around the middle of the 20th century [4]. After being first detected in France, it has rapidly spread to many European countries in Western, Southern, and Eastern Europe, Great Britain, and Iran [1]. The species was recorded in 2017 and 2018 in Bulgaria (Sofia) [6], Greece (Peloponnes) [7], and Romania (Bucharest) [8], which are closest to the Turkish borders. Considering the direction of its expansion within the continent, the species was expected to reach Türkiye eventually.

[1]ilyascan41@gmail.com (Corresponding Author)
[1]Department of Biology, Faculty of Arts and Sciences, Tokat Gaziosmanpaşa University, Tokat, Türkiye
Article History: Received: 22 Jul 2024 — Accepted: 27 Aug 2024 — Published: 31 Aug 2024

Thus, the current contribution aims to report the first record of *Isodontia mexicana* in Türkiye and present its important diagnostic characteristics. With this study, a species was added to the Sphecidae fauna of Türkiye, and the number reached 80.

2. Materials and Methods

Detecting a female specimen of the species was done incidentally rather than as a result of a planned study or regular monitoring. The author saw the female specimen carrying grass to the hole in the metal railing in front of the settlement (Figure 1). After the detected specimen entered the hole, the mouth of the hole was covered with a bottle, and the specimen was captured. The specimen has been identified based on the keys and diagnoses provided by Bitsch et al. [9], Pagliano and Negrisolo [10], and Notton [11]. The material studied is deposited in the Entomology Research Laboratory, Department of Biology, Tokat Gaziosmanpaşa University (Tokat, Türkiye). The photographs of the specimens were taken using a Leica M205C stereomicroscope controlled by the Leica Application Suite 3 software.

Figure 1. Nesting area and nest entrance (indicated with red arrow) of *Isodontia mexicana* specimen in Türkiye.

3. Results and Discussion

Family Sphecidae Latreille, 1802

Genus *Isodontia* Patton, 1880

***Isodontia mexicana* (de Saussure, 1867) (Figure 2)**

Material examined: One female collected from its nest entrance, 65 m above sea level, 40°43'04.6"N, 29°59'12.4"E, Başiskele district, Kocaeli province, 17 June 2024, collected and determined by İlyas CAN.

Figure 2. Lateral habitus of *Isodontia mexicana* (♀) (Scale bar: 2 mm).

Global Distribution. Europe, Iran, USA, Türkiye (new record in this study) [1,8] (Figure 3)

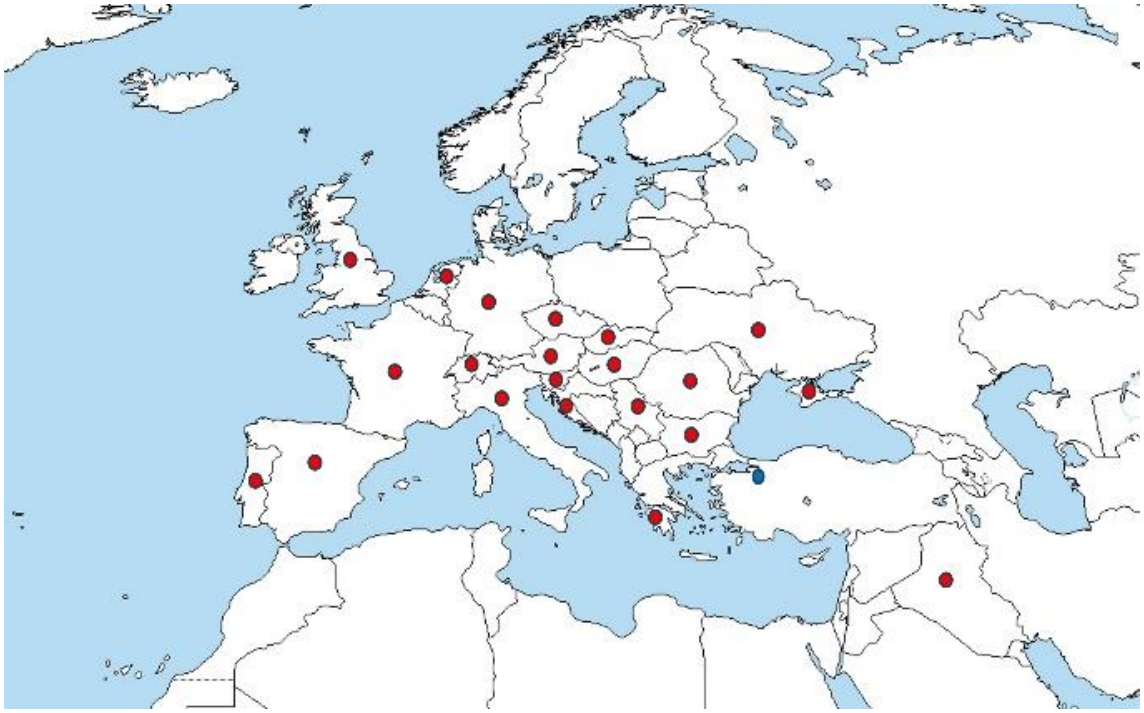


Figure 3. Distribution map of *Isodontia mexicana* in the Palaearctic region (red circle: previous records; blue circle: new record)

Differential diagnosis of female: Body length 16.8 mm. *Isodontia mexicana* can be easily distinguished from other native genus representatives as follows: it differs from *Isodontia splendidula* by the black metasoma (partly reddish in *Isodontia splendidula*). It differs from the closely related native species *Isodontia paludosa*

by the following characteristics: a smaller body size (larger in *Isodontia paludosa*); petiole strongly curved in profile, as long or longer than foretibia (petiole straight, clearly shorter than the foretibia in *Isodontia palludosa*) (Figures 4a and 4e); clypeus and lower frons with long erect black hairs (without black hairs in *Isodontia paludosa*) (Figures 4b and 4f); anterior clypeal margin of female with a small deep median notch (straight in *Isodontia paludosa*) (Figures 4c and 4g); wings smoky, darker along apical margin (wings tinged yellow in *Isodontia paludosa*) (Figures 4d and 4h) [11].

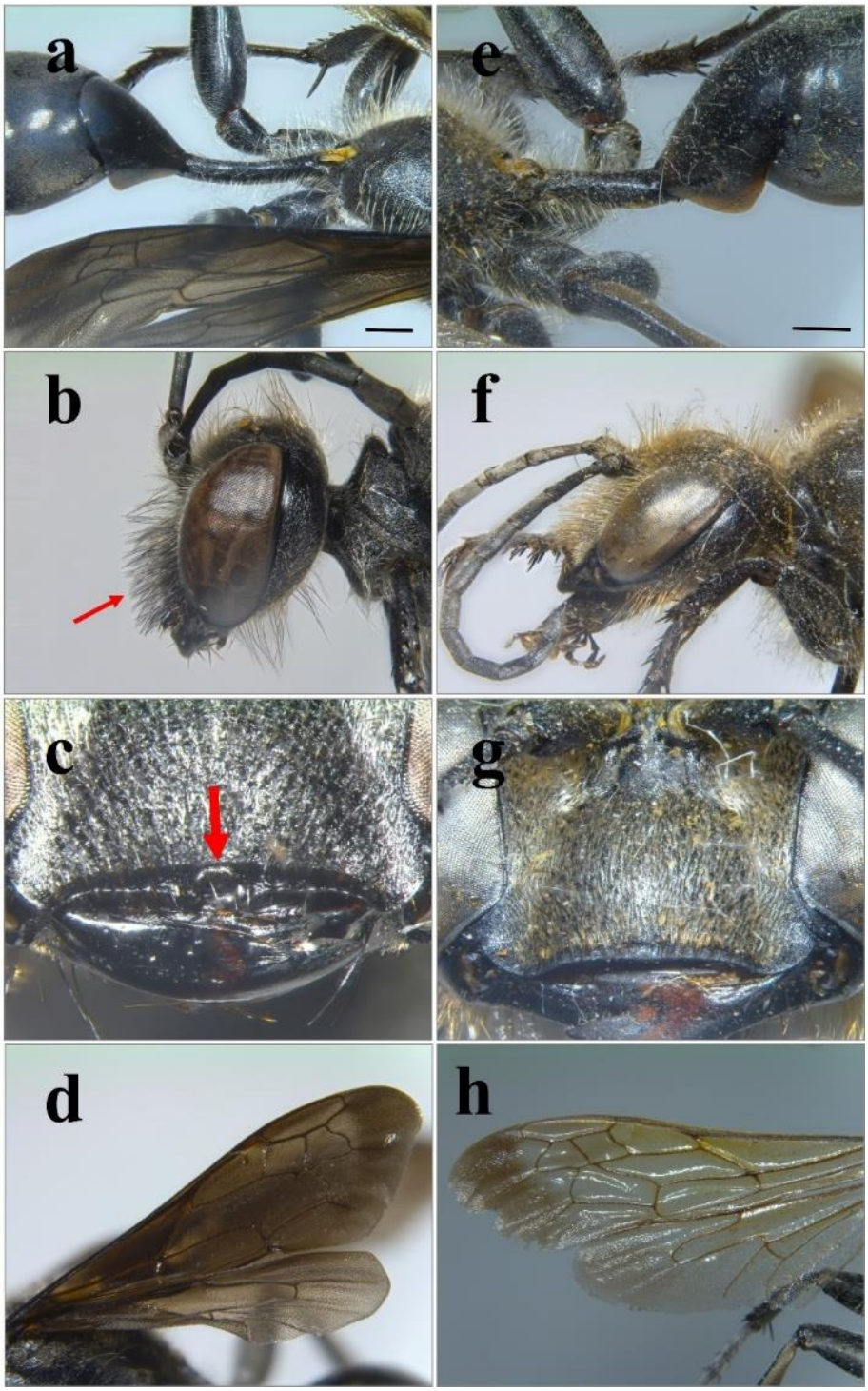


Figure 4. *Isodontia mexicana* ♀ (a - d) and *Isodontia paludosa* ♀ (e - h). a, e) Lateral view of the petiole (Scale bars: 2 mm); b, f) lateral view of head (erect black hairs indicated with red arrow); c, g) dorsal view of clypeus (deep median notch in the anterior border of the clypeus indicated with arrows); d, h) wings

In this study, *Isodontia mexicana* is reported for the first time in the Turkish fauna. This is the second invasive wasp species from the family Sphecidae to be introduced to Türkiye after *Sceliphron curvatum* (Drury, 1773) was discovered in 2015 [12]. So far, eight non-native species of Sphecidae have been successfully established in the Western Palaearctic fauna: Oriental *Chalybion bengalense* (Dahlbom, 1845), Nearctic *Chalybion californicum* (de Saussure, 1867), Palaearctic *Chalybion turanicum* (Gussakovskij, 1935), Nearctic *Chalybion zimmermanni* Dahlbom, 1843, Nearctic *Sceliphron caementarium* (Drury, 1773), Asian *Sceliphron curvatum* (F. Smith, 1870), Asian *Sceliphron deforme* (F. Smith, 1856), and Nearctic *Isodontia mexicana* [2]. *Sceliphron caementarium*, *Sceliphron deforme*, and *Chalybion turanicum* are relatively more common species than others, and their probability of being found in Türkiye is relatively high.

Isodontia mexicana was first recorded in Asia, in the Fars Province of southern Iran, in 2004 [13]. According to Ćetković et al. [3], the species was introduced to Western Asia in the early 1990s, and this occurrence seems likely to be a result of the import of equipment and cargo intended for American military units during the Gulf War. In this study, the species was detected in the Kocaeli Province of Türkiye, near the western edge of Asia Minor, one of the country's most important industrial cities. This city is located a great distance from the previously known localities of *Isodontia mexicana* in the surrounding regions: about 530 km, 600 km, and 750 km from Romania (Bucharest), Bulgaria (Sofia), and Greece (Peloponnese), respectively. There are geographical barriers between the detection point in Türkiye and these places, such as the sea and the Balkan Mountains. According to Polidori et al. [14], *Isodontia mexicana* could have spread further within Europe through active dispersal but could only have reached other continents through human trade. Since Kocaeli is a city with intensive imports due to industrial activities, the species may have reached the country through man, as in Iran.

4. Conclusion

In the future distribution projections of *Isodontia mexicana* prepared by Polidori et al. [14], the area invaded by the species is predicted to expand towards the North and East through active dispersal. The first discovery of the species in the Türkiye and its second record in the Asian continent show a tendency to further expand the species' range from Europe to the East. There are many suitable habitats in Türkiye for this wasp, which prefers to live in low-altitude regions with moderate temperatures and rainfall in the northern hemisphere. In this case, it can be predicted that *Isodontia mexicana* may spread eastwards soon in regions of Türkiye with this climate and geographical features, especially along the Black Sea coast.

Author Contributions

Drafted and wrote the manuscript, experimented, and result analysis. The author read and approved the final version of the paper.

Conflict of Interest

The author declares no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] W. J. Pulawski, Catalog of Sphecidae sensu lato (2024), https://www.calacademy.org/scientists/projects/catalog-of-sphecidae, Accessed 1 Jun 2024.

[2] A. V. Fateryga, S. P. Ivanov, M. V. Mokrousov, V. V. Fateryga, *The First record of the Far Eastern Grass-Carrying Wasp Isodontia nigella (F. Smith, 1856) (Hymenoptera: Sphecidae: Sphecinae) from the Crimea*, Far Eastern Entomologist (481) (2023) 14 - 24.

[3] R, Bohart, A, Menke, Sphecid wasps of the world. A generic revision. University California Press, Berkeley, Los Angeles, London, 1976.

[4] A. Ćetković, B. Čubrilović, M. Plećaš, A. Popović, D. Savić, L. Stanislavljević, *First records of the invasive American wasp Isodontia mexicana (Hymenoptera: Sphe-cidae) in Serbia,* Acta Entomologica Serbica 17 (1-2) (2012) 63–72.

[5] İ. Can, Y. Gülmez, *A faunistic study on the family Sphecidae (Hymenoptera) in the Upper Kelkit Valley with two new records and a checklist for Turkey,* Turkish Journal of Entomology 45 (3) (2021) 305-322.

[6] D. Gradinarov, *First records of the American wasps Sceliphron caementarium (Drury, 1770) and Isodontia Mexicana (de Saussure, 1867) (Hymenoptera: Sphecidae) from Bulgaria,* ZooNotes (118) (2017) 1- 4.

[7] W. Arens, *Die Grabwespen der Peloponnes (Hymenoptera: Apoidea) 1. Teil: Ampulicidae, Sphecidae, Crabronidae: Entomosericinae und Philanthinae; mit Beschreibung einer neuen Palmodes-Art,* Linzer Biologische Beitrage 49 (1) (2017) 619-655.

[8] A. F. Popa, I. E. Popescu, *First record in Romania of the genus Isodontia with Isodontia mexicana (Hymenoptera: Sphecidae)*, in: L. O. Popa, C. Adam, G. Chişamera, E. Iorgu, D. Murariu, O. P. Popa (Eds.), Proceedings of the International Zoological Congress of "Grigore Antipa" Museum, Bucharest, Romania, 2018, pp 53.

[9] J. Bitsch, Y. Barbier, S. F. Gayubo, K. Schmidt, M. Ohl, Hyménoptères Sphecidae d`Europe Occidentale, Volume 2. Faune de France 82, Paris, 1997.

[10] G. Pagliano, E. Negrisolo, Hymenoptera Sphecidae - Fauna d'Italia (in Italy), 40, Calderini 2005.

[11] D. G. Notton, *Grass-carrying wasp, Isodontia mexicana (de Saussure), genus and species new to Britain (Hymenoptera: Sphecidae),* British Journal of Entomology and Natural History (29) (2016) 241-245.

[12] Y. Gülmez, İ. Can, *First record of Sceliphron (Hensenia) curvatum (Hymenoptera: Sphecidae) from Turkey with notes on its morphology and biology*, North-Western Journal of Zoology 11(1) (2015) 174-177.

[13] M. Fallahzadeh, H. Ostovan, N. Saghaei, *First record of four Sphecid wasps from Iran,* Applied Entomology and Phytopathology 73 (2) 2005 43.

[14] C. Polidori, M. Nucifora, D. Sánchez-Fernández, *Environmental niche unflling but limited options for range expansion by active dispersion in an alien cavity-nesting wasp*, BMC Ecology 18 (2018) Article Number 36 12 pages.

Journal of New Results in Science

https://dergipark.org.tr/en/pub/jnrs

Research Article

Open Access

E-ISSN:1304-7981 https://doi.org/10.54187/jnrs.1522818

Genomic analysis of secondary metabolite biosynthesis gene clusters and structural characterization of terpene synthase and cytochrome P450 enzymes in *Zingiber officinale* Roscoe

Ummahan Öz[1]

Keywords:
Bioinformatic,
Gene cluster,
Protein homology,
Zingiber officinale

Abstract — This study uses bioinformatics approaches to elucidate the genetic basis of secondary metabolite biosynthesis in *Zingiber officinale* (*Z. officinale*). To this end, it identifies 44 secondary metabolite biosynthetic gene clusters and maps onto individual chromosomes, with chromosomes 1A and 8A exhibiting higher concentrations. Here, protein homology modeling provided insights into the structural characteristics of terpene synthases and Cytochrome P450 enzymes, shedding light on their potential roles in stress response and secondary metabolite production. Moreover, the identification of enzymes, such as (-)-kolavenyl diphosphate synthase TPS28 and cytochrome P450 93A3-like, opens up new possibilities for investigating the intricate pathways involved in terpene diversity and stress response mechanisms within *Z. officinale*. This study highlights the importance of understanding the molecular mechanisms underlying plant-derived bioactive compounds for pharmaceutical applications.

Subject Classification (2020):

1. Introduction

Plants synthesize two metabolites through metabolic pathways: primary and secondary metabolites [1]. Secondary metabolites play important roles in various functions, such as pharmaceutical production, plant protection, seed germination, signal transduction, and pollinator attraction [2]. Secondary metabolites can be grouped into three main categories according to their biosynthetic pathways: nitrogen-containing compounds synthesized in the tricarboxylic acid cycle pathway, phenolic compounds synthesized in the shikimate pathway, and terpenes synthesized in the mevalonic pathway [3]. Terpenes, the largest group of secondary metabolites with more than 22,000 compounds, are found in virtually all plants and are typically synthesized by isomer dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IDP) [4]. DMAPP and isopentenyl pyrophosphate (IPP) are converted to farnesyl diphosphate and geranyl diphosphate by farnesyl diphosphate synthase and geranyl diphosphate synthase enzymes, and these compounds serve as precursors for monoterpenes, sesquiterpenes, and triterpenes [5]. The enzymes terpene synthases (TPSs) are indicated to be responsible for the diversity of carbon structures in terpenoids [6].

Zingiber officinale (*Z. officinale*) Roscoe, a member of the Zingiberaceae family, is a plant native to Southeast Asia and has been utilized for over 3000 years, primarily in India, as both a flavorful spice in culinary practices and a traditional remedy in herbal medicine [7]. *Z. officinale* has anti-inflammatory, antimicrobial, antiviral,

[1]ummahanoz48@gmail.com (Corresponding Author)
[1] Department of Plant and Animal Production, Manisa Celal Bayar University, Manisa, Türkiye
Article History: Received: 26 Jul 2024 — Accepted: 26 Aug 2024 — Published: 31 Aug 2024

antioxidant, and antifungal effects. It is used in the treatment of nausea during pregnancy, fatigue or lack of energy, rheumatic diseases, and in the treatment of gastric, colorectal, ovarian, prostate, and breast cancers [8]. In addition to that, their uses have been identified in traditional medicine for conditions such as sprains, muscular aches, sore throats, cramps, vomiting, indigestion, dementia, helminthiasis, fever, infectious diseases, and hypertension [9]. *Z. officinale* contains over 400 compounds, including major biologically active compounds such as gingerols, paradols, and shogaols [10]. The components in *Z. officinale* inhibit the growth and proliferation of cancer cells, such as pancreatic, gastric, and colorectal cancer, by inducing apoptotic effects through various pathways on cancer cell lines [11]. The selection of *Z. officinale* as the material in the current article is also attributed to its diverse medicinal properties. Cytochrome P450 is a superfamily of enzymes that plays a crucial role in the metabolism of drugs [12], and *Z. officinale*, also known as ginger, is a plant used in pharmaceutical studies. For this reason, the current article provides a detailed examination of the Cytochrome P450 enzyme category. Additionally, more than half of all compounds present in the rhizome of *Z. officinale* are formed by terpenes [13]. Therefore, the enzymes belonging to the terpene synthase category, which play a role in terpene synthesis, have been detailed in the current article. Various investigations [14-25] have been conducted on the *Z. officinale*; however, as of now, no study has delved into the analysis of gene clusters related to secondary metabolite biosynthesis. As is well known, this species is a highly significant medicinal plant containing numerous important components in its composition. Identifying the gene clusters responsible for the synthesis of these components is crucial. Studying biosynthetic gene clusters is important for elucidating significant events such as synthesizing secondary metabolites, encoding drug resistance, and other critical biological processes. Identifying genes involved in the biosynthesis of secondary metabolites is essential for discovering new drug compounds and elucidating their biological activities [26].

Bioinformatics is a field that combines biology, mathematics, computer science, and statistics, playing a significant role in analyzing and interpreting large datasets [27]. Biosynthetic gene clusters typically encode the production of secondary metabolites and various software tools have been developed to detect biosynthetic gene clusters from DNA sequences [28]. In the determination of biosynthetic gene clusters for secondary metabolites, tools, such as 2metDB, antiSMASH, BAGEL, CLUSEAN, ClusterFinder, eSNaPD, EvoMining, GNP/Genome Search, GNP/PRISM, MIDDAS-M, MIPS-CG, NaPDoS, and SMURF are utilized [29]. Initially, gene annotations are obtained from the genome under study to identify conserved biosynthetic gene clusters. Following this, biosynthetic gene clusters are detected based on the core enzymes involved in the biosynthesis of secondary metabolites [30]. One of the aims of this study is to determine the gene clusters involved in the biosynthesis of secondary metabolites in *Z. officinale*.

Additionally, the study aims to utilize bioinformatics tools to identify these biosynthetic gene clusters and detect key enzymes, such as terpene synthases and Cytochrome P450. It also seeks to elucidate the protein structures of these enzymes. Identifying the individual gene clusters responsible for secondary metabolite biosynthesis on each chromosome will lay the groundwork for future research and support focused research efforts. The outcomes of this study are anticipated to significantly contribute to drug development research. This article represents a crucial step toward understanding plant-derived bioactive compounds at the molecular level and exploring their potential applications in the pharmaceutical industry.

2. Materials and Methods

To identify the gene clusters involved in the secondary metabolite biosynthesis of the *Zingiber officinale*, the species was initially scanned in the National Center of Biotechnology Information (NCBI) database. The genome assembly Zo_v1.1 was chosen due to the presence of the Reference Sequence (RefSeq). The obtained data was downloaded and uploaded to the plantiSMASH software [31] for detecting secondary metabolite biosynthetic gene clusters. PlantiSMASH software functions by identifying loci where at least two distinct enzyme classes from a minimum of three different enzyme subclasses converge, designating these as clusters.

Enzyme classes are assigned based on position-specific scoring matrices (pHMMs) tailored to each class. At the same time, sequence-based clustering uses the CD-HIT algorithm to count enzyme subclasses and assess sequence similarities above 50%.

The results obtained from the software have been downloaded to the computer, and each cluster has been separately filed. The data for each chromosome has been saved in Microsoft Excel, and the gene clusters related to secondary metabolite biosynthesis for each chromosome have been visualized in figure form. After identifying the gene clusters related to secondary metabolite biosynthesis in *Z. officinale*, enzymes in the Cytochrome P450 and Terpene synthase categories were copied into a separate Microsoft Excel file. Subsequently, on individual pages, "Enzyme ID" was assigned based on the chromosome number for each enzyme class. Subsequently, amino acid sequence data in .fasta format was downloaded from the NCBI database to determine protein homologies and uploaded to the Phyre2 (Protein Homology/Analogy Recognition Engine V 2.0) software [32]. The obtained results have been downloaded in a compressed (RAR) file format, and figures have been generated.

3. Results

According to information gathered from the NCBI platform, *Zingiber officinale* is reported to have 22 chromosomes, with a genome size of 3.1 Gb. As a result of the analyses conducted in the current article, no secondary metabolite biosynthetic gene cluster has been observed on chromosome 3B, chromosome 11A, and chromosome 11B. The secondary metabolite biosynthetic gene clusters found on other chromosomes, their respective locations, and the enzymes they contain are detailed in supplementary material (Tables 1 and 2). Based on the data acquired, the figures are presented below in sequential order of chromosomes.

Five secondary metabolite biosynthetic gene clusters have been identified on chromosome 1A. Cluster 1 is of the saccharide type, size 706.62 kb. Its location spans from 7328285 to 8034904 nt. In this cluster, enzyme categories include methyltransferase, CoA-ligase, and glycosyltransferase. Cluster 2 is of the polyketide type, spanning from 42651153 to 43781744 in location, with a size of 1130.59 kb. The enzyme categories detected within this cluster are ketosynthase, polyprenyl synthetase, and amino oxidase. Cluster 3 is of the saccharide type, while cluster 4 is of the polyketide type. Cluster 3 has a size of 179.36 kb, whereas cluster 4 is 217.54 kb. Cluster 3 is located between 146176694 and 146356055 nt. cluster 4 is found between 186274596 and 186492132 nt. In cluster 3, enzyme categories observed include CoA-ligase, glycosyltransferase, and fatty acid desaturase, while in cluster 4, aminotransferase, epimerase, and ketosynthase have been identified. Cluster 5 has been identified as the lignan type, with a size of 599.91 kb. It spans 192892652 to 193492563 and contains the enzyme categories dirigent enzymes and cytochrome P450 (Figure 1).

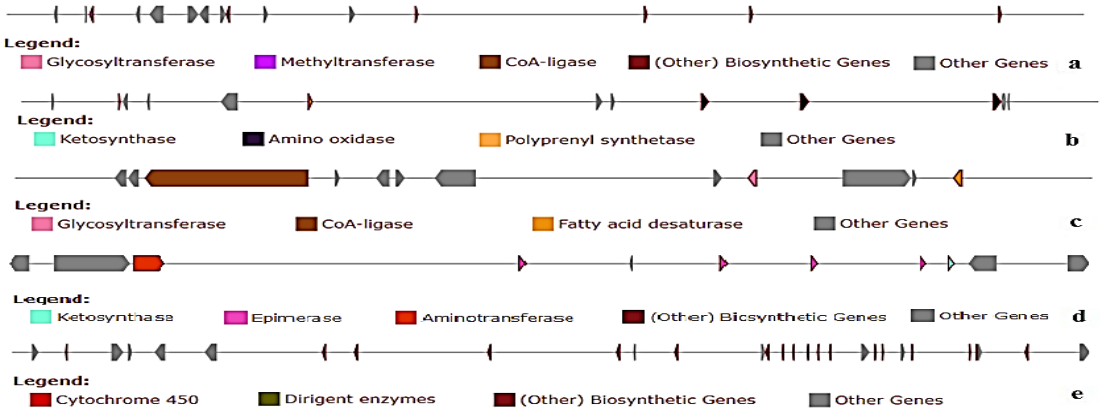


Figure 1. a) Cluster 1- saccharide, b) Cluster 2- polyketide, c) Cluster 3- saccharide, d) Cluster 4- polyketide, e) Cluster 5- lignan

Observations have revealed four clusters of secondary metabolite biosynthetic genes on Chromosome 1B, with the sequence of the first cluster being saccharide-terpene, the second cluster saccharide, the third cluster polyketide, and the fourth cluster lignan-type. Sizes of the clusters are as follows: Cluster 1 - 257.50 kb, Cluster 2 - 377.99 kb, Cluster 3 - 166.12 kb, and Cluster 4 - 483.87 kb. When evaluated in terms of locations, Cluster 1 spans from 129268146 nt. to 129525648 nt., Cluster 2 extends from 138835112 nt. to 139183102 nt., Cluster 3 covers the range from 163215785 nt. to 163381905 nt., and Cluster 4 is situated from 169767977 nt. to 170251844 nt. In addition, the enzyme categories they contain are as follows: Cluster 1 includes CoA ligase, terpene synthase, glycosyltransferase, and fatty acid desaturase; Cluster 2 consists of cytochrome P450 and glycosyltransferase; Cluster 3 involves aminotransferase, epimerase, and ketosynthase; Cluster 4 encompasses dirigent enzymes and cytochrome P450 (Figure 2).

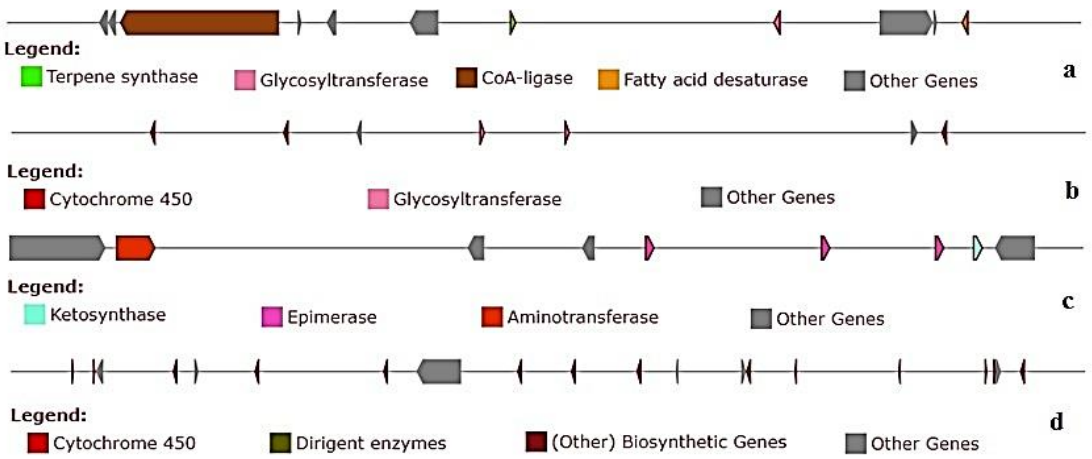


Figure 2. a) Cluster 1- saccharide-terpene, b) Cluster 2- saccharide, c) Cluster 3- polyketide, d) Cluster 4- lignan-type

According to the analysis results, two secondary metabolite biosynthetic gene clusters of polyketide and saccharide types have been identified on chromosome 2A. The size of Cluster 1 is 188.39 kb, while Cluster 2 is 327.19 kb. Cluster 1 spans from nucleotide position 77354198 to 77542590, whereas Cluster 2 is located from nucleotide position 125951799 to 126278987. The enzyme categories in Cluster 1 include ketosynthase and aminotransferase, while Cluster 2 comprises glycosyltransferase, cellulose synthase-like, and aminotransferase (Figure 3). A secondary metabolite biosynthetic gene cluster for polyketide has been identified on chromosome 2B, with a size of 188.97 kb. The location is 66468263-66657236 nt, and it contains enzymes belonging to the ketosynthase and aminotransferase categories (Figure 4).

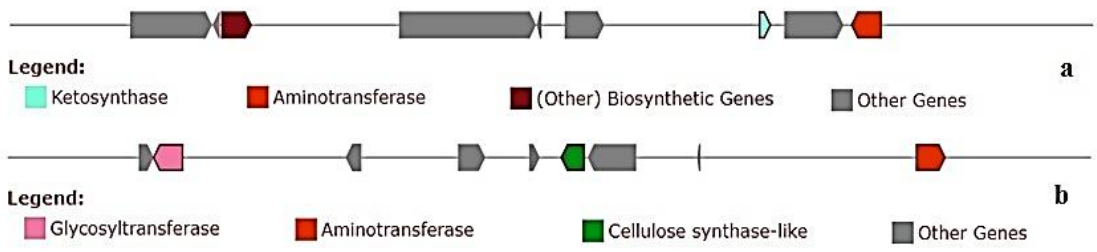


Figure 3. a) Cluster 1- polyketide, b) Cluster 2- saccharide

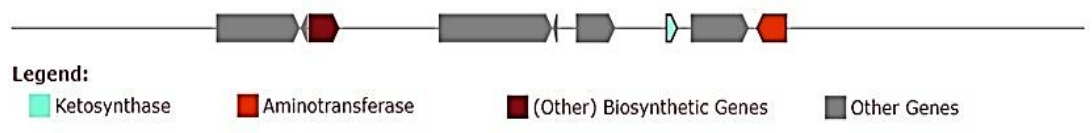


Figure 4. Polyketide gene cluster on chromosome 2B

On chromosome 3A, a secondary metabolite gene cluster of the terpene type has been identified, and its size is 186.24 kb. Furthermore, its location extends from nucleotide position 22347758 to 22533999. In addition, enzyme categories observed include dioxygenase, glycosyltransferase, and terpene synthase (Figure 5).

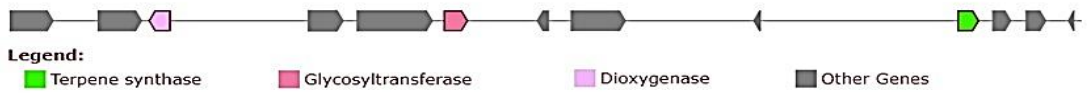


Figure 5. Terpene gene cluster on the chromosome 3A

Two secondary metabolite biosynthetic gene clusters have been identified on chromosome 4A. Cluster 1 is of the polyketide type with a size of 518.95 kb, while Cluster 2 is of the lignan-saccharide type with a size of 309.26 kb. The locations are as follows: Cluster 1 spans from nucleotide position 142297741 to 142816687, while Cluster 2 extends from nucleotide position 149885400 to 150194661. The enzyme categories are as follows: In Cluster 1, there are lipoxygenase, ketosynthase, oxidoreductase, and CoA-ligase; in Cluster 2, there are dirigent enzymes and glycosyltransferase (Figure 6). A secondary metabolite biosynthetic gene cluster of the polyketide type has been observed on chromosome 4B, with a size of 710.09 kb. This cluster is located between nucleotide positions 114712139 and 115422231 and has been found to contain enzyme categories such as ketosynthase, CoA-ligase, oxidoreductase, and lipoxygenase (Figure 7).

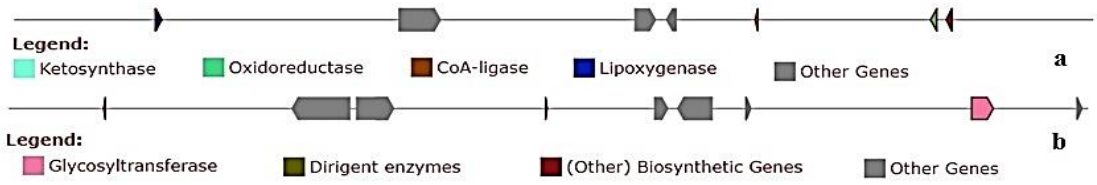


Figure 6. a) Cluster 1- polyketide, b) Cluster 2- lignan-saccharide

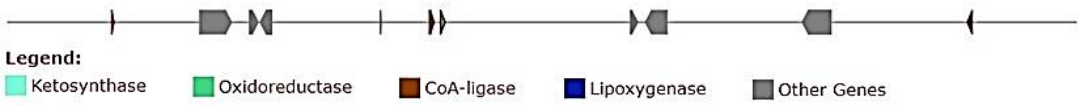


Figure 7. Polyketide gene cluster on chromosome 4B

On chromosome 5A, two distinct secondary metabolite biosynthetic gene clusters have been identified. Cluster 1 is of the polyketide type and has a size of 276.59 kb. Its location spans from nucleotide position 9574596 to 9851189 on the chromosome. Ketosynthase, methyltransferase, and fatty acid desaturase have been identified regarding enzyme categories. On the other hand, Cluster 2 is of the saccharide type and has a size of 177.40 kb. This cluster is located between nucleotide positions 124448054 and 124625455, and enzyme categories observed include glycosyltransferase and polyprenyl synthetase (Figure 8). A secondary metabolite gene cluster of the polyketide type has been analyzed on chromosome 5B, and its size is determined to be 247.96 kb. It spans from nucleotide position 10485332 to 10733290, and this cluster contains enzyme categories such as ketosynthase, methyltransferase, and fatty acid desaturase (Figure 9).

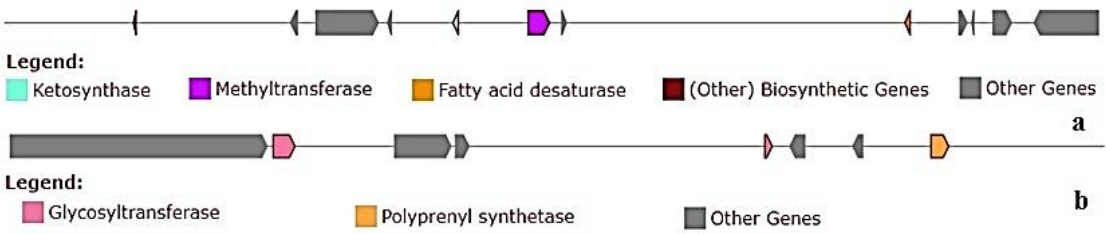


Figure 8. a) Cluster 1- polyketide, b) Cluster 2- saccharide

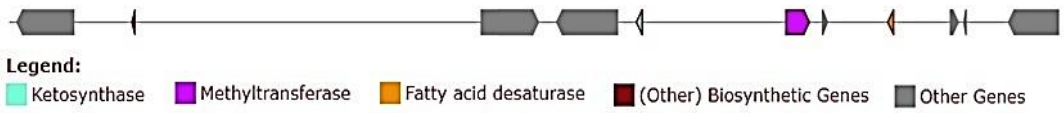


Figure 9. Polyketide gene cluster on chromosome 5B

On chromosome 6A, two types of secondary metabolite biosynthetic gene clusters have been identified, namely lignan and putative. Cluster 1 is of size 1384.62 kb, while Cluster 2 is 985.32 kb. The locations are as follows: Cluster 1 spans from nucleotide position 23617062 to 25001684, while Cluster 2 extends from nucleotide position 99578408 to 100563727. In addition, the enzyme categories are as follows: In Cluster 1, there are dirigent enzymes and Cytochrome 450, while in Cluster 2, there are lipoxygenase, oxidoreductase, and fatty acid desaturase (Figure 10). Lignan-type secondary metabolite biosynthetic gene cluster has been identified on Chromosome 6B. The size of this cluster is 727.33 kb, localized from 3845575 to 24572906 nt. Additionally, in this cluster, Scl acyltransferase, dirigent enzymes, and oxidoreductase enzyme categories have been identified (Figure 11).

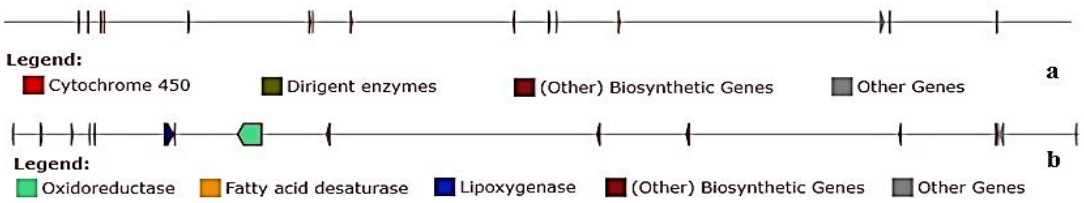


Figure 10. a) Cluster 1- lignan, b) Cluster 2- putative

Figure 11. Lignan gene cluster on chromosome 6B

Three clusters have been identified on Chromosome 7A, all of which are of saccharide type. Cluster 1 is 169.60 kb, Cluster 2 is 1514.08 kb, and Cluster 3 is 1267.58 kb in size. When evaluated in terms of their locations, Cluster 1 is observed between 84025 – 253630 nt., Cluster 2 between 46576688 – 48090770 nt., and Cluster 3 between 55068057 – 56335638 nt. Furthermore, the enzyme categories they contain are as follows: Cluster 1 includes glycosyltransferase, oxidoreductase, and BAHD acyltransferase; Cluster 2 includes Cytochrome 450, glycosyltransferase, CoA-ligase, and Scl acyltransferase; Cluster 3 includes oxidoreductase and glycosyltransferase (Figure 12). Two secondary metabolite biosynthetic gene clusters, one for saccharide and the other for polyketide type have been identified on Chromosome 7B. Cluster 1 is 355.22 kb in size, while Cluster 2 is 880.62 kb. Additionally, it has been identified that Cluster 1 extends from position 5277215 to 5632438 nucleotides, whereas Cluster 2 covers the region from position 26643208 to 27523833 nucleotides. Furthermore, the enzyme category in Cluster 1 is glycosyltransferase, while Cluster 2 includes Cytochrome 450, amino oxidase, and ketosynthase (Figure 13).

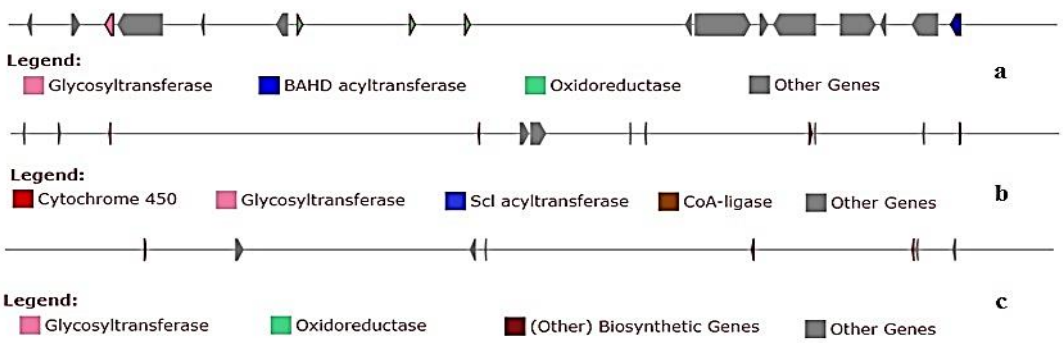


Figure 12. a) Cluster 1- saccharide, b) Cluster 2- saccharide, c) Cluster 3- saccharide

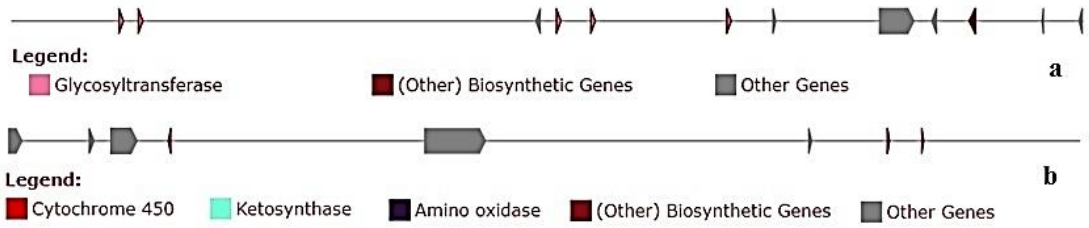


Figure 13. a) Cluster 1- saccharide, b) Cluster 2- polyketide

The analysis revealed five secondary metabolite biosynthetic gene clusters on chromosome 8A. Cluster 1, which is of the saccharide type, has a size of 240.37 kb and is located between nucleotides 10633148 and 10873514. The enzyme categories include aminotransferase, glycosyltransferase, and epimerase. Cluster 2 is of the saccharide type, with a size of 261.85 kb, and its location is between nucleotides 38297654 and 38559507. Additionally, this cluster contains enzymes in cytochrome P450, oxidoreductase, and glycosyltransferase categories. Cluster 3 is of the lignan-polyketide type, while Cluster 4 is of the saccharide type. Cluster 3 has a size of 414.08 kb, whereas Cluster 4 is 176.30 kb. Cluster 3 and Cluster 4 locations are 39117993-39532074 nt and 69732454-69908758 nt, respectively. Enzyme categories are as follows: Cluster 3 includes Ketosynthase, Dirigent enzymes, and Scl acyltransferase; Cluster 4, conversely, comprises epimerase, glycosyltransferase, and oxidoreductase. The final cluster, Cluster 5, is of the terpene-polyketide type and has a size of 819.03 kb. This cluster is located between nucleotides 130018371 and 130837398, and it contains enzymes in the categories of ketosynthase, terpene synthase, and methyltransferase (Figure 14). Four secondary biosynthetic gene clusters have been identified on Chromosome 8B, with Cluster 1 and Cluster 3 observed to be of saccharide type, Cluster 2 of polyketide type, and Cluster 4 of lignan-polyketide type. Cluster 1 is 232.74 kb, Cluster 2 is 196.19 kb, Cluster 3 is 372.31 kb and Cluster 4 is 205.09 kb in size. Cluster 1 spans from 8,488,692 to 8,720,436 nt, Cluster 2 from 27,405,491 to 27,601,678 nt, Cluster 3 from 32,315,114 to 32,687,419 nt, and Cluster 4 from 33,029,911 to 33,234,999 nt. The enzyme categories are as follows: Cluster 1 includes Aminotransferase, Glycosyltransferase, and Epimerase. Cluster 2 consists of Epimerase, Ketosynthase, and Dioxygenase. Cluster 3 encompasses Lipoxygenase, Cytochrome P450, Oxidoreductase, and Glycosyltransferase. Finally, Cluster 4 includes Ketosynthase, Dirigent enzymes, and Scl acyltransferase (Figure 15).

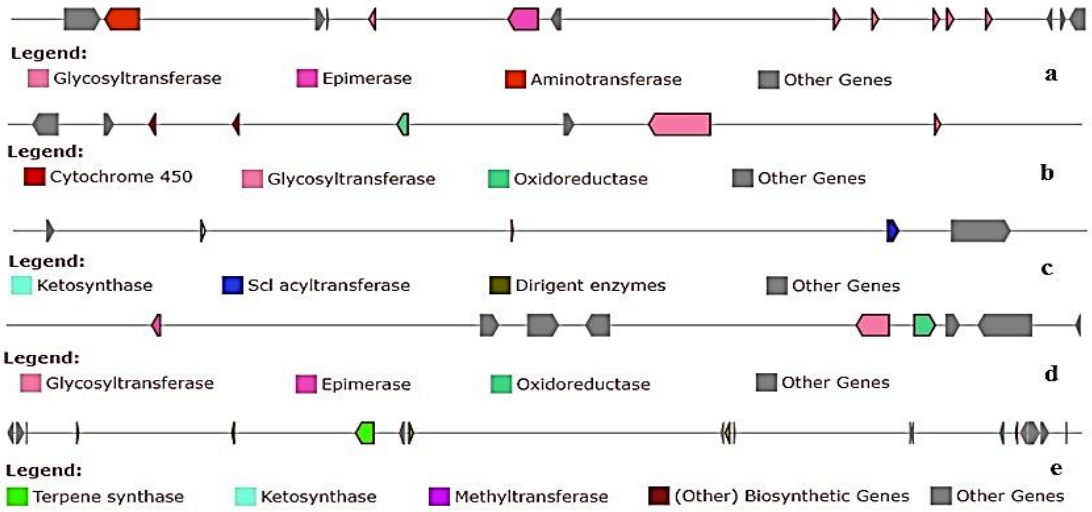


Figure 14. a) Cluster 1- saccharide, b) Cluster 2- saccharide, c) Cluster 3- lignan-polyketide, d) Cluster 4- saccharide, e) Cluster 5- terpene-polyketide

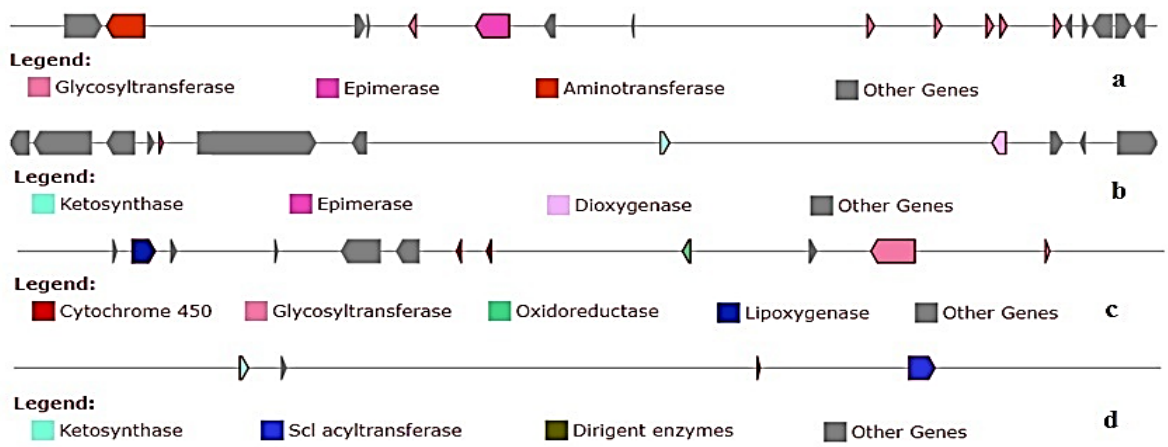


Figure 15. a) Cluster 1- saccharide, b) Cluster 2- polyketide, c) Cluster 3- saccharide, d) Cluster 4- lignan-polyketide

On Chromosome 9A, three secondary metabolite biosynthetic gene clusters have been identified. Cluster 1 is of saccharide type, Cluster 2 is of putative type, and Cluster 3 is of polyketide type. Cluster 1 is 400.98 kb, Cluster 2 is 228.58 kb, and Cluster 3 is 183.61 kb in size. When analyzed in terms of locations, it has been observed that Cluster 1 spans from nucleotide position 903,109 to 1,304,092, Cluster 2 extends from 114,121,928 to 114,350,508 nt, and Cluster 3 ranges from 133,439,172 to 133,622,782 nt. Additionally, in Cluster 1, enzyme categories include glycosyltransferase and BAHD acyltransferase; in Cluster 2, Cytochrome P450, Amino oxidase, Methyltransferase, and Scl acyltransferase are identified; while in Cluster 3, Cytochrome P450 and Methyltransferase enzyme categories have been determined (Figure 16). Chromosome 9B carries three secondary metabolite gene clusters, with the first being saccharide type and the third being polyketide type. In terms of size, Cluster 1 is 310.42 kb, Cluster 2 is 560.78 kb, and Cluster 3 is 172.30 kb. Cluster 1 is located between nucleotides 1418390 and 1728814, Cluster 2 spans from nucleotides 102414401 to 102975185, and Cluster 3 is positioned between nucleotides 117928197 and 118100500. Additionally, the enzyme categories they contain are as follows: Cluster 1 includes glycosyltransferase and BAHD acyltransferase; Cluster 2 comprises Cytochrome P450, amino oxidase, Scl acyltransferase, and glycosyltransferase; and Cluster 3 encompasses ketosynthase, Cytochrome P450, and methyltransferase (Figure 17).

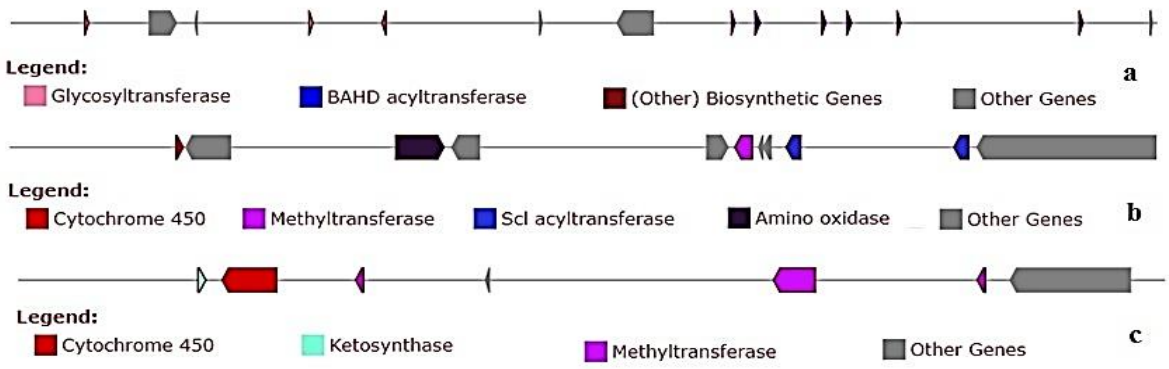


Figure 16. a) Cluster 1- saccharide, b) Cluster 2- putative, c) Cluster 3- polyketide

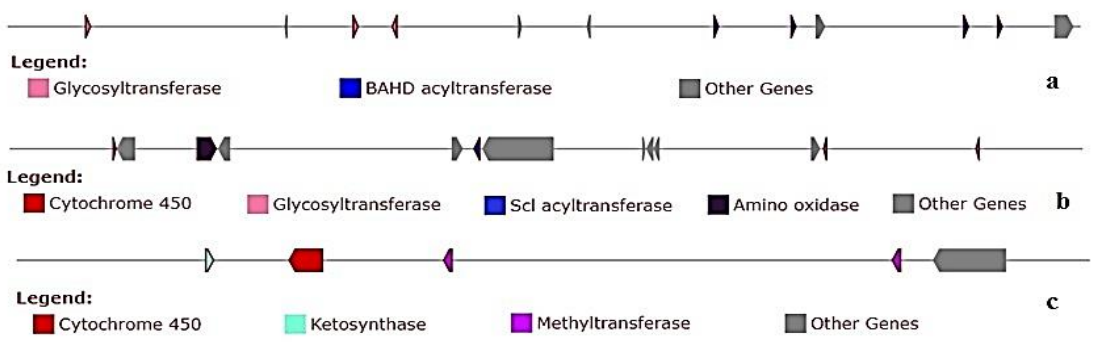


Figure 17. a) Cluster 1- saccharide, b) Cluster 2- saccharide, c) Cluster 3- polyketide

The terpene gene cluster has been detected on Chromosome 10A, with a size of 2278.99 kb. Its location spans from 89454170 to 91733163. The enzyme categories include terpene synthase and Cytochrome 450 (Figure 18). The secondary metabolite gene cluster on Chromosome 10B has been identified as terpene. It has a size of 1709.79 kb and is observed to be situated between nucleotides 86121756 and 87831544. Additionally, terpene synthase and Cytochrome 450 enzyme categories were detected within this cluster (Figure 19).

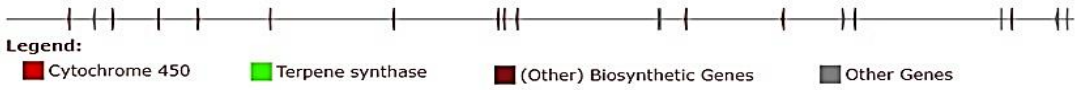


Figure 18. Terpene gene cluster on the chromosome 10A

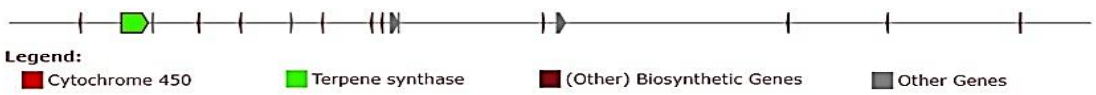


Figure 19. Terpene gene cluster on chromosome 10B

After analyzing secondary metabolite biosynthetic gene clusters, the locations of enzymes in the terpene synthase and Cytochrome P450 categories on the chromosome have been separately addressed in supplementary material (Tables 3 and 4). In the Cytochrome P450 category, the following enzymes have been observed: alpha-humulene 10-hydroxylase-like, cytochrome P450 93A3-like, flavonoid 3'-monooxygenase CYP75B137-like, alkane hydroxylase MAH1-like, cytochrome P450 734A6-like, cytochrome P450 71A9-like, cytochrome P450 77A2-like, abscisic acid 8'-hydroxylase 4-like, and cytochrome P450 94B3-like. Additionally, the presence of (-)-kolavenyl diphosphate synthase TPS28, chloroplastic-like, bifunctional isopimaradiene synthase, chloroplastic-like, (3S,6E)-nerolidol synthase 1-like, (3S,6E)-nerolidol synthase 2, chloroplastic/mitochondrial-like, and terpene synthase 13 within the category of terpene synthase enzymes has been identified.

Protein homology modeling was performed using the Phyre2 software, focusing on Cytochrome P450 and Terpene synthases. Observations reveal that the helix-loop-helix structure predominates in terpene synthases. ZoCYP1, ZoCYP2, ZoCYP3, ZoCYP4, ZoCYP5, ZoCYP6, ZoCYP8, ZoCYP10, ZoCYP11, ZoCYP12, ZoCYP13, ZoCYP14, ZoCYP15, ZoCYP17, ZoCYP19, ZoCYP33, and ZoCYP36 consist of α helices, numerous antiparallel β sheets, and long loops. ZoCYP7, ZoCYP16, ZoCYP22, ZoCYP24, ZoCYP25, ZoCYP26, and ZoCYP31 exhibit a higher presence of antiparallel β sheets. Likewise, ZoCYP23, ZoCYP29, and ZoCYP32 demonstrate the highest abundance of antiparallel β sheets, whereas ZoCYP9 and ZoCYP34 exhibit the least. Additionally, α helices, antiparallel β sheets, β turns, and long loops have been observed in ZoCYP9, ZoCYP18, ZoCYP20, ZoCYP21, ZoCYP27, ZoCYP28, and ZoCYP30. Furthermore, ZoCYP35 is characterized by numerous long loops, with a scarcity of α helix structure and antiparallel β sheets (Figure 20). After protein homology modeling, transmembrane helix prediction was performed. The analyses revealed a transmembrane helix structure in all ZoCYPs. The positions where they are located in the amino acid sequences

are given in Figure 21. When examined in terms of transmembrane helix percentage, it was determined that ZoCYP34 ranks the highest at 19%, followed by ZoCYP20 at 13%, ZoCYP25 at 13%, and ZoCYP26 at 13%. A notable point is that while the β-strand structure ranges between 6-9% in others, it is 20% in ZoCYP35.

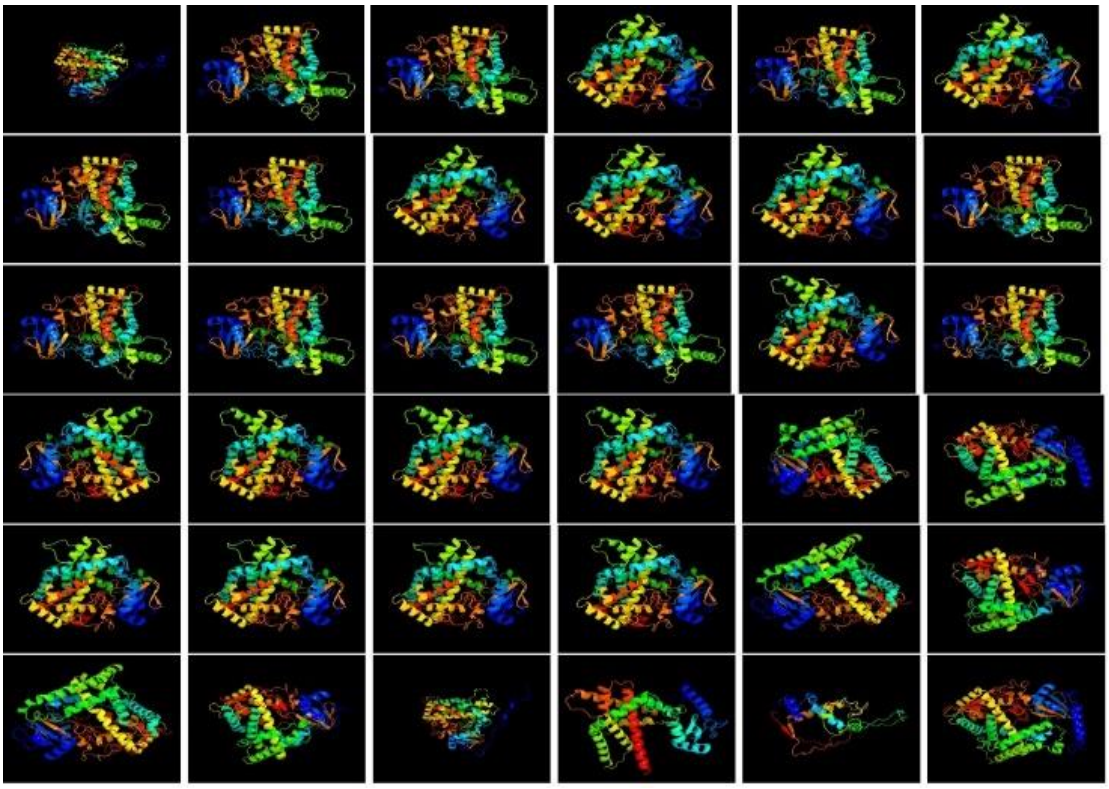


Figure 20. Predicted three-dimensional structures of ZoCYPs (sorted by protein number)

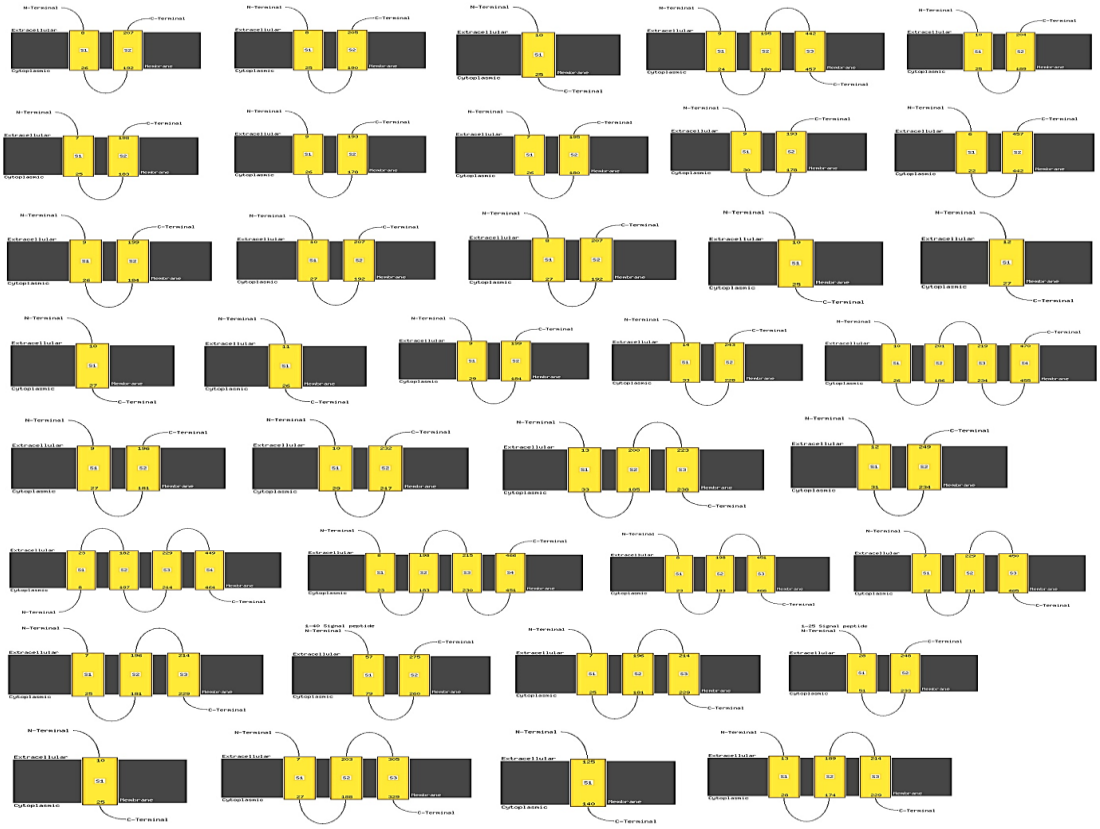


Figure 21. Transmembrane helix prediction for ZoCYPs (From ZoCYP1 to ZoCYP36 sequentially)

As a result of protein homology analysis in terpene synthases, it has been determined that the helix-loop-helix structure is dominant, and all ZoTPSs consist of α helices and long loops (Figure 22). As a result of transmembrane helix prediction analysis, a transmembrane helix structure has been identified in ZoTPS3, ZoTPS4, ZoTPS5, ZoTPS6, ZoTPS8, ZoTPS9, ZoTPS10, ZoTPS12, ZoTPS13, ZoTPS14, ZoTPS15, ZoTPS16, ZoTPS18, ZoTPS19, ZoTPS20, ZoTPS21, ZoTPS22, ZoTPS23, ZoTPS24, ZoTPS25, and ZoTPS26. The positions of these structures are detailed in Figure 23. When analyzing the amino acid sequences, ZoTPS5 exhibited the highest proportion of transmembrane helix at 13%, varying between 6% and 13% in others.

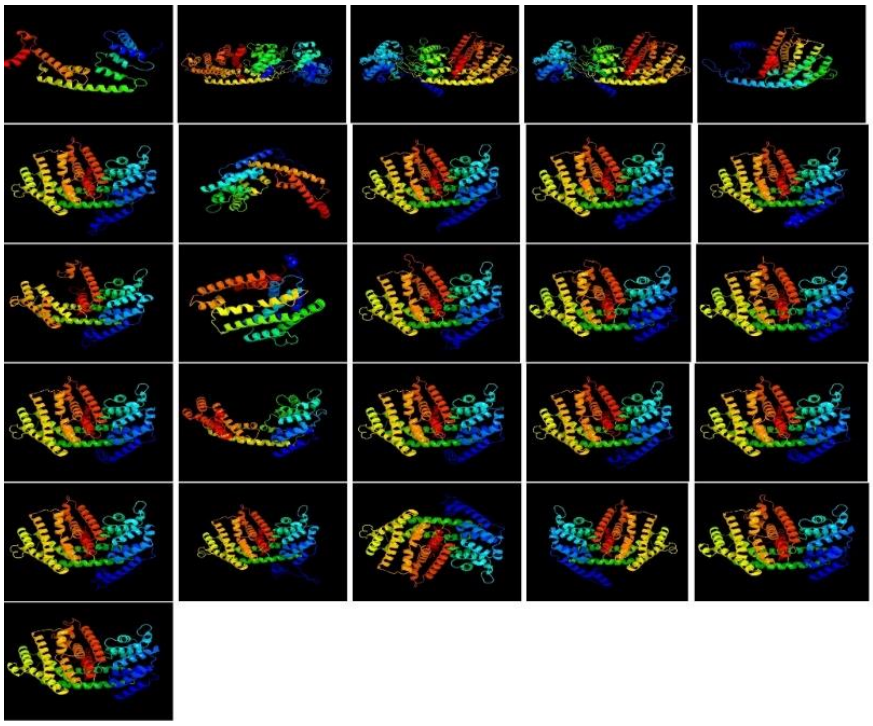


Figure 22. Predicted three-dimensional structures of ZoTPSs (sorted by protein number)

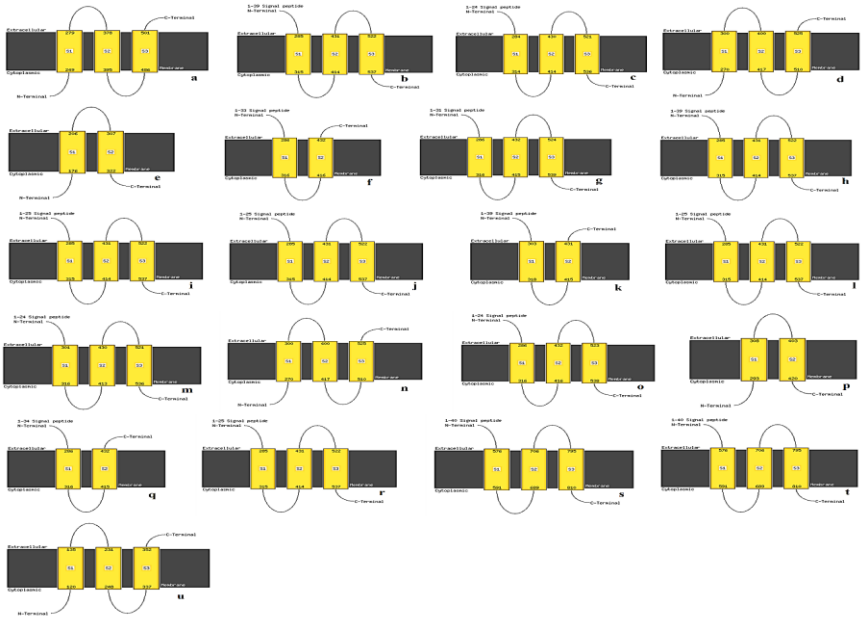


Figure 23. Transmembrane helix prediction for ZoTPSs: a)ZoTPS3, b)ZoTPS4, c)ZoTPS5, d)ZoTPS6, e)ZoTPS8, f)ZoTPS9, g)ZoTPS10, h)ZoTPS12, i)ZoTPS13, j)ZoTPS14, k)ZoTPS15, l) ZoTPS16, m)ZoTPS18, n)ZoTPS19, o) ZoTPS20, p)ZoTPS21, q)ZoTPS22, r)ZoTPS23, s)ZoTPS24, t)ZoTPS25, u)ZoTPS26

4. Discussion

Secondary metabolites in plants play important roles in innate immunity, plant growth, and development processes, developing responses to abiotic stress conditions, and signaling defense responses [33]. The current study involves the analysis of biosynthetic gene clusters responsible for these secondary metabolites, which are medically significant and involved in physiological processes within the plant. As a result of the analyses conducted, 44 secondary metabolite biosynthetic gene clusters were identified in *Z. officinale*, with the highest number of gene clusters being determined on chromosomes 1A and 8A. The observed types of secondary metabolite biosynthetic gene clusters in this plant include saccharide, polyketide, lignan, saccharide-terpene, terpene, lignan-saccharide, putative, lignan-polyketide, and terpene-polyketide. The highest number of clusters is in the saccharide type (17 clusters), while the least is in the saccharide-terpene (1), lignan-saccharide (1), and terpene-polyketide types. Research on plants' secondary metabolite biosynthetic gene cluster analyses is not extensively available. In another study [34], a total of 40 gene clusters were identified in *Citrus sinensis* (L.) Osbeck species, with 12 being saccharide, 11 putative, 6 terpene, 3 alkaloid, 3 lignan, 2 polyketide, 1 terpene-alkaloid, 1 saccharide-terpene, and 1 terpene-saccharide-polyketide type. Another study [31] found that the following numbers of biosynthetic gene clusters were detected in various plant species: *Arabidopsis thaliana* (L.) Heynh. (45), *Arabis alpina* Georgi (35), *Arachis duranensis* Krapov. and W. C. Greg. (34), *Arachis ipaensis* Krapov. & W.C.Greg. (34), *Beta vulgaris* L. (34), *Brachypodium distachyon* Roem. & Schult. (29), *Brachypodium stacei* Catalán, Joch. Müll., L. A. J. Mur and T. Langdon (43), *Brassica napus* L. (68), *Brassica oleracea* L. (34), *Brassica rapa* L. (51), *Cajanus cajan* (L.) Huth (23), *Camelina sativa* Boiss. (88), *Cicer arietinum* L. (28), *Cucumis sativus* L. (30), *Elaeis guineensis* Jacq. (18), *Fragaria vesca* L. (35), *Glycine max* (L.) Merr. (76), *Gossypium raimondii* Ulbr. (47), *Malus domestica* (Suckow) Borkh. (51), *Manihot esculenta* Crantz (36), *Medicago truncatula* Gaertn. (54), *Oryza brachyantha* A. Chev. and Roehr. (37), *Oryza sativa* Indica (54), *Oryza sativa* Japonica (46), *Panicum virgatum* L. (53), *Phaseolus vulgaris* L. (56), *Populus trichocarpa* Torr. and A. Gray ex Hook. (48), *Prunus persica* (L.) Batsch (33), *Salix purpurea* L. (33), *Sesamum indicum* L. (41), *Solanum lycopersicum* L. (45), *Solanum tuberosum* L. (51), *Sorghum bicolor* (L.) Moench (54), *Theobroma cacao* L. (48), *Vigna radiata* (L.) R. Wilczek (42), *Vitis vinifera* L. (46), *Zea mays* L. (34). When compared to other plants, *Z. officinale* has been observed to contain a greater number of secondary metabolite biosynthetic gene clusters in some plants, fewer in others, and a similar number in some. The superior aspect of the current article compared to others is the analysis of the location of each secondary biosynthetic gene cluster on individual chromosomes. In the current article, the location of each gene cluster has been determined.

Plant gene clusters possess regulatory enzymes such as Cytochrome P450s (CYP450s), UDP-glycosyl transferases, acyl transferases, short-chain alcohol dehydrogenases, transaminases, and decarboxylases to modify the backbone of signature metabolites. Reports indicate that certain CYP450s and terpene synthases exhibit non-random associations and are distributed among various gene clusters [35]. In the current article, enzyme categories such as Amino oxidase, Aminotransferase, BAHD acyltransferase, Cellulose synthase-like, CoA-ligase, Cytochrome P450, Dioxygenase, Dirigent enzymes, Epimerase, Fatty acid desaturase, Glycosyltransferase, Ketosynthase, Lipoxygenase, Methyltransferase, Oxidoreductase, Polyprenyl synthetase, Scl acyltransferase, and Terpene synthase have been identified in *Z. officinale*. The enzymes within each category and their locations on the chromosome are provided in supplementary material (Tables 1 and 2). Upon thoroughly examining this table. It is evident that apart from the terpene synthase and Cytochrome P450 categories discussed in detail in this article, *Z. officinale* also exhibits numerous important enzymes such as UDP-glycosyltransferase (uridine diphosphate-dependent glycosyltransferase, UGT), curcumin synthase 1, curcumin synthase 3-like, and others. UGTs are a crucial enzyme family involved in the biosynthesis of natural products and play significant roles in pharmaceuticals and developing responses to abiotic stress conditions [36]. In curcumin synthesis, it is mediated by the curcumin synthase gene family, which consists of three members: curcumin synthase 1, curcumin synthase 2, and curcumin synthase 3. Curcumin synthase 2 and

curcumin synthase 3 belong to the category of type III polyketide synthases and exhibit activity similar to that of curcumin synthase 1, participating in the synthesis of curcumin [37]. Curcumin is extracted from *Curcuma longa* L. and possesses numerous medical effects such as antioxidant, neuroprotective, anti-inflammatory, antiproliferative, proapoptotic, chemotherapeutic, chemopreventive, antimalarial, and antiparasitic effects [38]. Identifying curcumin synthase 1 in *Z. officinale* is an interesting and significant finding. No available data suggests the synthesis of curcumin synthase 1 from plants other than *C. longa*. Curcumin synthase 1 in *Z. officinale* has been determined to be located on chromosome 9A and chromosome 9B, with their respective localities provided in supplementary material (Tables 1 and 2). The oil of *Z.officinale* contains monoterpenoids such as cineole, b-phellandrene, (+)-camphene, geraniol, and sesquiterpenoids such as b-sesquiphellandrene, a-zingiberene, zingiberol, ar-curcumene [39]. The presence of ar-curcumene in *Z. officinale* may be associated with the data obtained in the current article. Advanced studies could be conducted to search for Curcumin Synthases on these chromosomes and compare them with those of *C. longa*. Additionally, the presence of curcumin in *Z. officinale* could be determined, and its medicinal effects could be tested on subjects to contribute to drug development efforts.

If we focus on the enzyme categories highlighted in the current article, it is indicated that Cytochrome P450 and terpene synthase gene pairs could form the basis of terpene diversity in eudicots [40]. Terpenes and terpenoids are structures that play significant roles in cellular processes such as photosynthesis, electron transport, cell wall formation, membrane fluidity, and signaling. They are synthesized through the action of terpene synthase enzymes [41]. In the current article, it has also been determined that terpene synthase enzymes are predominantly located on chromosomes 10A and 10B. When conducting studies related to terpene synthases in *Z. officinale*, it would be advantageous to focus on chromosomes 10A and 10B. The enzyme bifunctional isopimaradiene synthase, identified in *Z. officinale*, is also reported to be present in *Picea abies* (L.) H.Karst. (Norway spruce), *Picea sitchensis* (Bong.) Carrière (Sitka spruce), and *Abies balsamea* (L.) Mill. (Balsam fir) [42]. When research on isopimaradiene synthase is conducted, it is observed that studies involving this enzyme are predominantly focused on gymnosperms. The presence of this enzyme in *Z. officinale* will contribute to the literature. One of the enzymes identified in the current paper involved in terpene synthesis is (-)-kolavenyl diphosphate synthase TPS28. While this enzyme catalyzes the conversion of geranylgeranyl diphosphate to (-)- kolavenyl diphosphate, its presence is mentioned in both *Salvia divinorum* Epling & Játiva and *Scutellaria barbata* D.Don. [43-44]. In addition, the current article also identifies (3S,6E)-nerolidol synthase 1, (3S,6E)-nerolidol synthase 1-like, and (3S,6E)-nerolidol synthase 2 enzymes. Nerolidol, a sesquiterpene alcohol involved in numerous pharmacological activities, and supporting the current article, a study [45] mentions the presence of the *(3S,6E)-nerolidol synthase 1-like* gene in *Z. officinale*. Undeniably, the (3S,6E)-nerolidol synthases present in this plant contribute to the medicinal properties of *Z. officinale*. If these enzymes are to be isolated, chromosome 10A and chromosome 10B analysis should be conducted.

Cytochrome P450 enzymes participate in the biosynthesis of important secondary metabolites such as alkaloids, terpenoids, phenylpropanoids, and plant hormones while also providing tolerance to abiotic stress conditions [46]. One of these enzymes, cytochrome P450 93A3-like, has been identified in *Zingiber officinale.* Genes synthesizing the same enzyme have also been observed in *Sesamum indicum* L., an important oilseed plant, due to *Macrophomina phaseolina* interaction [47]. Additionally, it has been reported that cytochrome P450 734A6-like is expressed under salt stress in *Setaria italica* (L.) P.Beauv. and under drought stress in *Triticum aestivum* L. [48-49]. Cytochrome P450 71A9-like is up-regulated in *Triticum aestivum* under drought stress upon application of 5-aminolevulinic acid [50]. Similarly, in tomatoes, the association of cytochrome P450 71A9-like with other proteins in response development has been documented under salinity and viral stresses [51]. Additionally, CYP75B137, involved in flavonoid biosynthesis, has been reported to be up-regulated under drought stress [52]. In *Z. officinale*, a flavonoid 3'-monooxygenase CYP75B137-like enzyme has been identified, suggesting a potential similarity in function to the CYP75B137 enzyme. Another enzyme identified, alkane hydroxylase MAH1-like, increased expression when exposed to stress in a study [53],

indicating an effort to adapt to the environment. Additionally, another study [54] mentioned that alkane hydroxylase MAH1-like is involved in important functions in the biosynthesis of ketones secondary alcohols. Data also indicates this enzyme is down-regulated under heat stress in *Michelia macclurei* Dandy [55]. Abscisic acid 8'-hydroxylase 4 is an enzyme involved in abscisic acid biosynthesis [56]. Abscisic acid 8'-hydroxylase 4-like enzyme has also been identified in *Z. officinale*. This enzyme is also necessarily involved in abscisic acid biosynthesis. In the current article, another enzyme identified in the Cytochrome P450 category in *Z. officinale* is alpha-humulene 10-hydroxylase-like. In a study [57], it was mentioned that the gene named *Zoff283768* in *Z. officinale* encodes the alpha-humulene 10-hydroxylase enzyme, which catalyzes the conversion of alpha-humulene to 8-hydroxy-alpha-humulene in zerumbone biosynthesis. The reporting of the presence of the alpha-humulene 10-hydroxylase enzyme in this article supports the data obtained in the current study. When examining the enzymes in the Cytochrome P450 category in *Z. officinale*, it has been concluded that all of them are associated with stress response. This suggests the potential for these enzymes to enhance the resistance of *Z. officinale* to biotic and abiotic stress conditions.

In the current article, protein homologies of terpene synthases and enzymes in the cytochrome P450 category have also been identified. It has been observed that, unlike in ZoTPSs, antiparallel β sheets are present in the structure of ZoCYPs. Furthermore, while transmembrane helix structures have been identified in all ZoCYPs, they have been predicted in some ZoTPSs. Many crucial cellular functions are carried out through membrane proteins found in the cell membrane, and most transmembrane proteins cross the membrane with transmembrane helices [58-59]. The presence of these structures in ZoCYPs suggests significant roles in maintaining the balance between intracellular and extracellular. Additionally, it has been previously mentioned that enzymes in this enzyme category in *Z. officinale* have roles in stress response development, and it is also considered that the contribution of the transmembrane helix structure to this response development is possible.

5. Conclusion

This research focused on the gene clusters responsible for synthesizing secondary metabolites in *Zingiber officinale* and the important enzymes involved. The findings reveal the distribution of gene clusters across different chromosomes of *Z. officinale*, particularly highlighting the concentration of terpene synthases and Cytochrome P450 enzymes on specific chromosomes. This contributes to the localization and understanding of the genetic components involved in synthesizing secondary metabolites in the plant. Furthermore, the analysis of protein structures of these enzymes provides valuable insights into their functionality and interactions. For instance, protein modeling of terpene synthases elucidates their structural properties, shedding light on the potential effects on the plant's terpene synthesis. The results of this study could serve as a foundation for further research in plant biochemistry, genetics, and pharmacology. Investigating gene clusters and enzymes is an important step towards understanding the biological basis of herbal medicines and improving their utilization in the pharmaceutical industry. Deepening the understanding of the genetic structure of *Z. officinale* may also encourage applicable practices in herbal medicine development and agricultural industries to enhance productivity.

Author Contributions

The author read and approved the final version of the paper.

Conflict of Interest

The author declares no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

Supplementary Material

https://dergipark.org.tr/en/download/journal-file/31863

References

[1] Anjali, S. Kumar, T. Korra, R. Thakur, R. Arutselvan, A. S. Kashyap, Y. Nehela, V. Chaplygin, T. Minkina, C. Keswani, *Role of plant secondary metabolites in defence and transcriptional regulation in response to biotic stress,* Plant Stress 8 (2023) Article Number 100154 19 pages.

[2] X. Zhan, Z. Chen, R. Chen, C. Shen, *Environmental and genetic factors involved in plant protection-associated secondary metabolite biosynthesis pathways*, Frontiers in Plant Science 13 (2022) Article Number 877304 14 pages.

[3] R. Jan, S. Asaf, M. Numan, Lubna, K. M. Kim, *Plant secondary metabolite biosynthesis and transcriptional regulation in response to biotic and abiotic stress conditions*, Agronomy 11(5) (2021) Article Number 968 31 pages.

[4] J. M. Al-Khayri, R. Rashmi, V. Toppo, P. B. Chole, A. Banadka, W. N. Sudheer, P. Nagella, W. F. Shehata, M. Q. Al-Mssallem, F. M. Alessa, M. I. Almaghasla, A. A. S. Rezk, *Plant secondary metabolites: The weapons for biotic stress management*, Metabolites 13 (6) (2023) Article Number 716 37 pages.

[5] S. Zhang, L. Zhang, H. Zou, L. Qiu, Y. Zheng, D. Yang, Y. Wang, *Effects of light on secondary metabolite biosynthesis in medicinal plants*, Frontiers in Plant Science 12 (2021) Article Number 781236 16 pages.

[6] E. Alicandri, A. R. Paolacci, S. Osadolor, A. Sorgonà, M. Badiani, M. Ciaffi, *On the evolution and functional diversity of terpene synthases in the pinus species: A review*, Journal of Molecular Evolution 88 (3) (2020) 253-283.

[7] Z. Yang, Z. Guo, J. Yan, J. Xie, *Nutritional components, phytochemical compositions, biological properties, and potential food applications of ginger (Zingiber officinale): A comprehensive review*, Journal of Food Composition and Analysis 128 (2024) Article Number 106057 16 pages.

[8] F. Yousfi, F. Abrigach, J. D. Petrovic, M. Sokovic, M. Ramdani, *Phytochemical screening and evaluation of the antioxidant and antibacterial potential of Zingiber officinale extracts*, South African Journal of Botany 142 (2021) 433–440.

[9] S. Xiang, Q. Jian, W. Chen, Q. Xu, J. Li, C. Wang, R. Wang, D. Zhang, J. Lin, C. Zheng, *Pharmacodynamic components and mechanisms of ginger (Zingiber officinale) in the prevention and treatment of colorectal cancer*, Journal of Ethnopharmacology 324 (2024) Article Number 117733 20 pages.

[10] I. Boarescu, R. M. Pop, P. M. Boarescu, I. C. Bocșan, D. Gheban, A. E. Bulboacă, A. D. Buzoianu, S. D. Bolboacă, *Ginger (Zingiber officinale) root capsules enhance analgesic and antioxidant efficacy of diclofenac sodium in experimental acute inflammation*, Antioxidants 12 (3) (2023) Article Number 745 19 pages.

[11] N. T. Anh Nga, S. Sathiyavimal, L. A. Al-Humaid, N. D. Al- Dahmash, J. Lee, S. Barathi, G. K. Jhanani, *Deciphering the anticancer, anti-inflammatory and antioxidant potential of Ti nanoparticles fabricated using Zingiber officinale*, Environmental Research 236 (1) (2023) Article Number 116748 10 pages.

[12] L. T. D. Würger, J. Alarcan, A. Braeuning, *Effects of marine biotoxins on drug-metabolizing cytochrome P450 enzymes and their regulation in mammalian cells*, Archives of Toxicology, 98 (2024) 1311–1322.

[13] E. Gumbarewicz, A. Jarząb, A. Stepulak, W. Kukula-Koch, *Zingiber officinale Rosc. in the treatment of metabolic syndrome disorders-A review of in vivo studies*, International Journal of Molecular Sciences 23 (24) (2022) Article Number 15545 15 pages.

[14] P. Awasthi, A. Singh, G. Sheikh, V. Mahajan, A. P. Gupta, S. Gupta, Y. S. Bedi, S. G. Gandhi, *Mining and characterization of EST-SSR markers for Zingiber officinale Roscoe with transferability to other species of Zingiberaceae*, Physiology and Molecular Biology of Plants 23 (4) (2017) 925–931.

[15] Z. Chen, N. Tang, H. Li, G. Liu, L. Tang, *Genome-wide transcriptomic analysis during rhizome development of ginger (Zingiber officinale Roscoe.) reveals hormone and transcriptional regulation involved in cellulose production*, Scientia Horticulturae 264 (2020) Article Number 109154 11 pages.

[16] S. Han, X. Han, Y. Li, F. Guo, C. Qi, Y. Liu, S. Fang, J. Yin, Y. Zhu, *Genome-wide characterization and function analysis of ginger (Zingiber officinale Roscoe) ZoGRFs in responding to adverse stresses*, Plant Physiology and Biochemistry 207 (2024) Article Number 108392 13 pages.

[17] S. Han, X. Han, C. Qi, F. Guo, J. Yin, Y. Liu, Y. Zhu, *Genome-wide identification of DUF668 gene family and expression analysis under F. solani, chilling, and waterlogging stresses in Zingiber officinale*, International Journal of Molecular Sciences 25 (2) (2024) Article Number 929 15 pages.

[18] N. A. Ismail, M. Y. Rafii, T. M. M. Mahmud, M. M. Hanafi, G. Miah, *Molecular markers: A potential resource for ginger genetic diversity studies*, Molecular Biology Reports 43 (12) (2016) 1347–1358.

[19] D. Jiang, M. Xia, H. Xing, M. Gong, Y. Jiang, H. Liu, H. L. Li, *Exploring the heat shock transcription factor (HSF) gene family in ginger: A genome-wide investigation on evolution, expression profiling, and response to developmental and abiotic stresses*, Plants 12 (16) (2023) Article Number 2999 18 pages.

[20] Y. Jiang, D. Jiang, M. Xia, M. Gong, H. Li, H. Xing, X. Zhu, H. L. Li, *Genome-wide identification and expression analysis of the TCP gene family related to developmental and abiotic stress in ginger*, Plants 12 (19) (2023) Article Number 3389 21 pages.

[21] S. Tian, Y. Wan, D. Jiang, M. Gong, J. Lin, M. Xia, C. Shi, H. Xing, H. L. Li, *Genome-Wide Identification, Characterization, and Expression Analysis of GRAS Gene Family in Ginger (Zingiber officinale Roscoe)*, Genes 14 (1) (2023) Article Number 96 19 pages.

[22] H. Xing, Y. Jiang, Y. Zou, X. Long, X. Wu, Y. Ren, Y. Li, H. L. Li, *Genome-wide investigation of the AP2/ERF gene family in ginger: evolution and expression profiling during development and abiotic stresses*, BMC Plant Biology 21 (1) (2021) 1–21.

[23] H. Xing, Y. Li, Y. Ren, Y. Zhao, X. Wu, H. L. Li, *Genome-wide investigation of microRNAs and expression profiles during rhizome development in ginger (Zingiber officinale Roscoe)*, BMC Genomics 23 (1) (2022) 1–13.

[24] X. Yao, F. Meng, L. Wu, X. Guo, Z. Sun, W. Jiang, J. Zhang, J. Wu, S. Wang, Z. Wang, X. Su, X. Dai, C. Qu, S. Xing, *Genome-wide identification of R2R3-MYB family genes and gene response to stress in ginger*, Plant Genome 17(1) (2024) Article Number e20258 12 pages.

[25] P. Zhang, D. Liu, J. Ma, C. Sun, Z. Wang, Y. Zhu, X. Zhang, Y. Liu, *Genome-wide analysis and expression pattern of the ZoPP2C gene family in Zingiber officinale Roscoe*, BMC Genomics 25 (1) (2024) 1–20.

[26] C. L. M. Gilchrist, Y. H. Chooi, *Clinker & clustermap.js: Automatic generation of gene cluster comparison figures*, Bioinformatics 37 (16) (2021) 2473–2475.

[27] M. Tu, S. Cheng, W. Lu, M. Du, *Advancement and prospects of bioinformatics analysis for studying bioactive peptides from food-derived protein: Sequence, structure, and functions*, TrAC Trends in Analytical Chemistry 105 (2018) 7–17.

[28] B.R. Terlouw, K. Blin, J. C. Navarro-Muñoz, N. E. Avalon, M. G. Chevrette, S. Egbert, S. Lee, D. Meijer, M. J. J. Recchia, Z. L. Reitz, J. A. van Santen, N. Selem-Mojica, T. Tørring, L. Zaroubi, M. Alanjary, G. Aleti, C. Aguilar, S. A. A. Al-Salihi, H.E. Augustijn, J.A. Avelar-Rivas, L.A. Avitia-Domínguez, F. Barona-Gómez, J. Bernaldo-Agüero, V.A. Bielinski, F. Biermann, T.J. Booth, V. J. Carrion Bravo, R. Castelo-Branco, F. O. Chagas, P. Cruz-Morales, C. Du, K. R. Duncan, A. Gavriilidou, D. Gayrard, K. Gutiérrez-García, K. Haslinger, E. J. N. Helfrich, J. J. J. van der Hooft, A. P. Jati, E. Kalkreuter, N. Kalyvas, K. B. Kang, S. Kautsar, W. Kim, A. M. Kunjapur, Y. X. Li, G. M. Lin, C. Loureiro, J. J. R. Louwen, N. L. L. Louwen, G. Lund, J. Parra, B. Philmus, B. Pourmohsenin, L. J. U. Pronk, A. Rego, D. A. B. Rex, S. Robinson, L. R. Rosas-Becerra, E. T. Roxborough, M. A. Schorn, D. J. Scobie, K. S. Singh, N. Sokolova, X. Tang, D. Udwary, A. Vigneshwari, K. Vind, S. P. J. M. Vromans, V. Waschulin, S. E. Williams, J. M. Winter, T. E. Witte, H. Xie, D. Yang, J. Yu, M. Zdouc, Z. Zhong, J. Collemare, R. G. Linington, T. Weber, M. H. Medema, *MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters*, Nucleic Acids Research 51 (1) (2023) 603–610.

[29] T. Weber, H. U. Kim, *The secondary metabolite bioinformatics portal: Computational tools to facilitate synthetic biology of secondary metabolite production*, Synthetic and Systems Biotechnology 1 (2) (2016) 69–79.

[30] K. Blin, H. U. Kim, M. H. Medema, T. Weber, *Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters*, Briefings in Bioinformatics 20 (4) (2018) 1103–1113.

[31] S. A. Kautsar, H. G. Suarez Duran, K. Blin, A. Osbourn, M. H. Medema, *PlantiSMASH: Automated identification, annotation and expression analysis of plant biosynthetic gene clusters*, Nucleic Acids Research 45 (1) (2017) 55–63.

[32] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. Sternberg, *The Phyre2 web portal for protein modeling, prediction and analysis*, Nature Protocols 10 (6) (2016) 845–858.

[33] Z. Pang, J. Chen, T. Wang, C. Gao, Z. Li, L. Guo, J. Xu, Y. Cheng, *Linking Plant Secondary Metabolites and Plant Microbiomes: A Review*, Frontiers in Plant Science 12 (2021) Article Number 621276 22 pages.

[34] U. Öz, *Bioinformatics-based detection of secondary metabolite biosynthetic gene clusters and enzymes involved in terpene synthesis in Citrus sinensis (L.) Osbeck*, South African Journal Botany 168 (2024) 32–45.

[35] R. Bharadwaj, S. R. Kumar, A. Sharma, R. Sathishkumar, *Plant Metabolic Gene Clusters: Evolution, Organization, and Their Applications in Synthetic Biology*, Frontiers in Plant Science 12 (2021) 1–23.

[36] M. Wang, Q. Ji, B. Lai, Y. Liu, K. Mei, *Structure-function and engineering of plant UDP-glycosyltransferase*, Computational and Structural Biotechnology Journal 21 (2023) 5358–5371.

[37] R. Santhoshkumar, A. Yusuf, *In silico structural modeling and analysis of physicochemical properties of curcumin synthase (CURS1, CURS2, and CURS3) proteins of Curcuma longa*, Journal of Genetic Engineering and Biotechnology 18 (1) (2020) Article Number 24 9 pages.

[38] M. Urošević, L. Nikolić, I. Gajić, V. Nikolić, A. Dinić, V. Miljković, *Curcumin: Biological Activities and Modern Pharmaceutical Forms*, Antibiotics (Basel) 11 (2) (2022) Article Number 135 27 pages.

[39] K. Ashraf, A. Ahmad, A. Chaudhary, M. Mujeeb, S. Ahmad, M. Amir, N. Mallick, *Genetic diversity analysis of Zingiber officinale Roscoe by RAPD collected from subcontinent of India*, Saudi Journal of Biological Sciences 21 (2) (2014) 159–165.

[40] S. J. Smit, B. R. Lichman, *Plant biosynthetic gene clusters in the context of metabolic evolution*, Natural Product Reports 39 (7) (2022) 1465–1482.

[41] R. C. Rabara, C. Kudithipudi, M. P. Timko, *Identification of Terpene-Related Biosynthetic Gene Clusters in Tobacco through Computational-Based Genomic, Transcriptomic, and Metabolic Analyses*, Agronomy 13 (6) (2023) Article Number 1632 14 pages.

[42] S. Diao, Y. Zhang, Q. Luan, X. Ding, J. Sun, J. Jiang, *Identification of TPS-d subfamily genes and functional characterization of three monoterpene synthases in slash pine*, Industrial Crops and Products 188 (2022) Article Number 115609 11 pages.

[43] Y. Wang, Y. C. Liu, W. Y. Li, K. Guo, Y. Liu, S. H. Li, *Antifeedant, cytotoxic, and anti-inflammatory neo-clerodane diterpenoids in the peltate glandular trichomes and fresh leaves of Ajuga forrestii*, Phytochemistry 186 (2021) Article Number 112731 9 pages.

[44] T. Qiu, Y. Y. Li, H. Wu, H. Yang, Z. Peng, Z. Du, Q. Wu, H. Wang, Y. Shen, L. Huang, *Tandem duplication and sub-functionalization of clerodane diterpene synthase originate the blooming of clerodane diterpenoids in Scutellaria barbata*, The Plant Journal 116 (2) (2023) 375–388.

[45] Q. Wei, K. Lan, Y. Liu, R. Chen, T. Hu, S. Zhao, X. Yin, T. Xie, *Transcriptome analysis reveals regulation mechanism of methyl jasmonate-induced terpenes biosynthesis in Curcuma wenyujin*, PLoS One 17 (6) (2022) Article Number e0270309 20 pages.

[46] G. Nie, F. Wu, Z. Duan, S. Wang, B. Ao, P. Zhou, J. Zhang, *Genome-wide analysis of the cytochrome P450 superfamily suggests its roles in coumarin biosynthesis and salt stress response in Melilotus albus*, Environmental and Experimental Botany 220 (2024) Article Number 105718.

[47] N. Radadiya, N. Mangukia, V. Antala, H. Desai, H. Chaudhari, T. L. Dholaria, D. Dholaria, R. S. Tomar, B. A. Golakiya, M. K. Mahatma, *Transcriptome analysis of sesame-Macrophomina phaseolina interactions revealing the distinct genetic components for early defense responses*, Physiology and Molecular Biology of Plants 27 (8) (2021) 1675–1693.

[48] R. R. Al-Hindi, A. Bahieldin, S. Edris, M. G. Alharbi, A. A. Abulfaraj, *Expression Profiling of Transcription Factors Under Drought Stress At Seedling Stage in Bread Wheat (Triticum aestivum L.)*, Applied Ecology and Environmental Research 21 (5) (2023) 3965–3996.

[49] A. Sicilia, G. Testa, D. F. Santoro, S. L. Cosentino, A. R. Lo Piero, *RNASeq analysis of giant cane reveals the leaf transcriptome dynamics under long-term salt stress*, BMC Plant Biology 19 (1) (2019) 1–24.

[50] Y. Wang, X. Li, N. Liu, S. Wei, J. Wang, F. Qin, B. Suo, *The iTRAQ-based chloroplast proteomic analysis of Triticum aestivum L. leaves subjected to drought stress and 5-aminolevulinic acid alleviation reveals several proteins involved in the protection of photosynthesis*, BMC Plant Biology 20 (1) (2020) Article Number 96 17 pages.

[51] C. Gharsallah, S. Gharsallah Chouchane, S. Werghi, M. Mehrez, H. Fakhfakh, F. Gorsane, *Tomato contrasting genotypes responses under combined salinity and viral stresses*, Physiology and Molecular Biology of Plants 26 (7) (2020) 1411–1424.

[52] M. Salami, B. Heidari, J. Batley, J. Wang, X. L. Tan, C. Richards, H. Tan, *Integration of genome-wide association studies, metabolomics, and transcriptomics reveals phenolic acid- and flavonoid-associated genes and their regulatory elements under drought stress in rapeseed flowers*, Frontiers in Plant Science 14 (January) (2023) 1–25.

[53] U. Karki, P. Perez Sanchez, S. Chakraborty, B. Dickey, J. Vargas Ulloa, N. Zhang, J. Xu, *Intracellular trafficking and glycosylation of hydroxyproline-O-glycosylation module in tobacco BY-2 cells is*

dependent on medium composition and transcriptome analysis, Scientific Reports 13 (1) (2023) Article Number 13506 15 pages.

[54] J. Ma, Y. Geng, W. Pei, M. Wu, X. Li, G. Liu, D. Li, Q. Ma, X. S. Zang, S. Yu, J. Zhang, J. Yu, *Genetic variation of dynamic fiber elongation and developmental quantitative trait locus mapping of fiber length in upland cotton (Gossypium hirsutum L.)*, BMC Genomics 19 (2018) Article Number 882 18 pages.

[55] S. Wei, Z. Song, S. Luo, Y. Zhong, Y. Zhou, R. Lu, *Transcriptome Analysis Reveals the Heat Stress Response Genes by Fire Stimulation in Michelia macclurei Dandy*, Forests 14 (3) (2023) Article Number 610 19 pages.

[56] H. Wang, Z. Wang, M. Zhang, B. Jia, W. Heng, Z. Ye, L. Zhu, X. Xu, *Transcriptome sequencing analysis of two different genotypes of Asian pear reveals potential drought stress genes*, Tree Genetics & Genomes 14 (3) (2018) Article Number 40 14 pages.

[57] M. Huang, H. Xing, Z. Li, H. Li, L. Wu, Y. Jiang, *Identification and expression profile of the soil moisture and Ralstonia solanacearum response CYPome in ginger (Zingiber officinale)*, PeerJ 9 (2021) Article Number e11755 22 pages.

[58] J. Domanski, M. S. P. Sansom, P. J. Stansfeld, R. B. Best, *Atomistic mechanism of transmembrane helix association*, PLOS Computational Biology 16 (6) (2020) Article Number e1007919 14 pages.

[59] M. Bernhofer, B. Rost, *TMbed: transmembrane proteins predicted through language model embeddings*, BMC Bioinformatics 23 (1) (2022) Article Number 326 19 pages.

Journal of New Results in Science

https://dergipark.org.tr/en/pub/jnrs

Research Article

Open Access

E-ISSN:1304-7981 https://doi.org/10.54187/jnrs.1527152

An investigation of the enzymatic oligomerization of nitro-substituted phenylene diamine: Thermal and fluorescence properties

Feyza Kolcu[1]

Keywords:
p-phenylenediamine, Horseradish peroxidase, Enzyme-catalyzed oligomerization, Green photoluminescence emission, Limiting oxygen index

Abstract — 2-nitro-*p*-phenylenediamine, an aromatic diamine, was studied for its oxidative oligomerization with H_2O_2 using enzyme-catalyzed oligomer synthesis. Characterization of molecular structures was performed utilizing ultraviolet-visible (UV-Vis) spectrophotometer, Fourier-transform infrared spectroscopy (FT-IR), and nuclear magnetic resonance (NMR) techniques, identifying phenazinebridged oligomer resulting from the enzymatic oligomerization process. Based on the results of gel permeation chromatography (GPC) analysis, the synthesized compound was identified as being in an oligomeric form. Conversely, the number of repeating units, as determined by weight average molecular weight (M_w), was found to be 28. The solvent effect on the optical features of the synthesized oligomer in polar solvents was analyzed. The degradation of phenazine-type structures in the oligomer occurred at higher temperatures than that of the monomer. Under visible light excitation, the oligomer exhibited green light emission with a quantum yield (QY) of 6.2% in *N*,*N*-dimethylformamide (DMF). 2-nitro-*p*-phenylenediamine was readily oxidized into an oligomer with *ortho*-coupled constitutional units, having a lower electrochemical band gap than the monomer, via the enzymatic oligomerization route. Scanning electron microscopy revealed that enzyme-catalyzed oxidation of monomers exhibited a spongy morphology with some pores.

Subject Classification (2020): 80A50, 82D60

1. Introduction

π-conjugated polymers, particularly polyaniline and its derivatives have garnered significant interest in developing advanced materials. Among these, polyphenylenediamines are notable as important conducting polymers due to their facile synthesis. Aromatic diamine polymers offer unique functionalities compared to polyanilines, making them promising for applications in electrochromism and electrochromic cells [1,2]. The aromatic units within the conjugated polymer main chains, structured by an sp^2 carbon framework, facilitate charge transport along the polymer backbones [3,4]. Therefore, synthesizing novel conjugated polymers is crucial for technological applications [5,6]. Conducting polymers like polyaniline (PANI) and its derivatives are highly regarded for their well-defined structure and distinctive electrical and optical features [3]. PANI and poly(phenylenediamine) have seen increasing interest as polymeric adsorbents for heavy metals and various dyes due to their environmental stability [7].

Aromatic diamines are highly susceptible to oxidative oligomerization/polymerization, which can involve the oxidation of one or both -NH_2 units, forming ladder poly(phenazine) or poly(aminoaniline). These polymers exhibit distinct features compared to other conducting polymers like polyaniline. Poly (*p*-aminoaniline) is only

[1]feyzakolcu@comu.edu.tr (Corresponding Author)
[1]Department of Chemistry and Chemical Processing Technologies, Çanakkale Onsekiz Mart University, Çanakkale, Türkiye
Article History: Received: 02 Aug 2024 — Accepted: 30 Aug 2024 — Published: 31 Aug 2024

slightly soluble in most organic solvents, and its molecular weight is typically in the range of thousands, which is more characteristic of oligomers rather than polymers. These polymers can improve solubility by introducing polar side groups like NO_2 substituents [8].

From the perspective of oligomer/polymer mechanisms, enzymatic polymer synthesis is a crucial aspect of green polymer chemistry [9]. Horseradish peroxidase (HRP) facilitates the decomposition of H_2O_2 and catalyzes the oxidative oligomerization/polymerization of aromatic amine derivatives [10]. Enzymes are highly effective catalysts for synthesizing macromolecules, operating under mild reaction conditions. Amine functional groups act as initiators for various crosslinking reactions, essential in forming plastic thermosets, adhesives, and biologically active substances [11].

This study researched synthesizing and characterizing the enzymatically oligomerized 2-nitro-*p*-phenylenediamine (monomer) compound. The study provides insights into designing structures to improve the solubility of nitro-bearing oligomers. The synthesized oligomer exhibited green light emission upon visible light excitation. In the enzyme-catalyzed oxidative oligomerization of 2-nitro-*p*-phenylenediamine, the oxidation of *p*-phenylenediamine produces cationic radicals, leading to the formation of *ortho*-coupled or head-to-tail units via N-C linkage formation, culminating in oligomer. Konyushenko et al. suggested that phenazine units play a part in gathering polyaniline-related nanostructures [12]. The study aimed to elucidate the feasibility of an enzyme-catalyzed oligomerization mechanism for 2-nitro-*p*-phenylenediamine. Comparative analyses were conducted using photoluminescence (PL), thermogravimetric analysis (TGA), ultraviolet-visible (UV-Vis) spectrophotometer, differential scanning calorimetry (DSC), and cyclic voltammograms (CV) techniques to evaluate the oligomer's photoluminescence, thermal, and electrochemical features.

2. Materials and Methods

2.1. Materials

Horseradish Peroxidase (HRP) and 2-nitro-*p*-phenylenediamine (monomer) were procured from Sigma-Aldrich Chemical Co. and Across Organics, respectively. Additionally, Sigma-Aldrich Chemical Co. provided ethanol (EtOH), methanol (MetOH), dimethyl sulfoxide (DMSO), acetonitrile (CH_3CN), dichloromethane (CH_2Cl_2), tetrahydrofuran (THF), heptane, hexane, 1,4-dioxane, *N*,*N*-dimethylformamide (DMF), Tetrabutylammonium hexafluorophosphate (NBu_4PF_6), pH=7 disodium hydrogen phosphate, and 35% hydrogen peroxide (H_2O_2). No further purification was needed for the chemicals used as received.

2.2. Enzymatic oligomerization of 2-nitro-*p*-phenylenediamine

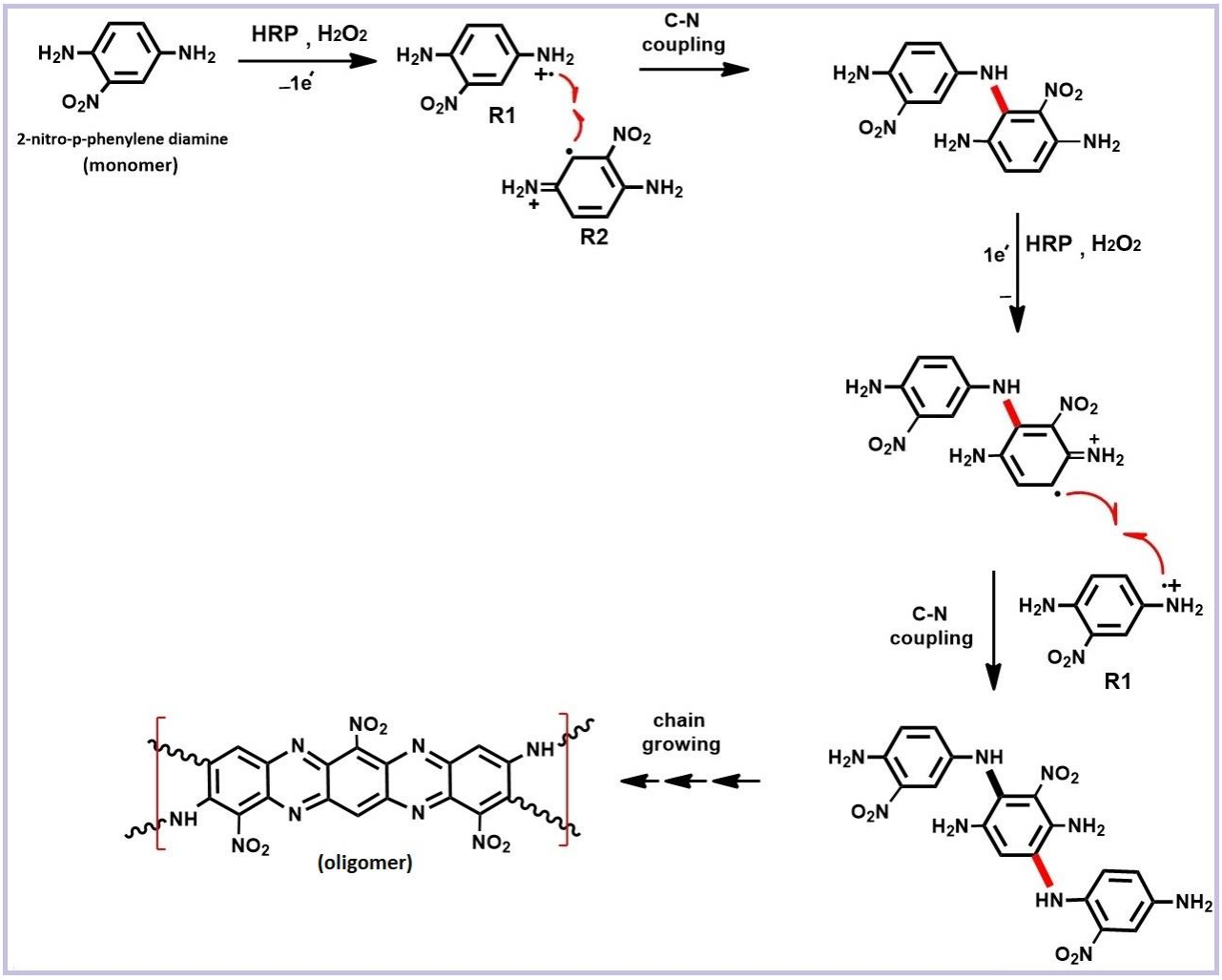
A mixture of monomer (2.0 mmol, 0.306 g) and 5 mg of HRP was prepared in 30 mL of 1,4-dioxane, comprising 0.1 M disodiumhydrogen phosphate buffer solution (10 mL). The enzymatic oligomerization was commenced by putting in 230 µL of H_2O_2. Subsequently, 1 mL of H_2O_2 was added in 1 hour at room temperature. After approximately 24 hours, a dark brown precipitate formed at the bottom of the flask, indicating a successful oligomerization at room temperature. The resulting product was filtered to remove unreacted reactants, which were washed away with 100 mL of distilled H_2O three times. Drying the oligomer was performed in a vacuum oven at 45 °C, yielding 61%. The suggested mechanism for the oligomer formation is outlined in Scheme 1.

2.3. Instruments

The solubility test was conducted with 1 mg of oligomer, thoroughly dispersed in 1 mL of solvent. The functional groups of the monomer and oligomer were analyzed using a PerkinElmer FT-IR Spectrum One, including an ATR accessory, covering the range 4000-400 cm^{-1}. With DMSO-d_6 and tetramethylsilane (TMS), ^1H and ^{13}CNMR spectra were obtained using an Agilent 600 MHz and 150 MHz Premium COMPACT NMR Magnet, respectively. The molecular weight distribution of the oligomer was specified using Gel Permeation

Chromatography (GPC) with Light Scattering and Refractive Index detectors (Malvern Viscotek GPC Dual 270 max). The GPC column was calibrated at 55 ºC, using polystyrene standards (162-60.0000 Dalton (Da), Polymer Laboratories).

DMF containing 40 mM lithium bromide (LiBr) was used as the mobile phase. The optical features of the monomer and its oligomer were assessed using an Analytikjena Specord 210 spectrometer in the wavelength range of 280-700 nm. The measurements were conducted using sample solutions placed in a 1 cm quartz cell. Photoluminescence (PL) measurements were performed by a Shimadzu RF-5301PC spectrofluorophotometer (slit width: 5 nm). Thermogravimetric-Differential Thermal Analysis (TG-DTA) was performed using Perkin Elmer Diamond Thermal Analyzer from 25 °C to 1000 °C. Differential Scanning Calorimetry (DSC) analysis was performed using a Perkin Elmer Sapphire Differential Scanning Calorimeter between 25 °C and 420 °C (N_2 atmosphere and heating rate: 10 °C min^{-1}). A CHI 660C Electrochemical Analyzer was used to monitor CV with a 25 mV s^{-1} scan rate. The measurements under argon employed a cell consisting of a silver and a platinum wire as a reference and auxiliary electrode, respectively, as well as a GCE as a working electrode, in a 0.1 M NBu_4PF_6 in CH_3CN solution as the supporting electrolyte. The band gap ($E_g^{'}$) between the electrochemical highest occupied molecular orbital - lowest unoccupied molecular orbital (HOMO-LUMO) levels and were obtained from the cyclic voltammograms [13]. Scanning Electron Microscopy (SEM) photographs were received to analyze the morphology of the oligomeric particles.



Scheme 1. Synthesis procedure for the oligomer

3. Results and Discussion

3.1. Spectral Comments on Monomer and Oligomer

2-nitro-*p*-phenylenediamine was enzymatically oligomerized, as shown in Scheme 1. The yield of brown-colored oligomer was 51%. Fine solubility of the oligomer in MetOH, EtOH, DMSO, DMF, CH_2Cl_2, and CH_3CN was observed, whereas it was insoluble in apolar solvents, such as heptane and hexane.

The coupling of R1 and R2 radicals resulted in the formation of the enzyme-catalyzed oligomer. Two monomer radicals were coupled to initiate the chain extension of the oligomer. R1 and R2 have the potential to interact, forming a concurrent double C-N-C bond because of intramolecular oxidative cyclization, leading to the formation of phenazine units in a ladder-type oligomer (Scheme 1) [14].

Figure 1 presents the FT-IR spectra of the monomer and its enzymatically oligomerized product. The N-H stretching wavenumbers observed at 3440 and 3326 cm^{-1}, corresponding to the primary amine groups in monomer, were detected at 3466 and 3366 cm^{-1} for oligomer. The peaks at 3455 and 3359 cm^{-1}, associated with the N-H stretching modes, indicate the presence of NH-/-NH_2 in the oligomer. The C-H stretching modes of monomer and oligomer were displayed at 3170 and 3179 cm^{-1}, respectively. The N-H bending vibration of primary amine in the monomer was observed at 1591 cm^{-1}. The C=C stretching of the monomer was noticed within the range of 1568 to 1513 cm^{-1}. The stretching mode of C=N in the phenazine unit showed itself as an appearing peak at 1630 cm^{-1} in the oligomer's spectrum [14]. The asymmetrical and symmetric N-O stretching bands were monitored at 1475 and 1334 cm^{-1} and 1475 and 1331 cm^{-1}, respectively. Additionally, bands at 820 at 826 cm^{-1} are related to ring hydrogen deformation vibrations.

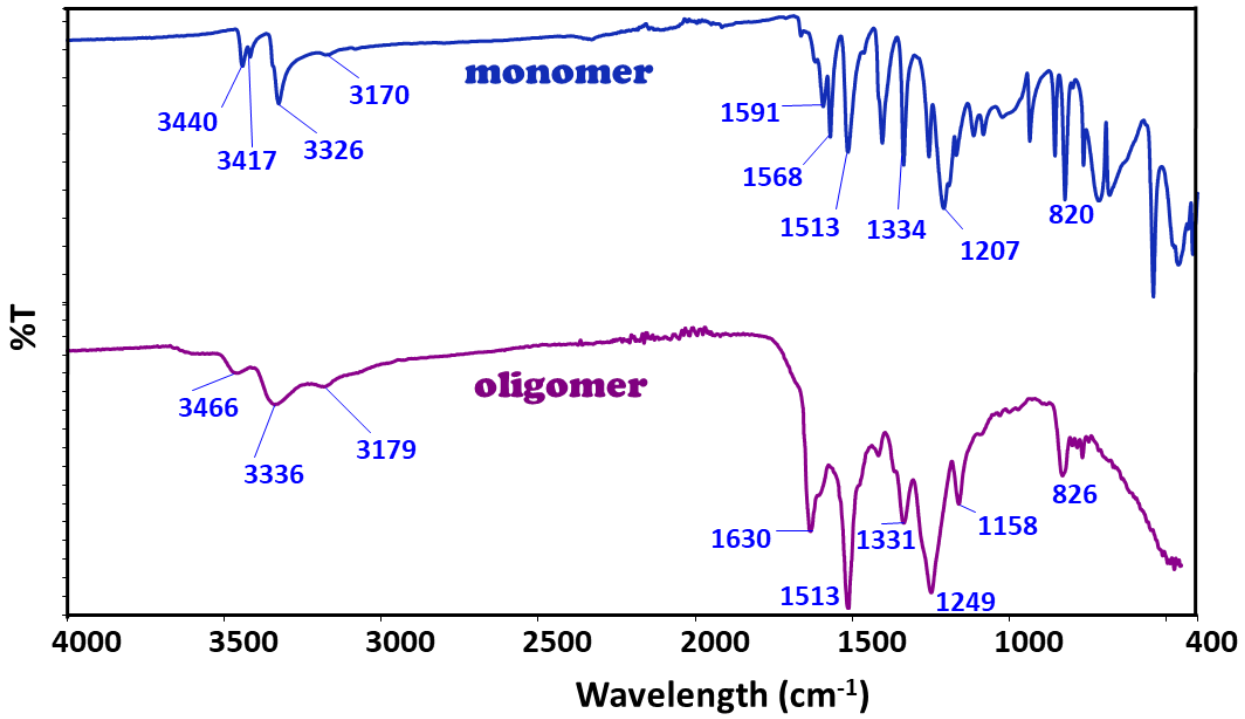


Figure 1. FT-IR spectra of monomer and oligomer

^{1}H NMR spectra of the monomer and its oligomer were analyzed to elucidate their molecular structures, as shown in Figure 2. For the monomer, the peaks at 4.80 (singlet), 6.81 (doublet), 6.86 (doublet), and 7.11 (singlet) ppm were related to three aromatic hydrogen protons and four amino hydrogens, respectively, as seen in Figure 2A. The *ortho*-coupled monomer units underwent intramolecular cyclization, forming a phenazine-type structure. The transformation of primary to secondary amine via N-C coupling is evidenced by the appearance of a singlet at 7.88 ppm, as depicted in Figure 2B. The terminal -NH_2 groups were ascribed to the peak at 6.94 ppm.

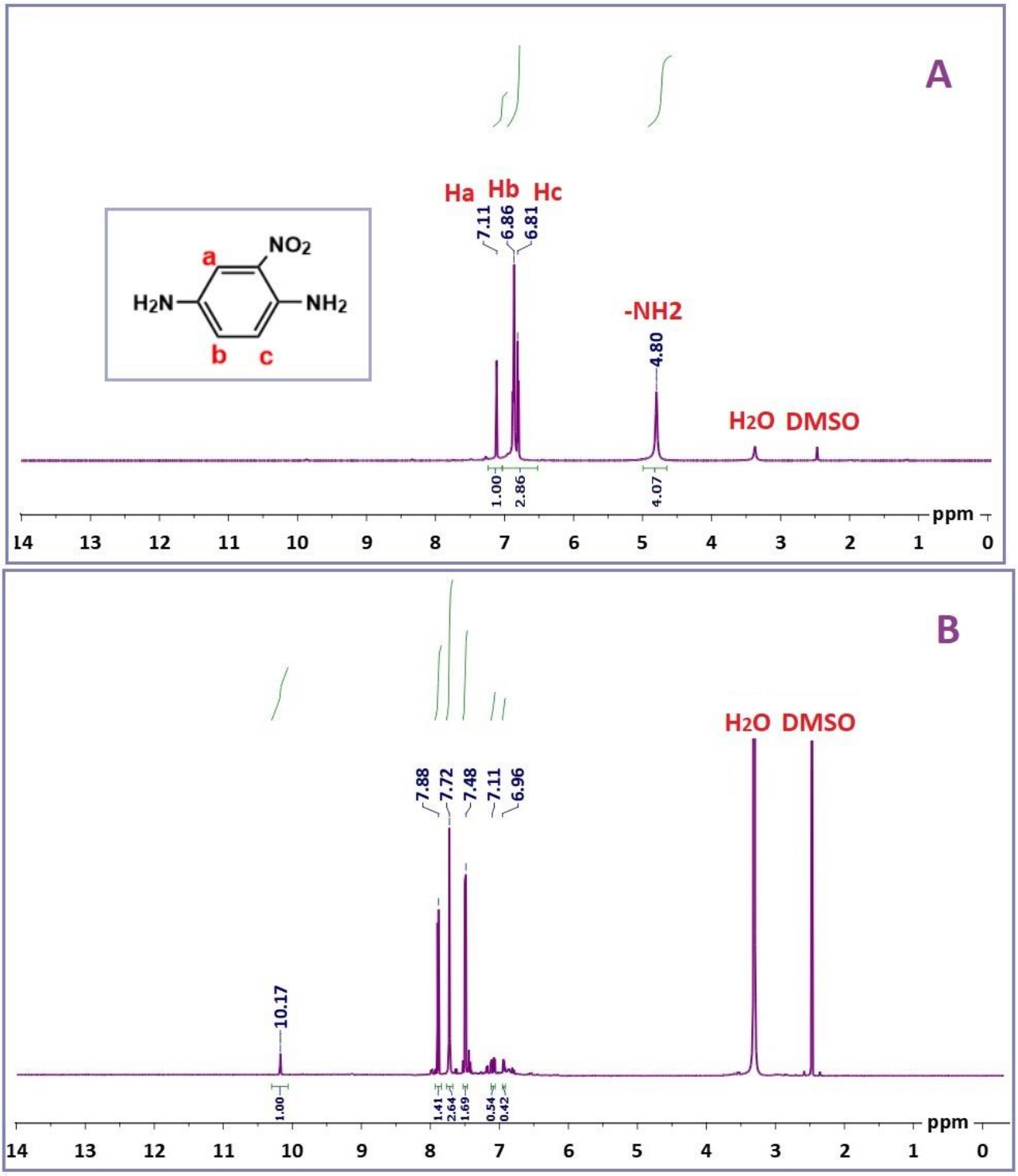


Figure 2. [1]H NMR spectra of (A) monomer and (B) oligomer

[13]C NMR spectrum of the monomer indicated the signals for C1, C3, and C4, which carry -NH_2, -NO_2, and -NH_2, appearing at downfield shifts of 139.65, 138.98, and 130.67 ppm, respectively, as seen in Figure 3A. Due to the +M effect of NH_2 groups, the carbon signals of C6, C5, and C2 at the *ortho* position shifted upfield at 127.53, 120.33, and 105.72 ppm, respectively. According to Figure 3B, two distinct peaks at 160.02 and 150.62 ppm could be evidence of two C=N bonds due to the phenazine ring formation in the enzyme-catalyzed oligomerization.

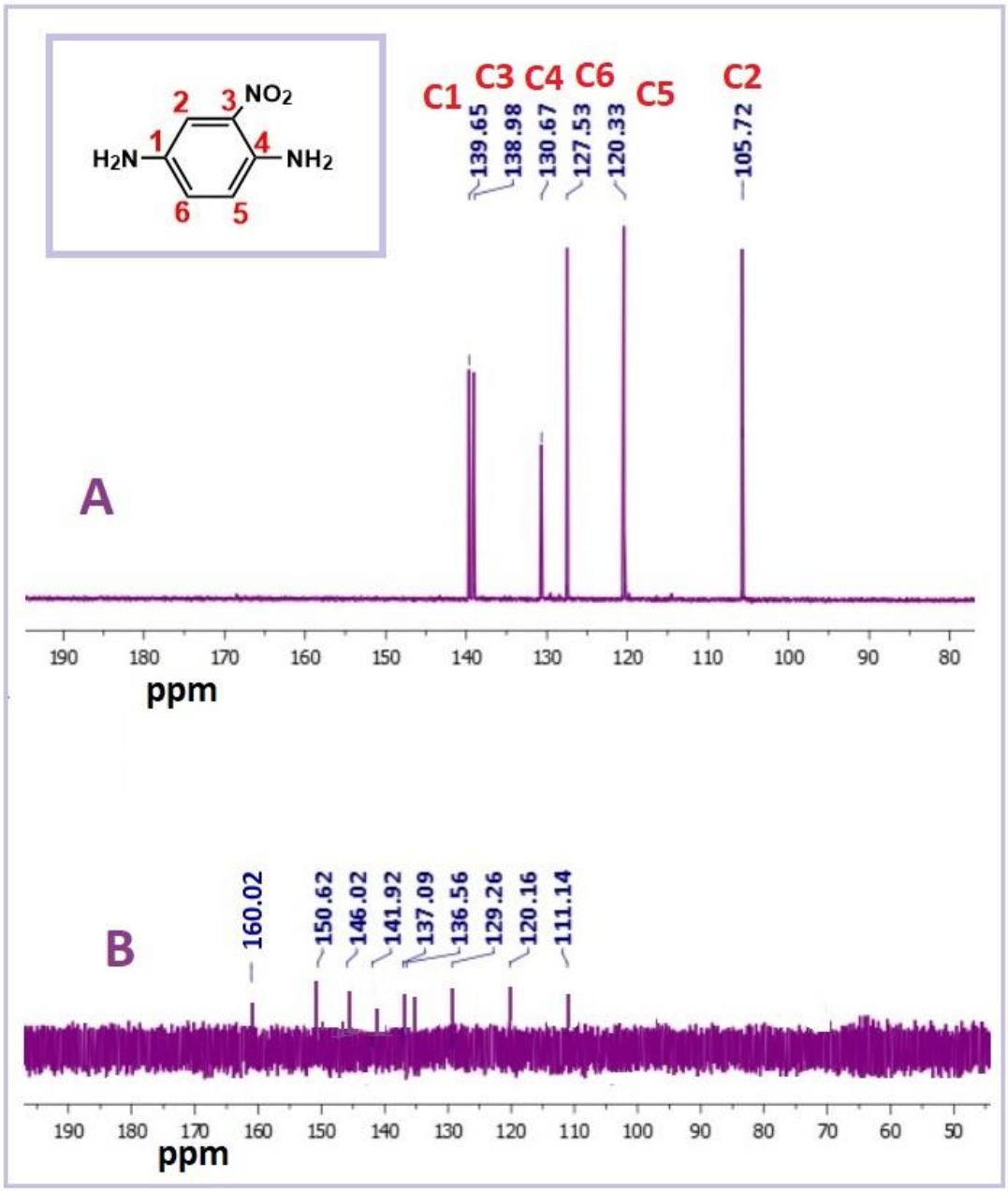


Figure 3. ^{13}C NMR spectra of (A) monomer and (B) oligomer

Based on the GPC analysis of the oligomer, weight average molecular weight (M_w) and number average molecular weight (M_n) were found to be 4350 and 3800 Da, respectively, confirming that the synthesized compound possesses an oligomeric structure. PDI value, calculated as the ratio of M_w to M_n, was determined as 1.14. Based on M_w and M_n, the average number of repeating units was 28 and 25, respectively.

3.2. Electrochemical and Optical properties

Due to phenazine units appearing in the enzymatic oligomerization, the π→π* electronic transition for the oligomer was observed between 280 and 348 nm. The synthesized oligomer's UV-vis properties were analyzed and presented in Figure 4 to investigate the solvent effect. A broad signal, ascribed to the n→π* electronic transition of NH_2 and NO_2 groups, was detected in the 400-600 nm range. The π→π* electronic transition in the oligomer remained unaffected by changing the solvent polarity. The maximum absorption wavelength (λ_{max}) values of the oligomer were observed to be 474 nm, 476 nm, 482 nm, 483 nm, 488 nm, and 493 nm in CH_2Cl_2, CH_3CN, MetOH, EtOH, DMF, and DMSO, respectively, indicating a bathochromic shift of n→π* band as the solvent polarity increased, as seen in Figure 4.

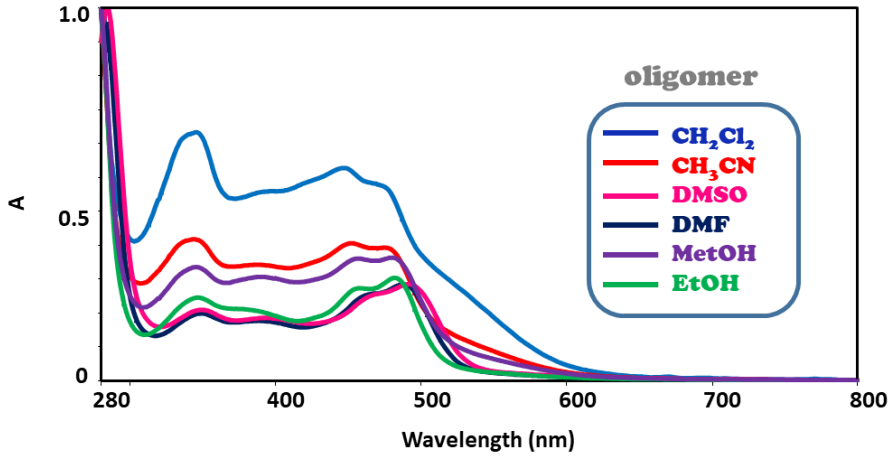


Figure 4. Normalized UV-vis spectra of enzyme-catalyzed oligomer in solvents.

Protic solvent molecules orient themselves around the chromophore of the oligomer, forming hydrogen bonds with the unshared electron pairs on the oxygen and nitrogen atoms. Upon establishing the intermolecular hydrogen bonding, the energy gap between the HOMO and the LUMO levels decreases with increasing solvent polarity, leading to a shift of (λ_{max}) towards longer wavelengths. The π→π* transition in the aromatic moiety of the oligomer remains unaffected by the solvent.

The peaks for oxidation and reduction potentials of the compounds are determined using Cyclic Voltammograms, as presented in Figure 5. During oxidation, one electron is removed from the HOMO level. A radical cation, generated through CV oxidation, subsequently underwent a radical-radical coupling related to deprotonation, forming a dimer [15]. The anodic peaks can be ascribed to the oxidation of -NH_2 of monomer and oligomer, forming a polaron structure ($^{\bullet+}$NH) through coupling with the parent molecule. Following the formation of the oligomer, the terminal amino groups undergo oxidation due to their lower oxidation potentials.

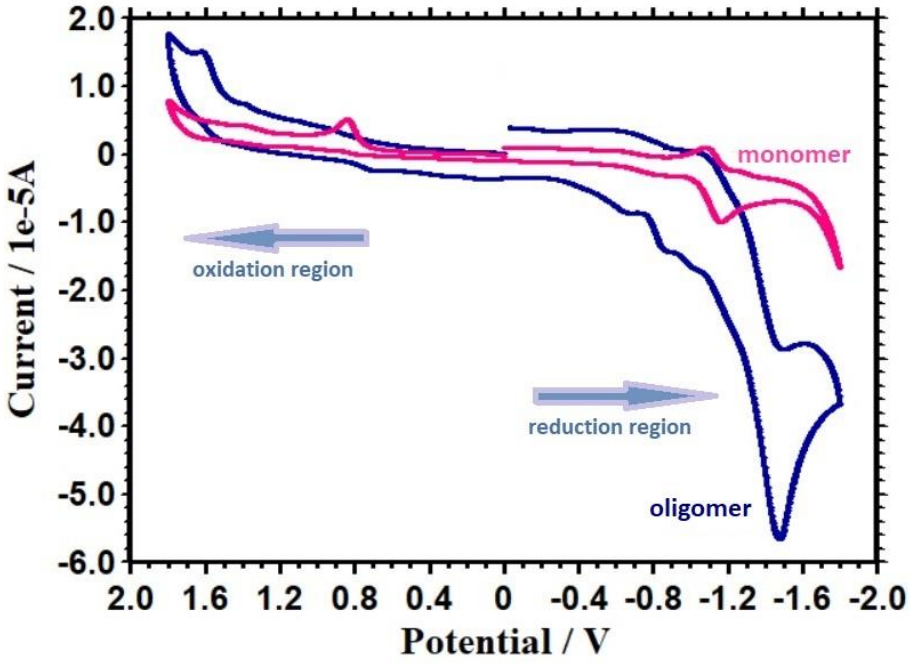


Figure 5. CVs of monomer and oligomer

The -C=N- group exhibits high sensitivity to reduction in the synthesized oligomer. During the negative scan, the -C=N- group undergoes reduction at -1471 mV. The elevation of HOMO from -5.58 V to -5.87 V and the reduction of the LUMO from 2.97 V to 3.01 V during the oligomer formation was observed. The enzyme-catalyzed oligomer had a lower $E_g^{'}$ values than monomer.

3.3. Fluorescence (PL) Characteristics

The PL spectrum of the obtained oligomer was recorded to study its photoluminescence property. Green light emission was observed in DMF for the enzyme-catalyzed oligomer, as presented in Figure 6. The maximum emission wavelength (λ_{max}) occurred at 507 nm, corresponding to the emission of green light when excited by 410 nm. Fluorescent light emission intensity was measured at 410 a.u. PL characteristics could be attributed to quantum yield (QY) [16]. The oligomer synthesized enzymatically exhibited QY of 6.2% for green photoluminescence emission. Notably, the enzyme-catalyzed oligomer, synthesized using an environmentally friendly method, appears promising for future development in optoelectronic devices. The inset displays the time-dependent (0-3600 s) fluorescence measurement under 410 nm light excitation. No changes in fluorescence intensity were observed over 3600 s under identical conditions, indicating that the enzymatically synthesized oligomer exhibits excellent photostability when excited by 410 nm light. Consequently, the synthesized oligomer shows promising potential for producing display devices and light-emitting diodes.

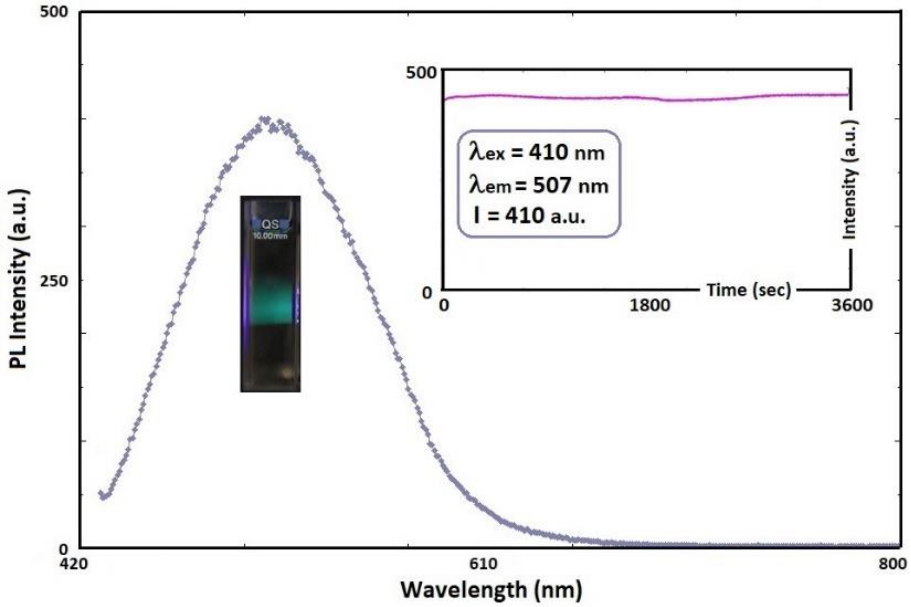


Figure 6. PL emission spectrum of oligomer Inset: Time-dependent PL spectrum of oligomer

3.4. Thermal Stability

The thermal properties of the monomer and the synthesized oligomer were investigated using TGA-DTG instruments from 25 °C to 1000 °C, as presented in Figure 7. The thermal degradation of the monomer and oligomer occurred in two steps, as noticed in Figure 7. Although solvent evaporation was not observed in the monomer, it was evident in the synthesized oligomer as an initial weight loss of 9.30% from 30 °C to 120 °C, attributed to the moisture trapped within the oligomer chains. The degradation (T_{on}) of the monomer and oligomer started at 231 and 310 °C, respectively. The oligomer's higher T_{on} value than the monomers can be explained by the conjugation resulting from the linkage of monomer units via phenazine formation, as depicted in Scheme 1.

Consequently, the phenazine-type structures' degradation within the oligomer occurs at higher temperatures than the monomer (Figure 7A). The maximum weight loss (T_{max}) occurred for the monomer and oligomer were 266 °C and 796 °C, and 329 °C and 751 °C, respectively, as observed in Figure 7B. Furthermore, the temperatures corresponding to 20% weight loss (T_{20}) and 50% weight loss (T_{50}) were found at 259 °C and 331 °C for the monomer and 341 °C and 728 °C for the oligomer, respectively.

The monomer decomposition occurred between 150 °C and 417 °C, resulting in a 52.18% weight loss, and continued to 1000 °C with an additional 18.40% weight loss. For the oligomer, two decomposition stages were

observed: the first ranged from 110 °C to 377 °C with a 31.14% weight loss, and the second stage ranged from 377 °C to 1000 °C with a 36.82% weight loss. The char% was calculated to be 28.93% for the monomer and 33.04% for the oligomer. This residual content is attributed to the decomposition of the oligomer's rigid chain structure, organized into a structural order stabilized by phenazine units. T_{max} values indicate that the thermal steadiness of the synthesized oligomer is superior to that of the monomer. Therefore, the synthesized oligomer demonstrates the potential for applications across diverse industrial sectors, including electrical applications necessitating high-temperature insulation in manufacturing aircraft components, weapon systems, and space vehicles [11].

The flame retardancy value of the oligomer can be obtained using the TGA curve following Van Krevelen's calculation, utilizing the percentage of residue at 1000 °C. The formation of carbon during thermal decomposition restricts the emission of organic volatile compounds. This carbonaceous layer, produced in the process, impedes the propagation of flammable gases by diminishing heat transfer to the material. In this regard, Van Krevelen proposed a formula of LOI = 17.5 + 0.4 (*s*) where LOI represents the limiting oxygen index and (*s*) denotes the percentage of residue in the polymer, respectively [17]. According to Van Krevelen's theory, an increase in residue formation correlates with the flammability of polymers. For polymers to be self-extinguishing, their LOI values must be at least 26 or higher [18]. The LOI value of the synthesized oligomer was calculated as 30.72, indicating that the oligomer exhibits excellent thermal stability and lower flammability. From DSC measurements, the glass transition temperature (T_g) and specific heat capacity (ΔC_p) values of the oligomer were calculated to be 208 °C and 0.384 J/g °C, respectively. An exothermic peak at 340 °C was observed in the DSC curve of the oligomer.

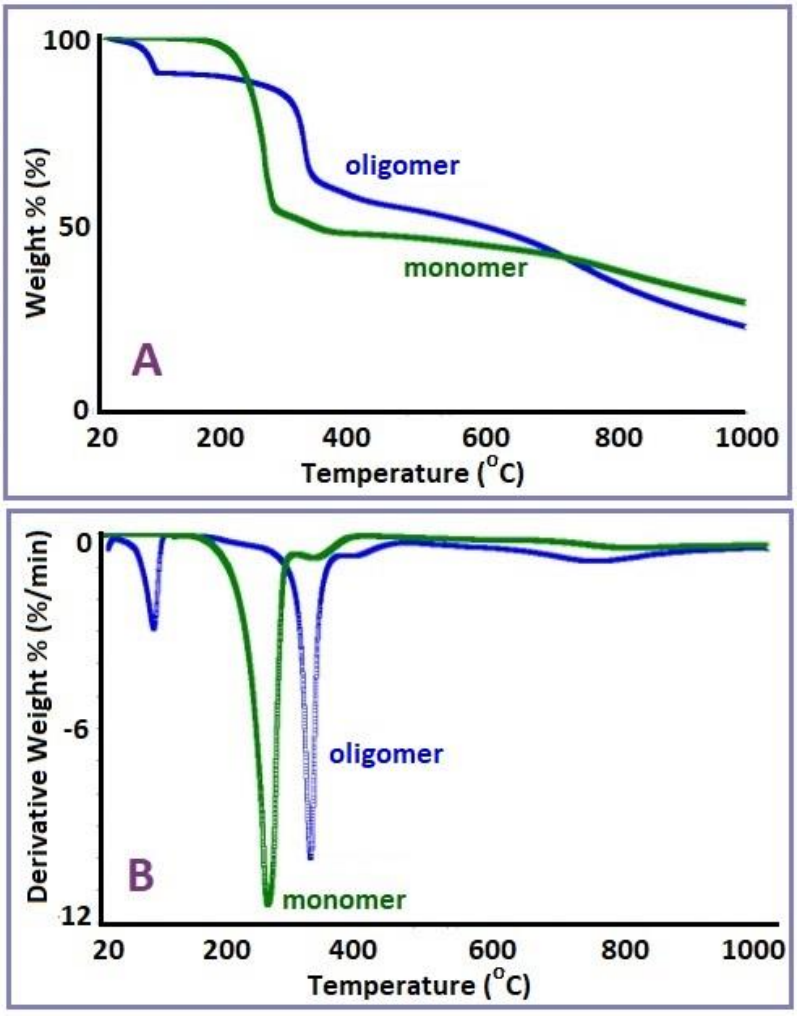


Figure 7. (A) TGA and (B) DTG curves of monomer and its oligomer

3.5. Morphological Properties

The sample was sprinkled onto the stub and subsequently subjected to a sputter coating with a 5 nm layer of gold. Morphological properties of the enzyme-catalyzed oligomer were examined using SEM at different magnifications, as illustrated in Figure 8. As depicted in Figure 8, the surface morphology of the oligomer exhibited a sponge-like structure with an irregular network. The accumulated oligomer particles formed a dense, shrunken mass with a rough surface.

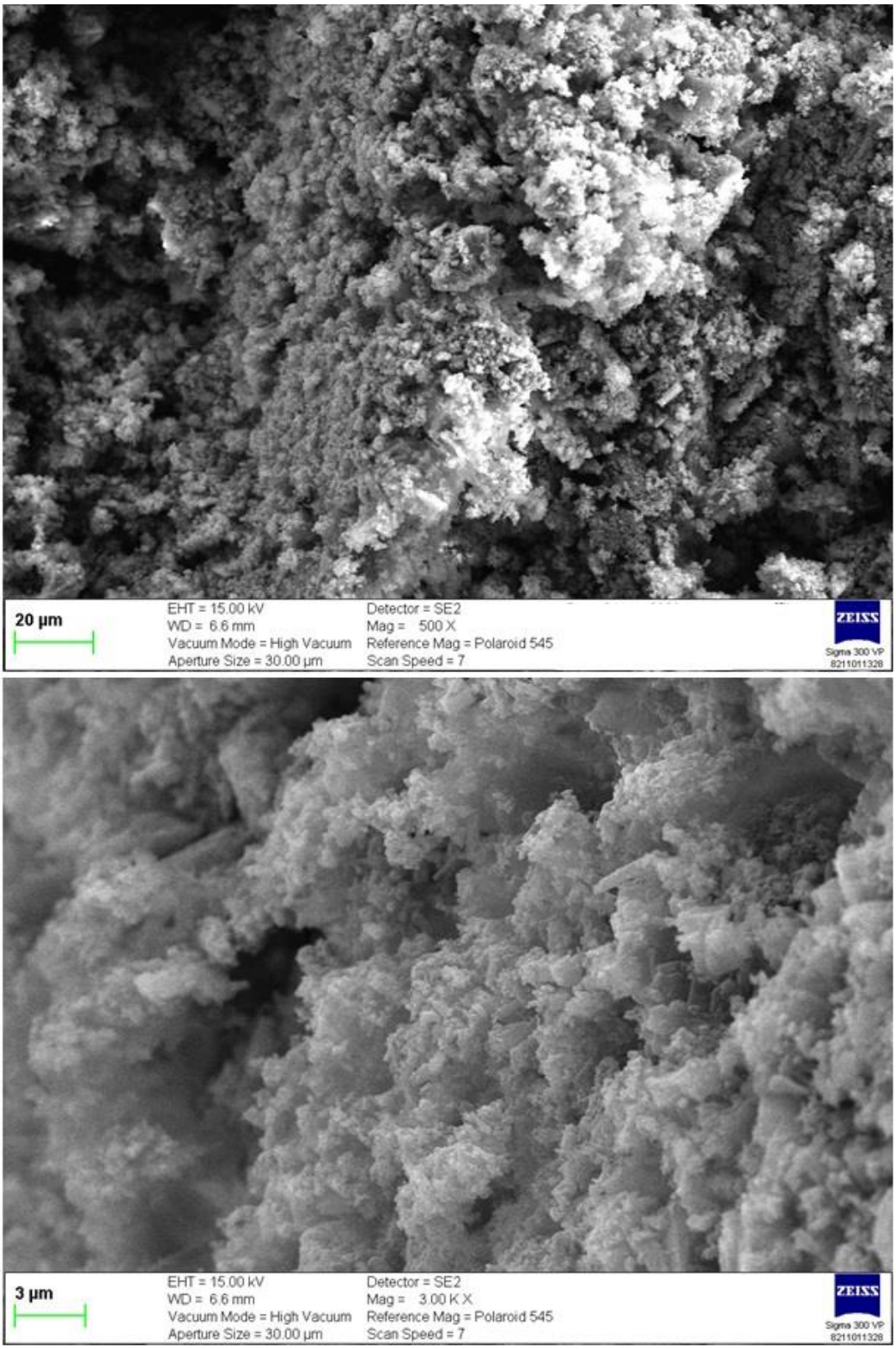


Figure 8. SEM views of the oligomer

4. Conclusions

The oxidation of 2-nitro-*p*-phenylenediamine, an aromatic diamine compound, resulted in an enzyme-catalyzed oligomer. The solvent's polarity significantly influenced the synthesized oligomer's UV-vis absorption. Due to its potential as a green light emitter, the enzyme-catalyzed oligomer can be a promising compound for the fabrication of display devices and applications in bioscience. Additionally, the exceptional thermal stability of the oligomer suggested its suitability for use in the manufacture of electrical appliances requiring high-temperature insulation. These attributes highlighted the advantages of the enzymatic synthetic process, which provides simplicity in synthesis, environmental compatibility, chemical stability, and notable thermal and photophysical features. Obtaining and characterizing macromolecules using different catalysts and oxidative polymerization methods for the monomer used in this study will shed light on future studies.

Author Contributions

The author read and approved the final version of the paper.

Conflict of Interest

The author declares no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] T. Qin, L. Deng, P. Zhang, M. Tang, C. Li, H. Xie, S. Huang, X. Gao, *Enhancement of Electrochromic Properties of Polyaniline Induced by Copper Ions*, Nanoscale Research Letters 17 (2022) Article Number 51 13 pages.

[2] K. Chiba, T. Ohsaka, N. Oyama, *Electrode kinetics of electroactive electropolymerized polymers deposited on graphite electrode surfaces*, Journal of Electroanalytical Chemistry and Interfacial Electrochemistry 217 (2) (1987) 239-251.

[3] G. E. J. Garion, S. Li, N. K. Obhi, C. N. Jarrett-Wilkins, D. S. Seferos, *Programmable assembly of π-conjugated polymers*, Advanced Materials 33 (46) (2021) Article Number 2006287 21 pages.

[4] M. Li, X. Zhang, S. Zhen, J. Xu, *Multichannel transport in conjugated polymers based on through-space conjugated naphthalene*, New Journal of Chemistry 45 (2021) 1795-1799.

[5] A. Khosravi, M. Vossoughi, S. Shahrokhian, Iran Alemzadeh, *HRP-dendron nanoparticles: the efficient biocatalyst for enzymatic polymerization of poly (2,5-dimethoxyaniline)*, Journal of Molecular Catalysis B: Enzymatic 90 (2013) 139-143.

[6] A. Douka, S. Vouyiouka, L. M. Papaspyridi, C. D. Papaspyrides, *A review on enzymatic polymerization to produce polycondensation polymers: the case of aliphatic polyesters, polyamides and polyester amides*, Progress in Polymer Science 79 (2018) 1-25.

[7] I. Sapurina, J. Stejskal, *The mechanism of the oxidative polymerization of aniline and the formation of supramolecular polyaniline structures*, Polymer International 57 (12) (2008) 1295-1325.

[8] X. G. Li, X. L. Ma, J. Sun, M. R. Huang, *Powerful reactive sorption of Silver(I) and Mercury(II) onto Poly(o-phenylenediamine) microparticles*, Langmuir 25 (2009) 1675-1684.

[9] A. G. El-Shekeil, H. A. Al-Saady, F. A. Al-Yusufy, *Synthesis and characterization of some soluble conducting polyazomethine polymers*, Polymer International 44 (1) (1997) 78-82.

[10] S. Kobayashi, A. Makino, *Enzymatic polymer synthesis: An opportunity for green polymer chemistry*, Chemical Reviews 109 (11) (2009) 5288-5353.

[11] P. Xu, A. Singh, D. L. Kaplan, *Enzymatic catalysis in the synthesis of polyanilines and derivatives of polyanilines*, in: S. Kobayashi, H. Ritter, D. Kaplan (Eds), Enzyme-Catalyzed Synthesis of Polymers, Vol 194 of Advances in Polymer Science, Springer, Berlin, Heidelberg, 2006, pp. 69-94.

[12] E. N. Konyushenko, J. Stejskal, I. Šeděnková, M. Trchová, I. Sapurina, M. Cieslar, J. Proke, *Polyaniline nanotubes: conditions of formation*, Polymer International 55 (1) (2006) 31-39.

[13] R. Cervini, X. C., Li, G. W. C., Spencer, A. B., Holme, S. C. Moratti, R. H. Friend, *Electrochemical and optical studies of PPV derivatives and poly(aromatic oxadiazoles)*, Synthetic Metals 84 (1997) 359-360.

[14] L. L. Wu, J. Luo, Z. H. Lin, *Spectroelectrochemical studies of poly-o-phenylamine. Part 1. In situ resonance Raman spectroscopy*, Journal of Physical Chemistry B 112 (23) (2008) 6976-6987.

[15] H. Yang, D. O. Wipf, A. J. Bard, *Application of rapid scan cyclic voltammetry to a study of the oxidation and dimerization of N,N-dimethylaniline in acetonitrile*, Journal of Electroanalytical Chemistry 331 (1-2) (1992) 913-924.

[16] A. T. R. Williams, S. A. Winfield, J. N. Miller, *Relative fluorescence quantum yields using a computer-controlled luminescence spectrometer*, Analyst 108 (1290) (1983) 1067-1071.

[17] D. W. van Krevelen, *Some basic aspects of flame resistance of polymeric materials*, Polymer 16 (8) (1975) 615-620.

[18] İ. Kaya, R. E. İriş, H. K. Yağmur, *Synthesis and characterization of cage structured flame-resistant melamine and pentaerythritol based polymer networks*, Journal of Polymer Research 30 (2023) Article Number 388 15 pages.

Journal of New Results in Science

https://dergipark.org.tr/en/pub/jnrs

Research Article

Open Access

E-ISSN:1304-7981 https://doi.org/10.54187/jnrs.1529229

Quantitative analysis of phenolics in *Trifolium pratense* L. flowers and evaluation of antioxidant activity by sensory

Ramazan Erenler[1], İbrahim Hosaflioğlu[2], İlyas Yıldız[3], Mehmet Nuri Atalar[4], Süleyman Muhammed Çelik[5], Mehmet Hakkı Alma[6]

Keywords:
Trifolium pratense,
Natural compounds,
Quantification,
Sensory,
Antioxidant activity.

Abstract – *Trifolium pratense* L. flowers (TPF) were collected and dried in shade in this study. After extraction in methanol, a diluted solution was applied to the liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) device to determine the bioactive compounds quantitatively. Isoquercitrin (38.64 mg/g extract), coumarin (13.66 mg/g extract), and catechin (12.52 mg/g extract) were verified as major products. Antioxidant activity of TPF was performed using a potentiometric PVC membrane sensor to evaluate 2,2-Diphenyl-1-picrylhydrazyl (DPPH) radical scavenging activity and total phenolic content of TPF. TPF inhibited the DPPH radical as 31.4 ± 0.02% at the 500-ppm concentration. However, the DPPH activity of gallic acid was determined as 81.43 ± 0.07% and 92.44 ± 0.1% at the TPF concentrations of 12.5 ppm and 25 ppm, respectively. In addition, the total phenolic content was calculated to be 82.4 ± 0.15 mg gallic acid eq/g extract. It was observed that TPF has the potential to be an antioxidant and a valuable source of isoquercitrin, coumarin, and catechin.

Subject Classification (2020):

1. Introduction

Plants play a substantial role in drug development since they contain bioactive compounds called secondary metabolites [1-3]. Quantitative analysis of phenolics in plants is crucial for reflecting their significant roles in plant biology, agriculture, medicine, and industry [4,5]. Phenolic compounds are essential in plant growth, development, and defense mechanisms. Quantifying phenolics provides an understanding of how plants respond to various stresses such as drought, salinity, and pests, enabling the development of stress-resistant plant varieties. Phenolic content affects fruits' and vegetables' taste, color, and nutritional quality [6-8]. Determining phenolic compositions helps improve post-harvest storage and processing techniques, enhancing agricultural products' shelf life and quality [9-12]. Many phenolic compounds have strong antioxidant properties, contributing to the health benefits of plant-based foods [13,14]. Quantifying these compounds helps in assessing and promoting dietary sources of antioxidants. Phenolic-rich extracts are used in cosmetics for their antioxidant and anti-aging properties [15-17]. Natural compounds have inspired many synthetic chemists to synthesize them in laboratory conditions [18-23].

Reactive oxygen species (ROS) are chemically reactive molecules in cell signaling pathways that regulate cell proliferation and apoptosis [24,25]. ROS are produced by immune cells to destroy invading pathogens, playing

[1]rerenler@gmail.com (Corresponding Author); [2]hosafli@hotmail.com; [3]lysyldz60@gmail.com; [4]mehmetnuriatalar@hotmail.com, [5]suleymancelik10@gmail.com; [6]mhakki.alma@igdir.edu.tr
[1-6]Research Laboratory Practice and Research Center, Iğdır University, Iğdır, Türkiye
Article History: Received: 06 Aug 2024 — Accepted: 30 Aug 2024 — Published: 31 Aug 2024

a crucial role in the body's defense mechanism. Excessive ROS can cause oxidative damage to lipids, proteins, and DNA, contributing to cellular dysfunction and death [26]. Constant oxidative stress is correlated with the promotion of various chronic diseases, including cardiovascular diseases, neurodegenerative disorders, diabetes, and cancer. To alleviate the harmful effects of ROS, the body develops a complex antioxidant defense system, including enzymatic and non-enzymatic antioxidants [27]. The body's antioxidant system can become insufficient in certain situations, leading to an imbalance between the production of ROS and the body's ability to neutralize them [28]. Therefore, antioxidants are taken from natural products obtained from plants and are rich sources of antioxidants. These compounds help neutralize ROS and reduce oxidative stress, thereby protecting cells from damage [29]. Antioxidants are added to the food to prevent deterioration. However, synthetic antioxidants are restricted as food additives due to suspected carcinogenicity. Hence, the interest in natural products used in food as natural antioxidants has increased considerably [30].

Natural products play an important role in nanotechnology [31-35]. Nanoparticles can be synthesized from natural products [36-39]. Plant extracts act as reducing, capping, and stabilizing agents [40-44]. Many silver nanoparticles were synthesized by plant extracts that revealed considerable biological activities [45-54].

For the assessment of antioxidant assays, spectroscopic techniques like UV-Vis spectroscopy are widely used [55-57]. This conventional method is expensive, time-consuming, necessitates pre-treatment, and requires costly equipment [58]. The potentiometric method is new and efficient and has many advantages for measurement, such as applicability in heterogeneous solutions, simple design, inertness, robustness, integration into computer systems, fast response, and selectivity [59].

Herein, quantitative analysis of phenolic compounds was carried out in *Trifolium pratense* L. flowers by LC-ESI/MS/MS, and a potentiometric PVC membrane sensor determined antioxidant activity and total phenolic content.

2. Materials and Methods

2.1. Plants Materials

Trifolium pratense was obtained from Iğdır University Campus in July 2023 and identified by Dr. Belkız Muca Yiğit, Iğdır University. A voucher specimen was deposited in the herbarium of Iğdır University (No: INWM00000113).

2.2. PVC Membrane Biosensor

The biosensors were designed in two stages. The first one is the preparation of solid contact, in which the end of the copper wires was coated with graphite, epoxy, and hardener. The mixture was prepared by adding THF (3.0 mL), graphite (50%), epoxy (35%), and hardener (15%). Then, the copper wires were immersed into the solid-contact mixture to get the appropriate viscosity and coating. After the copper wires were covered with solid contact, they were kept dark for 24 hours to dry. The second stage includes the preparation of the membrane surface. PVC, gallic acid and plasticizer were mixed in a watch glass, and THF (1.0 mL) was added to homogenize. Afterward, the homogenized mixture was conditioned at rt for 4.0 hours. DPPH-selective and FCR-selective PVC membrane biosensors were prepared, and the activity of plant extract was carried out [59]. The schematic representation is given in Figure 1.

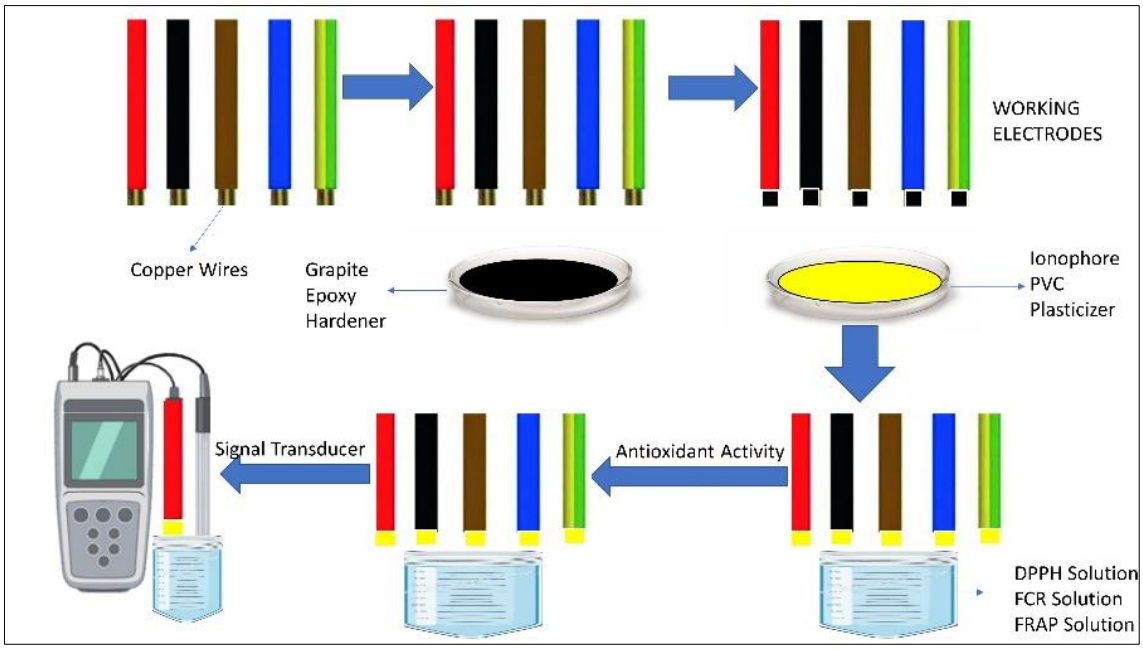


Figure 1. The measurement scheme of antioxidant activity

2.3. DPPH Free Radical Scavenging Effect by PVC Membrane Biosensor

DPPH• free radical effect of TPF was carried out using the biosensor. TPF extract (10 mL, 500 mg/L) was treated with DPPH• solution (10 mL of 100 µg/mL). The potential was measured by immersing the DPPH-SPMB into the solution. Gallic acid was used as a standard. The experiment was repeated three times. DPPH activity was calculated using the following equation (2.1) [59].

$$\%\text{Activity} = \frac{[(E1-E0)-(E2-E0)]}{E1-E0} \times 100 \quad (2.1)$$

Here, E_0 is the potential value of the plant sample, E_1 is the potential value of the standard DPPH solution, and E_2 is the potential value of the DPPH• activity remaining in the medium after a 30-minute reaction [59].

2.4. Total phenolic content analysis

The gallic acid- Folin-Ciocalteu calibration curve was plotted to reveal the gallic acid equivalent to the amount of FCR reduced by the TPF extract. The gallic acid solution was prepared (1.0 - 0.75 - 0.50 - 0.375 - 0.375 - 0.250 - 0.125 - 0.050 - 0.025 mg mL^{-1}). The potential was generated by measurement of the potentiometric responses of these FCR solutions. The TPF solution was prepared (50 mL, 0.25 mg/mL) to determine the total phenolic content of the TPF using the calibration curve. TPF (5.0 mL) was mixed with the deionized water (40.0 mL) and FCR solution (5.0 mL, 0.5 mmol/L) and vortexed. The potential of TPF was measured after the reduction reaction. Total phenolic content was calculated concerning GA equivalent using the equation (E= 0.0513 [GA] + 8.771) R^2 = 0.9996 obtained from the FCR - gallic acid calibration graph [60].

2.5. Quantitative Analysis of Phenolic Compounds by LC-ESI-MS/MS

The bioactive compounds in TPF were determined quantitatively using the Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry (LC-ESI-MS/MS) (An Agilent Technologies 1260 Infinity II, jointed 6460 Triple Quad mass spectrometer) instrument. The sample (50 mg) was dissolved in methanol (1.0 mL), and hexane (1.0 mL) was added. Then, the mixture was subjected to the ultrasonic bath at 10000 rpm for 15 minutes. A sample from the methanol phase (100 µL) was mixed with methanol and water

(each 450 µL). After filtration of the mixture, it was transferred to the vial. Poros hell 120 EC-C18 column was used. An Electrospray ionization (ESI) source was employed with positive and negative ionization modes to detect the compounds' mass-to-ion ratio (m/z). The injection volume was kept at 4.0 µL. The water, including formic acid (0.1%), ammonium format (5.0 mM) A, methanol consisting of formic acid (0.1%), and ammonium format (5.0 mM) B were used as mobile phase. The gradient program was set as 80% for 0-5 min, 70% for 6-15 min, 15% for 16-20 min, 10% for 21-25 min, and 26-30 min 95% was applied in the mobile phase A. The flow rate was 0.4 mL/min, and the gas flow (Nitrogen) was 11 L/min [61].

2.6. Statistical Analysis

GraphPad Prism (8.0.1) was used for statistical analysis. After approving the normality of distribution and homogeneity of the data, the differences of the means of the standard and sample in the same column were submitted to analysis of variance (one-way ANOVA), followed by Tukey's test. Different letters (a, b, c) reveal the significantly different mean in the column. The results were indicated as mean ± standard deviation (SDs). The antioxidant assay and total phenolic content assay were executed in triplicate. The statistical significance level was accepted at $p < 0.05$.

3. Results and Discussion

Antioxidant activity, including DPPH radical scavenging assay, was executed by a novel potentiometric PVC membrane sensor developed by the Isıldak et al. [59]. TPF inhibited DPPH radical by 31.4% at the 500-ppm concentration. Whereas gallic acid inhibited the DPPH radical by 92.4% at 25.0 ppm. Compared to the standard gallic acid, there is a significant difference between gallic acid and TPF. TPF has lower activity than the standard gallic acid statistically. TPF can be considered to have moderate activity on the DPPH radical (Table 1). In the reported study, *Origanum onites* inhibited the DPPH radical by 50.5% at 40 ppm, *Thymus praecox* inhibited the DPPH by 99.55 at 40 ppm, and *Origanum bilgeri* activity was reported as 52.25% at the same concentration [59]. In total phenolic content analysis, TPF was determined to include total phenolic with the value of 82.44 (mg GA/g plant extract). There is an agreement between the total phenolic content and antioxidant activity. Quantitative analysis of bioactive compounds was carried out by LC-MS/MS. Isoquercitrin (38.64 mg/g extract), coumarin (13.66 mg/g extract), and catechin (12.52 mg/g extract) were established as major products (Table 2, Figure 2). Signals marked with an asterisk (*) in Figure 2 indicate compounds not found in the standards.

Table 1. Antioxidant activity and total phenolic content

Sample	Conc. (ppm)	DPPH˙ scavenging effect (%)	Total phenolic mg GA/g plant extract)
TPF	500	31.43 ± 0.02^{a}	82.44 ± 0.15^{a}
Gallic Acid	25	92.44 ± 0.1^{c}	555.625 ± 0.05^{c}
Gallic Acid	12.5	81.43 ± 0.07^{b}	325.625 ± 0.05^{b}

GA: Gallic acid. Different letters (a,b,c) indicate the significantly different of the mean in the column

Flavonoids are a class of plant-derived polyphenolic compounds widely recognized for their beneficial health effects. These compounds are known for their anti-inflammatory, antimicrobial, anticancer, and antiallergic activities. Flavonoids are found in various fruits, vegetables, tea, and other plant-based foods, contributing to the health benefits of a diet rich in these foods. Flavonoids play a crucial role in plant biology, including UV filtration, symbiotic nitrogen fixation, and floral pigmentation [62].

Isoquercitrin is a type of flavonoid, specifically a flavonol glycoside, found in many plants. It is a quercetin molecule bound to a glucose molecule. Isoquercitrin is known for its potent antioxidant properties and

contributes to the overall health benefits attributed to flavonoids. It was reported that Isoquercitrin exhibited biological activities such as antioxidant, anti-inflammatory, antimicrobial, and anticancer properties [63].

Table 2. Quantitative analysis of standard compounds' in *Trifolium pratense* flowers by LC-MS/MS (mg/g extract)

No	Compound	RT	Quantity
1	Catechin	6.904	12.520
2	4-hydroxybenzaldehyde	7.697	0.272
3	Caffeic Acid	7.891	0.494
4	Caffeine	8.498	0.113
5	Vanillin	8.678	0.496
6	p-coumaric acid	9.495	1.249
7	Salicylic Acid	9.871	1.145
8	Coumarin	11.173	13.668
9	Isoquercitrin	11.735	38.642
10	Kaempferol-3-glucoside	13.282	1.766

RT: Retention time

Coumarin is a naturally occurring fragrant organic compound found in many plants, notably in the tonka bean, vanilla grass, sweet woodruff, and some species of cinnamon. It is known for its sweet, vanilla-like aroma and is used in the fragrance and flavor industry [64].

Catechin is a type of natural phenolic compound and antioxidant belonging to the flavonoid family, specifically a subgroup known as flavan-3-ols. Catechins are widely found in various foods and beverages, with particularly high concentrations in tea, cocoa, and certain fruits. Catechins help neutralize free radicals, reducing oxidative stress and cellular damage. Regular consumption of catechin-rich foods is associated with improved heart health, including reduced blood pressure, improved blood lipid profiles, and decreased risk of heart disease. Catechins have been studied for their potential to inhibit cancer cell growth and induce apoptosis in various cancer types. Catechins may protect brain health, potentially reducing the risk of diseases like Alzheimer's and Parkinson's [65].

The medicinal effects of salicylic acid (SA) have been known for years. SA is the phenolic compound plants synthesized and contained in many regulatory pathways. SA has been shown to regulate cell growth, stomatal aperture, respiration, seed germination, fruit yield, nodulation in legumes, and the expression of senescence-related genes. Moreover, it is mostly known for its central role in defense responses [66].

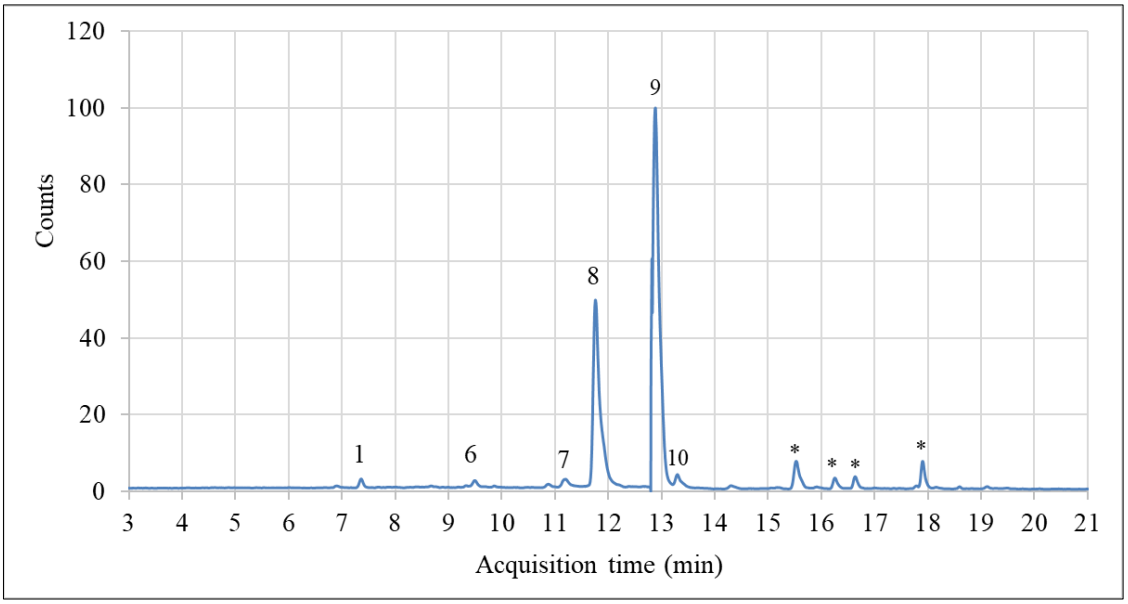


Figure. 2. The MRM chromatogram of *Trifolium pratense* flowers

4. Conclusion

The phytochemistry of *Trifolium pratense* flowers was determined. Quantitative analysis of phenolic compounds in *Trifolium pratense* flowers demonstrated that the corresponding plant included significant compounds for food and pharmaceuticals. This study proved that this plant could be a valuable source of important compounds, including isoquercitrin, coumarin, and catechin. In addition, a new, efficient, sensitive, and fast technique was utilized to determine the antioxidant activity and total phenolic content of TPF. It was presented that there was a correlation between the total phenolic content and antioxidant activity. TPF revealed moderate antioxidant activity in comparison to the standard gallic acid. Due to the importance of antioxidants in food and drugs, investigation of the antioxidant activity and total phenolic content of TPF will make significant contributions to related fields. This study will inspire scientists to study natural products to isolate high concentrations of bioactive compounds from this plant. This is the first report to determine the antioxidant activity and total phenolic content by sensory. Moreover, this study first presented the quantitative analysis of phenolic compounds in TPF.

Author Contributions

All the authors equally contributed to this work. The author read and approved the final version of the paper.

Conflict of Interest

All the authors declare no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] G. Topçu, R. Erenler, O. Çakmak, C. B. Johansson, C. Çelik, H.-B. Chai, J. M. Pezzuto, *Diterpenes from the berries of Juniperus excelsa*, Phytochemistry 50 (7) (1999) 1195-1199.

[2] M. Elmastas, L. Ozturk, I. Gokce, R. Erenler, H. Y. Aboul-Enein, *Determination of antioxidant activity of marshmallow flower (Althaea officinalis L.)*, Analytical Letters 37 (9) (2004) 1859-1869.

[3] A. Sahin Yaglioglu, B. Akdulum, R. Erenler, I. Demirtas, I. Telci, S. Tekin, *Antiproliferative activity of pentadeca-(8E, 13Z) dien-11-yn-2-one and (E)-1,8-pentadecadiene from Echinacea pallida (Nutt.) Nutt. roots*, Medicinal Chemistry Research 22 (6) (2013) 2946-2953.

[4] R. Erenler, I. Telci, M. Ulutas, I. Demirtas, F. Gul, M. Elmastas, O. Kayir, *Chemical constituents, quantitative analysis and antioxidant activities of Echinacea purpurea (L.) Moench and Echinacea pallida (Nutt.) Nutt*, Journal of Food Biochemistry 39 (5) (2015) 622-630.

[5] R. Erenler, I. Demirtas, T. Karan, F. Gul, O. Kayir, O. C. Karakoc, *Chemical constituents, quantitative analysis and insecticidal activities of plant extract and essential oil from Origanum onites L*, Trends in Phytochemical Research 2 (2) (2018) 91-96.

[6] R. Erenler, İ. Telci, M. Elmastaş, H. Akşit, F. Gül, A. R. Tüfekçi, İ. Demirtaş, Ö. Kayir, *Quantification of flavonoids isolated from Mentha spicata in selected clones of Turkish mint landraces*, Turkish Journal of Chemistry 42 (6) (2018) 1695-1705.

[7] N. Genç, M. Elmastaş, İ. Telci, R. Erenler, *Quantitative analysis of phenolic compounds of commercial basil cultivars (Ocimum basilicum L.) by LC-TOF-MS and their antioxidant effects*, International Journal of Chemistry and Technology 4 (2) (2020) 179-184.

[8] R. Erenler, E. N. Geçer, M. H. Alma, İ. Demirtas, *Quantitative analysis of bioactive compounds by high-performance liquid chromatography in origanum bilgeri*, Journal of Integrative and Anatolian Medicine 4 (1) (2022) 15-20.

[9] C. Yaman, Ş. Önlü, H. Ahmed, R. Erenler, *Comparison of phytochemicals and antioxidant capacity of Hypericum perforatum; wild plant parts and in vitro samples*, Journal of Animal & Plant Sciences 32 (2) (2022) 596-603.

[10] M. N. Atalar, R. Erenler, F. Turkan, M. H. Alma, I. Demirtas, A. Baran, S. I. Kandemir, A. I. Kekec, F. Z. Saltan, *Phytochemical analysis and biological activity of Corchorus olitorius L.: Quantitative analysis of bioactive compounds by LC–MS/MS, antibacterial, enzyme inhibition, and cytotoxic activities*, European Journal of Integrative Medicine 62 (2023) Article ID 102290 7 pages.

[11] R. Erenler, M. N. Atalar, İ. Yıldız, E. N. Geçer, A. Yıldırım, İ. Demirtas, M. H. Alma, *Quantitative analysis of bioactive compounds by LC-MS/MS from Inula graveolens*, Journal of Integrative and Anatolian Medicine 4 (3) (2023) 3-10.

[12] A. Saidi, L. Hambaba, B. Kucuk, E. Cacan, R. Erenler, *Phenolic profile, acute toxicity, and hepatoprotective and antiproliferative activities of algerian ruta tuberculata forssk*, Current Bioactive Compounds 18 (3) (2022) 72-83.

[13] Z. Heyem, B. Nassima, M. Ahmed, R. Erenler, Z. Lahcene, B. Fadila, S. Ameddah, *Phytochemical profile, anti-lipid peroxidation and anti-diabetic activities of Thymus algeriensis Boiss. & Reut*, Egyptian Journal of Chemistry 65 (12) (2022) 695-705.

[14] F. Z. Houari, R. Erenler, A. Hariri, *Biological activities and chemical composition of Rubia tinctorum (L.) root and aerial part extracts thereof*, Acta Biológica Colombiana 27 (3) (2022) 403-414.

[15] R. Erenler, T. Karan, İ. Hosaflioglu, *Phytochemical analysis of Syringa vulgaris: Quantitative analysis of natural compounds by LC-ESI-MS/MS*, Turkish Journal of Biodiversity 6 (2) (2023) 75-78.

[16] R. Erenler, C. Yaman, I. Demirtas, M. H. Alma, *Phytochemical Investigation of Hypericum heterophyllum Flowers: LC-ESI-MS/MS analysis, total phenolic and flavonoid contents, antioxidant activity*, The Natural Products Journal 13 (7) (2023) Article ID 120123212672 8 pages.

[17] Y. Başar, R. Erenler, *Phytochemical analysis of Silybum marianum flowers: Quantitative analysis of natural compounds and molecular docking application*, Turkish Journal of Biodiversity 7 (1) (2024) 20-31.

[18] R. Erenler, O. Cakmak, *Synthesis of hexabromo, hydroxy, epoxy, methoxy and nitroxy derivatives of tetralins and naphthalenes*, Journal of Chemical Research (8) (2004) 566-569.

[19] R. Erenler, J. F. Biellmann, *Facile conversion of pyridine propargylic alcohols to enones: Stereochemistry of protonation of allenol*, Tetrahedron Letters 46 (34) (2005) 5683-5685.

[20] O. Cakmak, R. Erenler, A. Tutar, N. Celik, *Synthesis of new anthracene derivatives*, Journal of Organic Chemistry 71 (5) (2006) 1795-1801.

[21] R. Erenler, J. F. Biellmann, *Synthesis of pentafluorophenyl- and pyridinyl-3 allenes*, Journal of the Chinese Chemical Society 54 (1) (2007) 103-108.

[22] R. Erenler, M. Uno, T. V. Goud, J. F. Biellmanna, *Preparation of some heterocyclic enones and ynones by isomerisation of the propargylic alcohols*, Journal of Chemical Research 2009 (7) (2009) 459-464.

[23] J. Lu, I. Maezawa, S. Weerasekara, R. Erenler, T. D. T. Nguyen, J. Nguyen, L. Z. Swisher, J. Li, L.W. Jin, A. Ranjan, S. K. Srivastava, D. H. Hua, *Syntheses, neural protective activities, and inhibition of glycogen synthase kinase-3β of substituted quinolines*, Bioorganic and Medicinal Chemistry Letters 24 (15) (2014) 3392-3397.

[24] R. Erenler, S. Yilmaz, H. Aksit, O. Sen, N. Genc, M. Elmastas, I. Demirtas, *Antioxidant activities of chemical constituents isolated from Echinops orientalis Trauv*, Records of Natural Products 8 (1) (2014) 32-36.

[25] E. A. T. Khodja, R. A. Abd El Hamid Khabtane, D. Benouchenne, M. S. Bensaad, C. Bensouici, R. Erenler, *In vitro assessment of antioxidant, neuroprotective, anti-urease and anti-tyrosinase capacities of Tamarix africana leaves extracts*, Journal of Traditional Chinese Medicine 43 (2) (2023) 252.

[26] M. Elmastaş, İ. Telci, H. Akşit, R. Erenler, *Comparison of total phenolic contents and antioxidant capacities in mint genotypes used as spices*, Turkish Journal of Biochemistry 40 (6) (2015) 456-462.

[27] R. Erenler, O. Sen, H. Aksit, I. Demirtas, A. S. Yaglioglu, M. Elmastas, İ. Telci, *Isolation and identification of chemical constituents from Origanum majorana and investigation of antiproliferative and antioxidant activities*, Journal of the Science of Food and Agriculture 96 (3) (2016) 822-836.

[28] R. Erenler, T. Adak, T. Karan, M. Elmastas, I. Yildiz, H. Aksit, G. Topcu, M. A. Sanda, *Chemical Constituents isolated from Origanum solymicum with Antioxidant activities*, The Eurasia Proceedings of Science, Technology, Engineering & Mathematics 1 (2017) 139-145.

[29] R. Erenler, B. Meral, O. Sen, M. Elmastas, A. Aydin, O. Eminagaoglu, G. Topcu, *Bioassay-guided isolation, identification of compounds from Origanum rotundifolium and investigation of their antiproliferative and antioxidant activities*, Pharmaceutical Biology 55 (1) (2017) 1646-1653.

[30] A. Guzel, H. Aksit, M. Elmastas, R. Erenler, *Bioassay-guided isolation and identification of antioxidant flavonoids from Cyclotrichium origanifolium (Labill.) Manden. and Scheng.*, Pharmacognosy Magazine 13 (50) (2017) 316-320.

[31] R. Erenler, E. N. Gecer, N. Genc, D. Yanar, *Antioxidant activity of silver nanoparticles synthesized from Tagetes erecta L. leaves*, International Journal of Chemistry and Technology 5 (2) (2021) 141-146.

[32] E. N. Gecer, *Green synthesis of silver nanoparticles from Salvia aethiopis L. and their antioxidant activity*, Journal of Inorganic and Organometallic Polymers and Materials 31 (11) (2021) 4402-4409.

[33] N. Genc, *Biosynthesis of silver nanoparticles using Origanum onites extract and investigation of their antioxidant activity*, Particulate Science and Technology 39 (5) (2021) 562-568.

[34] N. Genc, I. Yildiz, R. Chaoui, R. Erenler, C. Temiz, M. Elmastas, *Biosynthesis, characterization and antioxidant activity of oleuropein-mediated silver nanoparticles*, Inorganic and Nano-Metal Chemistry 51 (3) (2021) 411-419.

[35] B. Dag, *Green synthesis, characterization, and antioxidant activity of silver nanoparticles using Stachys annua L. subsp. annua var. annua*, Particulate Science and Technology 40 (4) (2022) 512-520.

[36] R. Erenler, B. Dag, *Biosynthesis of silver nanoparticles using Origanum majorana L. and evaluation of their antioxidant activity*, Inorganic and Nano-Metal Chemistry 52 (4) (2022) 485-492.

[37] R. Erenler, E. N. Gecer, *Green synthesis of silver nanoparticles from Astragalus logopodioides L. leaves*, Turkish Journal of Agriculture-Food Science and Technology 10 (6) (2022) 1112-1115.

[38] R. Erenler, E. N. Gecer, *Synthesis of Silver Nanoparticles Using Sideritis montana L. Leaf Extract: Characterization, catalytic degradation of methylene blue and antioxidant activity*, Journal of Nano Research 75 (2022) 17-28.

[39] E. N. Gecer, R. Erenler, *Biosynthesis of silver nanoparticles using Dittrichia graveolens (Asteraceae) leaves extract: Characterisation and assessment of their antioxidant activity*, Turkish Journal of Biodiversity 5 (1) (2022) 50-56.

[40] E. N. Gecer, R. Erenler, *Biosynthesis of silver nanoparticles using Echinacea pallida (Nutt.) Nutt. and antioxidant activity thereof*, Journal of the Chemical Society of Pakistan 44 (6) (2022) 610-617.

[41] E. N. Gecer, R. Erenler, C. Temiz, N. Genc, I. Yildiz, *Green synthesis of silver nanoparticles from Echinacea purpurea (L.) Moench with antioxidant profile*, Particulate Science and Technology 40 (1) (2022) 50-57.

[42] T. Karan, R. Erenler, B. Bozer Moran, *Synthesis and characterization of silver nanoparticles using curcumin: Cytotoxic, apoptotic, and necrotic effects on various cell lines*, Zeitschrift für Naturforschung C 77 (7-8) (2022) 343-350.

[43] A. Sahin Yaglioglu, R. Erenler, E. N. Gecer, N. Genc, *Biosynthesis of silver nanoparticles using astragalus flavesces leaf: Identification, antioxidant activity, and catalytic degradation of methylene blue*, Journal of Inorganic and Organometallic Polymers and Materials 32 (2022) 3700-3707.

[44] R. Erenler, R. Chaoui, I. Yildiz, N. Genc, E. N. Gecer, C. Temiz, S. Akkal, *Biosynthesis, characterisation, and antioxidant activity of silver nanoparticles using Schinus molle L*, Trends in Sciences 20 (10) (2023) 6105-6105.

[45] R. Erenler, E. N. Gecer, I. Hosaflioglu, L. Behcet, *Green synthesis of silver nanoparticles using Stachys spectabilis: Identification, catalytic degradation, and antioxidant activity*, Biochemical and Biophysical Research Communications 659 (2023) 91-95.

[46] R. Erenler, I. Hosaflioglu, *Green synthesis of silver nanoparticles using Onobrychis sativa L.: Characterization, catalytic degradation of methylene blue, antioxidant activity, and quantitative analysis of bioactive compounds*, Materials Today Communications 35 (2023) 105863.

[47] R. Erenler, R. A. Ojelade, T. Karan, E. N. Gecer, N. Genc, C. Yaman, *Facile, efficient synthesis of silver nanoparticles using Salvia absconditiflora: Assessment of their antioxidant capacity and catalytic activity*, Inorganic Chemistry Communications 158 (2023) 111623.

[48] E. N. Gecer, *Synthesis and characterization of silver nanoparticles using Origanum onites leaves: Cytotoxic, apoptotic, and necrotic effects on Capan-1, L929, and Caco-2 cell lines*, Green Processing and Synthesis 12 (1) (2023) Article ID 20228126 8 pages.

[49] E. N. Gecer, R. Erenler, *Biogenic synthesis of silver nanoparticles using Echium vulgare: Characterisation, quantitative analysis of bioactive compounds, antioxidant activity and catalytic degradation*, Journal of the Indian Chemical Society 100 (5) (2023) Article ID 101003 8 pages.

[50] T. Karan, R. Erenler, Z. Gonulalan, U. Kolemen, *Biogenic synthesis of silver nanoparticles using Sambucus nigra leaves: elucidation, antimicrobial, antioxidant activities and quantification of phenolics*, Chemical Papers 78 (1) (2024) 473-481.

[51] T. Karan, Z. Gonulalan, R. Erenler, U. Kolemen, O. Eminagaoglu, *Green synthesis of silver nanoparticles using Sambucus ebulus leaves extract: Characterization, quantitative analysis of bioactive molecules, antioxidant and antibacterial activities*, Journal of Molecular Structure 1296 (2024) Article ID 136836 8 pages.

[52] Y. B. Karan, R. Erenler, *Synthesis of silver nanoparticles using potato cultivars with their antioxidant activity*, Communications in Soil Science and Plant Analysis 55 (9) (2024) 1249-1257.

[53] N. Nayel, S. Marah, T. Ozen, I. Yıldız, R. Erenler, L. Behcet, *Characterization and comparative investigation of in vitro bioactivities for Lactuca Anatolica root aqueous extract and their green-chemical*

synthesized nanoparticles; Molecular Docking Studies, BioNanoScience (2024) https://doi.org/10.1007/s12668-024-01353-9

[54] A. Seçme, B. M. Bozer, A. Y. Kocaman, R. Erenler, M. H. Calimli, *Synthesis, characterization, and anticancer properties of Ag nanoparticles derived from walnut leaves tested on cells of L929, MCF-7 and H1299*, Journal of Drug Delivery Science and Technology 94 (2024) Article ID 105478 9 pages.

[55] C. Yaman, R. Erenler, M. N. Atalar, Ş. Adem, U. K. Çalişkan, *Phytochemical properties, antioxidant and in vitro/in silico anti-acetylcholinesterase activities of hypericum heterophyllum leaf from Türkiye*, Brazilian Archives of Biology and Technology 67 (2024) Article ID 24230043 14 pages.

[56] F. Z. Houari, M. Brahmi, R. Erenler, A. Hariri, *Chemical profile, antibacterial and antioxidant properties of Rubia tinctorum L. essential oils*, Natural Product Research (2024) 1-10.

[57] B. Bakchiche, G. Hadjira, R. Erenler, İ. Yildiz, A. S. Mohamed, M. A.m. El-shazly, *Mineral contents, antioxidant and antimicrobial activities of algerian terfezia claveryi extracts*, Fabad Journal of Pharmaceutical Sciences 49 (1) (2024) 19-36.

[58] M. Elmastas, S. M. Celik, N. Genc, H. Aksit, R. Erenler, İ. Gulcin, *Antioxidant activity of an Anatolian herbal tea—Origanum minutiflorum: Isolation and characterization of its secondary metabolites*, International Journal of Food Properties 21 (1) (2018) 374-384.

[59] Ö. Isildak, I. Yildiz, N. Genc, *A new potentiometric PVC membrane sensor for the determination of DPPH radical scavenging activity of plant extracts*, Food Chemistry 373 (2022) Article ID 131420 6 pages.

[60] Ö. Işıldak, İ. Yıldız, *New potentiometric sensor for total phenolic assay of plant extracts*, Bulletin of the Chemical Society of Japan 97 (5) (2024) Article ID uoae043 8 pages.

[61] Y. Başar, İ. Hosaflıoğlu, R. Erenler, *Phytochemical analysis of Robinia pseudoacacia flowers and leaf: quantitative analysis of natural compounds and molecular docking application*, Turkish Journal of Biodiversity 7 (1) (2024) 1-10.

[62] I. Demirtas, R. Erenler, M. Elmastas, A. Goktasoglu, *Studies on the antioxidant potential of flavones of Allium vineale isolated from its water-soluble fraction*, Food Chemistry 136 (1) (2013) 34-40.

[63] W. Hassan, G. Rongyin, A. Daoud, L. Ding, L. Wang, J. Liu, J. Shang, *Reduced oxidative stress contributes to the lipid-lowering effects of isoquercitrin in free fatty acids induced hepatocytes*, Oxidative Medicine and Cellular Longevity 2014 (1) (2014) Article ID 313602 18 pages.

[64] X.-M. Peng, G. L. V. Damu, H. Zhou, *Current developments of coumarin compounds in medicinal chemistry*, Current Pharmaceutical Design 19 (21) (2013) 3884-3930.

[65] A. Ganeshpurkar, A. Saluja, *The pharmacological potential of catechin*, Indian Journal of Biochemistry and Biophysics (IJBB) 57 (5) (2020) 505-511.

[66] J. L. Boatwright, K. Pajerowska-Mukhtar, *Salicylic acid: An old hormone up to new tricks*, Molecular Plant Pathology 14 (6) (2013) 623-634.