Neutrosophic Sets and Systems {Special Issue: Neutrosophy and Plithogeny: Fundamentals and Applications}, Vol. 69, 2024

University of New Mexico

Neutrosophic Clustering Analysis with Data from Cysticercus Tenuicollis DNA Samples

Blanca Mercedes Toro Molina[1], Rafael Alfonso Garzón Jarrin[2], Mauricio Rafael Aguilera Vizuete[3], Christian Geovanny Ávila Rocha[4], Luis Favian Cartuche Macas[5], and Edilberto Chacón Marcheco[6]

[1] Technical University of Cotopaxi, Cotopaxi. Ecuador. E-mail: blanca.toro@utc.edu.ec
[2] Technical University of Cotopaxi, Cotopaxi. Ecuador. E-mail: rafael.garzon@utc.edu.ec
[3] Technical University of Cotopaxi, Cotopaxi. Ecuador. E-mail: mauricio.aguilera@utc.edu.ec
[4] Technical University of Cotopaxi, Cotopaxi. Ecuador. E-mail: christian.avila@utc.edu.ec
[5]Universidad Intercultural de las Nacionalidades y Pueblos Indígenas Amawtay Wasi, Quito, Ecuador. E-mail: luis.cartuche@uaw.edu.ec
[6] Technical University of Cotopaxi, Cotopaxi. Ecuador. E-mail: edilberto.chacon@utc.edu.ec

Abstract. Parasitic diseases caused by *Cysticercus tenuicollis* significantly impact the livestock industry. The quantification of such impacts presents challenges due to the handling of data obtained through modern molecular techniques. This investigation explores the application of neutrosophic cluster analysis to the study of *Cysticercus tenuicollis* DNA. Neutrosophic clustering and the use of linguistic terms help interpret analyses, employing a five-step neutrosophic clustering algorithm: selection of the clustering algorithm, definition of distance metrics, data preparation, algorithm execution, and examination of results. Samples from 20 sheep exhibiting liver cysts were collected in Cotopaxi Province, Ecuador. The analysis resulted in a neutrosophic cluster where one group displayed low truth and falsity values but a moderate level of indeterminacy. The stability of the clusters was evaluated using the Adjusted Rand Index (ARI) with 100 bootstrapping repetitions, yielding a moderate value of approximately 0.52, which implies that the detected cluster structures are robust and representative of the underlying patterns in the data. The clusters show a clear differentiation based on the initial DNA concentration. Samples with high DNA concentrations tend to group together, suggesting that the quality and quantity of DNA may be critical indicators of the samples' condition and possibly the presence and state of the parasite.

Keywords: Neutrosophic Clustering Analysis, DNA, clustering algorithm, Cysticercus tenuicollis, molecular techniques.

1 Introduction

The study of parasitic diseases, such as those caused by *Cysticercus tenuicollis*, a significant parasite in the livestock sector, is crucial for the development of effective control and eradication strategies [1]. This parasite, whose prevalence in animals like sheep can lead to substantial economic losses, has been the subject of numerous studies aiming to better understand its biology and transmission. Recently, molecular identification has enabled significant advancements in this area, providing more accurate methods for studying and diagnosing the parasite [2].

However, analyzing the vast amount of genetic data obtained through modern molecular techniques presents challenges, especially in terms of managing uncertainty and interpreting genetic variability [3]. Neutrosophic cluster analysis emerges as a powerful tool, facilitating the grouping of complex DNA data from Cysticercus tenuicollis into clusters that reflect significant similarities and differences, even under conditions of uncertainty [4]. This approach enhances understanding of the parasite's genetic structure and variability, supporting the development of more targeted and efficient intervention measures [5].

This is a globally occurring disease with a low incidence, but when it does occur, it can cause losses due to death ranging from 3% to 25%. The disease was first reported in California, USA in 1916 and has since been described in other parts of the world. Thus, this disease has a global distribution, concentrating in humid areas and lowlands with predominantly alkaline soils and waters [5].

This article explores the application of neutrosophic cluster analysis in the study of Cysticercus tenuicollis DNA, highlighting how this method can overcome the limitations of traditional analytical approaches and significantly contribute to the field of veterinary parasitology [6]. By integrating neutrosophic set theory into cluster analysis, a new dimension is opened in the study of pathogens, enhancing our ability to interpret complex data and make informed decisions in the fight against parasitic diseases.

2 Background of Cysticercus tenuicollis in sheep

Currently, sheep are highly valued mammals in our region, especially in the Central Sierra of Ecuador. Their geographic distribution is broad because these animals exist worldwide, and their adaptability allows them to be constantly managed in rustic conditions across various climates and ecologies; therefore, this species has been produced and exploited in extensive pasture areas [7]. Moreover, sheep have very distinct and specific feeding habits; unlike most species, they refuse to eat or drink from dirty feeders or troughs. They require clean food and water, free of feces or any other contaminants [8].

Animal behavior is the ultimate expression of adaptability or adjusting internal conditions to the external environment, considering that the animal's body responds as a whole to a stimulus. Thus, every manifestation of behavior is directed toward satisfying one of the three basic needs of life: feeding, defense, and reproduction. For sheep, nutritional requirements are stronger than defensive ones, and defensive needs are stronger than reproductive ones [9].

In this sense, most sheep are found infested by external parasites and, more so, by internal parasites, which have secondary life cycles that occur across different pastures. Without a proper parasite control program or a deworming schedule, sheep can suffer severe parasitism, leading to weight gain deficiencies, especially in lambs and adult animals intended for reproduction, resulting in complex cases of anemia, diarrhea, and in the most neglected cases, death at any age [9].

Cysticercosis in wild and domestic animals is caused by the larval stages (metacestodes) of tapeworms (cestodes), with these adult phases located in the intestines of dogs and wild canids, and even findings in cats [10]. This type of cysticercosis particularly causes economic losses due to the seizure of meats and residues infected with these types of parasites, as well as specific organs such as livers and even lungs [11].

Based on various research studies and investigations on Cysticercus tenuicollis, it is determined that a parasite is any living being that spends all or part of its life at the expense of another living being, called a host (intermediate or definitive), from which the parasite lives, causing harm or not, and on which it has an obligatory dependence. Similarly, a parasite is an animal or plant organism that lives only at the expense of another organism, either on it or inside it [12]. In this regard, there are currently two types of parasites; internal parasites that maintain a part of their cycle in the bodies of animals and eventually humans, causing disorders [13], and external parasites that live outside an animal organism, such as fleas, ticks, lice, among others, which are easier to detect with some exceptions, like the agents producing scabies, but which can be effectively treated with a simple clinical diagnosis [13].

Indeed, the etiological agent causing hepatoperitoneal cysticercosis in sheep, which typically affects other mammalian animals including humans, is Cysticercus tenuicollis, the metacestode of Taenia Hydatigena [10]. According to recent studies, the type of parasites Taenia Hydatigena and in its larval stages such as Cysticercus tenuicollis and Toxocara canis can sometimes be transmitted to humans and can be widely distributed in the cells of the parenchyma, the basal membrane of the tegument, and the apical surface of the epithelial cells, which surround and cover the ceca; this type of cysticercus in large quantities can lead to severe parasitic infections that may induce the development of specific cancer [10].

In the case of Cysticercus tenuicollis, it has also been found in the pudú, a native deer of South America commonly found in countries like Chile and Argentina. This parasite has been observed in the bronchi of this deer, in both males and females, young and adults; these cases are among the first records of deer acting as hosts, due to the establishment of European deer populations, which presumably causes the spread of Cysticercus tenuicollis by sharing various pasture sites that practically cause a high additional impact, both in exotic mammals and in farm mammals [7].

In the case of cattle, diagnosis, as in sheep, is normally carried out during post-mortem inspection, where the cysticerci (vesicles) found present whitish or slightly turbid vesicles, where a single scolex is clearly recognized. In certain cases, these vesicles may present a reddish hue, with the scolex difficult to identify due to its coloring, and in some Cysticercus tenuicollis, rounded formations are found, already encapsulated or calcified in specific organs such as the liver and lungs [9].

2.1 Description of clinical symptoms

Due to the acute to subacute progression of this disease, it is very difficult to observe clinical symptoms, especially under field conditions. What was observed upon arrival at the facility was that the animals exhibited bloody diarrhea, colored urine, a retracted abdomen, and a staggering gait. Animals showing clinical symptoms would fall and exhibit pedaling motions, opisthotonos, nystagmus, bloody foam from the nose, and vocalizations. Rectal temperature remained within normal parameters. The oral, ocular, and vulvar mucosa exhibited an anemic coloration with a jaundiced tint. Subsequently, it was observed that these symptoms led to the facility manager requesting the euthanasia of the animal for subsequent necropsy.

3 Neutrosophic analysis of the data

Neutrosophy is a new branch of philosophy that studies the origin, nature, and scope of neutralities, created by Professor Florentin Smarandache. Its incorporation ensures that the inherent uncertainty in decision-making is considered, including indeterminacies where experts will issue their judgments evaluating linguistic rather than numerical terms, which constitutes the most natural form of measurement in humans [14]. Neutrosophic logic and sets, on the other hand, represent a generalization of Zadeh's fuzzy logic and sets, particularly of Atanassov's intuitionistic logic, with multiple applications in the field of decision-making and machine learning [14].

The veracity of a proposition in neutrosophic analysis is as follows [15]: Let $N = \{(T, I, F): T, I, F \subseteq [0,1]\}n$, a neutrosophic evaluation is a mapping of a group of proposition formulas to N, and therefore p:

$$v(p) = (T, I, F) \quad (1)$$

All this facilitates practical application in real-world problems [16], where the analysis of Single-Value Neutrosophic Sets (SVNS) also provides linguistic terms to achieve a greater interpretation of them [14]. In this context, X represents the universe as an SVNS. Thus, X can be determined in the following way [14]:

$$A = \{\langle x, u_a(x), r_a(x), v_a(x)\rangle: x \in X\} \quad (2)$$

Where $u_a(x): X \to [0,1], r_a(x): X \to [0,1]$ y $v_a(x): X \to [0,1]$

With

$$0 \leq u_a(x), r_a(x), v_a(x) \leq 3, \forall\, x \in X \quad (3)$$

The interval $u_a(x), r_a(x)$ and $v_a(x)$ denotes the neutrosophic true, indeterminate, and false values from x in A, respectively [15]. For convenience, an SVNN is expressed as A = (a, b, c), where a, b, c are in the range [0, 1] and $0 \leq a + b + c \leq 3$ [17].

Neutrosophic numbers and linguistic terms in the form (T, I, F) are shown in Table 1.

Table 1: Linguistic terms according to the method and single-valued neutrosophic numbers.

Initial Concentration		DNA Volume		Water Volume	
[0.35, 0.65, 0.05]	(Medium, Medium, Minimum)	0.40, 0.60, 0.05]	(Medium, Medium, Minimum)	[0.60, 0.40, 0.05]	(Medium, Medium, Minimum)
[0.60, 0.40, 0.05]	(Medium, Medium, Minimum)	[0.13, 0.87, 0.05]	(Low, High, Minimum)	[0.87, 0.13, 0.05]	(High, Low, Minimum)
[1.00, 0.00, 0.05]	(High, Low, Minimum)	[0.00, 1.00, 0.05]	(Low, High, Minimum)	[1.00, 0.00, 0.05]	(High, Low, Minimum)
[0.61, 0.39, 0.05]	(Medium, Medium, Minimum)	[0.12, 0.88, 0.05]	(Low, High, Minimum)	[0.87, 0.13, 0.05]	(High, Low, Minimum)
[0.35, 0.65, 0.05]	(Medium, Medium, Minimum)	[0.33, 0.67, 0.05]	(Low, High, Minimum)	[0.67, 0.33, 0.05]	(High, Low, Minimum)

These terms help interpret the degrees of truth, falsehood, and indeterminacy in more accessible terms such as "Low", "Medium", and "High" for T and F, and "Minimal", "Moderate", and "High" for I. This structure enriches the analysis by allowing not only the viewing of numerical data but also a quick understanding of their meaning in the context of neutrosophic analysis.

3.1 Neutrosophic Cluster Analysis

Cluster Analysis[18, 19] is the generic name for a wide variety of procedures that can be used to create a classification. More specifically, a clustering method is a multivariate statistical procedure that starts with a dataset

containing information about a sample of entities and attempts to reorganize them into relatively homogeneous groups, which we will call clusters [20].

In Cluster Analysis, little or no information is known about the structure of the categories, which differentiates it from multivariate assignment and discrimination methods. What is available is a collection of observations, with the operational objective in this case being to discover the structure of the categories into which the observations fit. More specifically, the goal is to organize the observations into groups such that the degree of natural association is high among members of the same group and low among members of different groups. Although little or nothing is known about the structure of the categories a priori, there are often some notions about desirable and undesirable characteristics when establishing a particular classification scheme. Operationally, the analyst is sufficiently informed about the problem such that they can distinguish between good and bad category structures when they encounter them [21].

To carry out the neutrosophic cluster analysis with the DNA sample data of Cysticercus tenuicollis in sheep from Ecuador, the following steps were taken:

- **Step 1.** Selecting the Clustering Algorithm

Given that we are working with neutrosophic data, a clustering method that can adapt or be modified to handle this type of data is necessary. The K-means algorithm [22]is commonly used, but it must be adapted to work with the T, I, and F components.

- **Step 2**. Definition of the Distance Metric

Euclidean distance is not suitable for neutrosophic numbers due to their three-dimensional nature and the interpretation of T, I, and F. It will be required to define a metric that can effectively measure the distance between neutrosophic data points considering all components.

- **Step 3.** Data Preparation

The original measurements of each sample were converted into neutrosophic values as we have already calculated (T, I, F). We will use these values for clustering.

- **Step 4**. Execution of the Algorithm

The selected algorithm is implemented in Python, using the defined distance metric, and we apply the clustering to the prepared data.

- **Step 5.** Analysis of the Results

Evaluate the formed clusters to understand how the samples are grouped according to their neutrosophic characteristics and to interpret the results biologically.

3.2 Collection and processing of Cysticercus tenuicollis DNA samples in sheep

Twenty cyst samples (ranging from 2 to 5 cm in size) were collected from sheep livers slaughtered at the Metropolitan Public Company of Rastro Quito (EMRAQ-EP) and in the communities of the Cantons Pujilí and Saquisilí in the Province of Cotopaxi, Ecuador. Molecular analysis of the samples was conducted at the Research Laboratory of the Universidad de las Américas (UDLA), Quito-Ecuador. DNA was extracted from a portion of tissue from each cysticercus using the Phenol-Chloroform method [23]:

- Dissect 2 mm of tissue sample and transfer it to a new 1.5 mL tube.
- Add 500 µL of Extraction Buffer and macerate with a pipette tip or pestle for 5 minutes inside a cold block.
- Vortex for 1 minute and add 5 µL of Proteinase K (20 mg/mL).
- Incubate the samples at 56°C overnight with shaking at 300 rpm.
- Remove the samples from the heating block, and let them rest for 5 minutes at room temperature.
- Add 750 µL of Phenol/Chloroform/Isoamyl and vortex until a milky emulsion forms.
- Centrifuge for 10 minutes at 4°C at maximum speed. Transfer all the aqueous phases to a new 1.5 mL microtube.
- To this phase, add 500 µL of Chloroform: Isoamyl, and homogenize by vortexing.
- Centrifuge for 10 minutes at 4°C at maximum speed. Transfer all the aqueous phases to a new 1.5 mL microtube.
- To the new microtube, add 400 µL of 100% Isopropanol and 40 µL of 3M NaCl.
- Mix well and let it sit at room temperature for two hours.
- Centrifuge for 30 minutes at 4°C at 15,000 rpm.
- Carefully discard the supernatant without disturbing the pellet.
- Add 1000 uL of ice-cold 70% Ethanol and centrifuge for 15 min at 15,000 rpm.
- Discard the supernatant and allow it to dry.
- Resuspend in 40 µL of Milli Q Water or TE and incubate the sample at 37°C for two hours.
- Quantify the DNA concentration.

The quantification of DNA concentrations was performed using a spectrophotometer-nanodrop, after which it was diluted to a final concentration of 5 ng/µL as shown in Table 2. The samples were stored in a freezer at -20°C until use.

Table 2: Dilutions of Cysticercus tenuicollis DNA samples.

CYSTICERCUS DILUTIONS			
FINAL VOLUME (µL)			20
FINAL CONCENTRATION (ng/µL)			5
SAMPLE ID	**INITIAL C.**	**V. DNA**	**V. WATER**
C001	73.3	1.3	18.64
C002	212.8	0.47	19.53
C003	305.4	0.33	19.67
C004	215.5	0.46	19.54
C005	150	0.67	19.33
C006	275.5	0.36	19.64
C007	126.2	0.79	19.21
C008	273.5	0.37	19.63
C009	236.5	0.42	19.58
C010	109.3	0.91	19.09
C011	136.3	0.73	19.27
C012	73.1	1.37	18.63
C013	124.8	0.8	19.2
C014	155.2	0.64	19.36
C015	107	0.93	19.07
C016	76.8	1.3	18.7
C017	101.6	0.98	19.02
C018	95.9	1.04	18.96
C019	109.3	0.91	19.09
C020	126.4	0.79	19.21

3.3 Amplification of DNA fragments

In the study, a total of 454 bp of subunit 1 of the NADH dehydrogenase (nad1) from the mitochondrial DNA of T. hydatigena isolated from sheep were sequenced [11]. PCR amplification of the mitochondrial nad1 gene was performed using the following pair of primers specific for cestode parasites, described by [24]:

5´-CARTTTCGTAAGGGBCCWAAWAAGGT-3´ forward
5´-CCAATTTCYTGAAGTTAACAGCATCA-3´ reverse

The PCR reactions were carried out in a final volume of 15 µl for each sample as shown in Table 3.

Table 3: Components for the nad1 gene amplification PCR. Source: own elaboration.

COMPONENT	C. Initial	C. End	Vol. 1 Rx (µL)	Vol. 60 Rx (µL)
MQ water			8.8	528
Buffer	10X	1X	1.5	90
MgCl2	50mM	1.5mM	0.45	27
dNTP Mix	10mM	0.2	0.3	18
First Fw	20 µM	0.5µM	0.375	22.5
First RV	20 µM	0.5µM	0.375	22.5
Taq Polymerase*	10 U/uL	2 U/uL	0.2	12
DNA	5ng/µL	1ng/µL	3	
Final Volume			15	

The PCR reactions were performed with a final volume of 15 µl for each sample. The fragments were generated by PCR using thermocyclers, under the following conditions: initial denaturation temperature of 94°C for 2 min (1 cycle), followed by 35 cycles of denaturation at 94°C for 1 min, primer annealing at 58°C for 30 seconds, extension at 72°C for 1 min, and a final extension of the products at 72°C for 5 min, ending with a 10 min rest at 4°C.

The PCR products were visualized on 2% agarose gels to identify possible failures. The gels were stained with SYBR Safe, analyzed under ultraviolet light, and photographed. In each case, a negative control was run to check for contamination. The amplified product was sequenced using the BigDye v3.1 sequencing kit (Applied Biosystems). The sequencing reactions were purified using an enzymatic purification protocol based on gel filtration resin. The Sanger sequencing matrix (BigDye - 3.1) was analyzed by capillary electrophoresis. The sequences were aligned and edited with the MEGA X software, using the reference sequence (MN175584.1), available on GenBank. The determination of the identified polymorphisms (number of haplotypes, number of mutations, as well as haplotype diversity and nucleotide diversity), was performed using DnaSP version 6.

Phylogenetic trees generated among sequences of the Nad1 gene (454 bp) from Cysticercus tenuicollis and the comparison with other sequences deposited in GenBank with accession numbers: MN175584.1 for Taenia hydatigena from Nigeria and FJ440841.1 for Taenia regis from Uganda, were constructed using the UPGMA (unweighted pair group method using arithmetic averages) and NJ (neighbor-joining) methods [12]. The reliability of the obtained trees was determined with a bootstrap test (500 replicates), shown next to the branches [11].

4 Results

Figure 1 displays the neutrosophic cluster obtained from the analysis of DNA data from Cysticercus tenuicollis in sheep from Ecuador. The charts reveal groupings based on the neutrosophic characteristics of the samples, aiding in identifying relationships and differences among data groups. The variability among the clusters in each chart could indicate different behaviors or characteristics inherent to the samples, useful for more detailed analysis or specific decisions based on the identified groups.

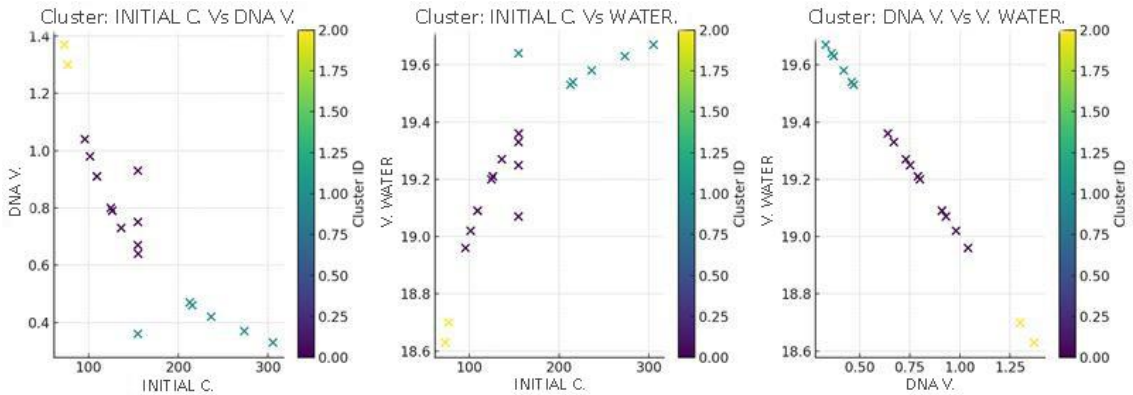

Figure 1: Neutrosophic clusters obtained from the analysis of DNA data of Cysticercus tenuicollis in sheep from Ecuador. Source: own elaboration.

1. INITIAL C. vs V. ADN: This chart shows how the samples are distributed in terms of ' INITIAL C. ' and 'V. ADN'. The clusters are clearly differentiated, indicating distinct patterns in these two parameters.

2. INITIAL C. vs V. WATER: Similar to the first chart, this one displays the distribution of samples for ' INITIAL C. ' and 'V. WATER'. A separation of clusters is observed, suggesting differences in how these two variables interact.

3. V. ADN vs V. WATER: This chart compares 'V. ADN' with 'V. WATER', also shows a clear distinction between clusters.

To assess the stability of the neutrosophic clusters, a resampling or bootstrapping approach is applied. This method involves repeating the clustering process multiple times with different subsets of data and observing how consistently the samples are grouped into the same clusters in each iteration. The stability of the clusters can indicate how robust and reliable the detected grouping patterns are.

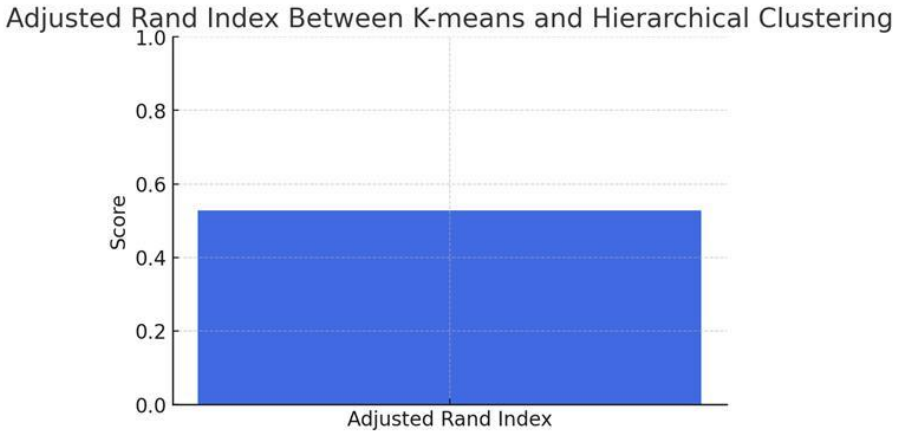

Figure 2: Stability analysis of the neutrosophic cluster obtained from DNA of Cysticercus tenuicollis in sheep from Ecuador. Source: Own elaboration.

The evaluation of cluster stability, using the Adjusted Rand Index (ARI) over 100 bootstrapping iterations, resulted in an average value of approximately 0.52 as shown in Fig. 2. This suggests that, although there is some consistency between the clusters generated by both methods, there are also significant differences. This may be due to the nature of the algorithms and how they handle the structure of the data in the clusters generated from the neutrosophic DNA samples of Cysticercus tenuicollis in sheep from Ecuador [19]. A high ARI suggests that samples tend to be grouped consistently into the same clusters across different subsets of data, which supports the reliability of the clustering analysis performed. This implies that the detected cluster structures are robust and representative of the underlying patterns in the data.

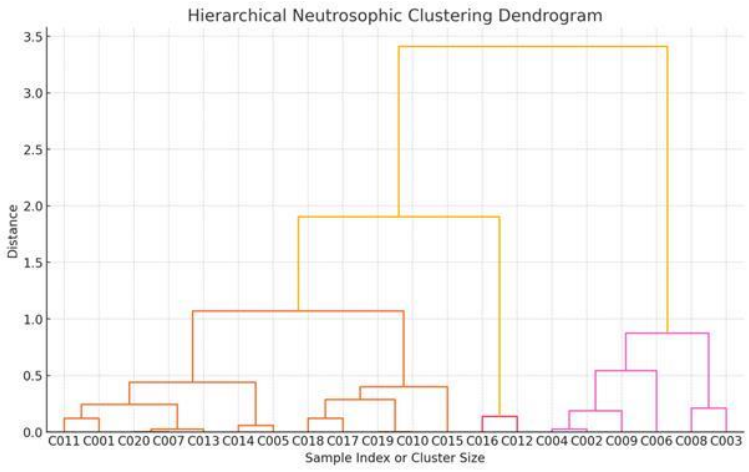

Figure 3: Hierarchical clustering analysis with the neutrosophic data from DNA samples of Cysticercus tenuicollis isolated in sheep from Ecuador. Source: Own elaboration.

Figure 3 shows the hierarchical clustering analysis of the neutrosophic data from DNA samples of Cysticercus tenuicollis isolated in sheep from Ecuador. This graph demonstrates how the samples are grouped based on their neutrosophic characteristics (T, I, F). The height in the dendrogram where two groups join reflects the clustering distance between them, which is a measure of their dissimilarity. Different 'cuts' in the dendrogram can be observed to decide the most appropriate number of clusters. For example, a cut at a height of about 10 could indicate the formation of two or three main groups, depending on how you prefer to interpret the separation between the groups.

A total of 2 polymorphic sites were detected in the Nad1 gene of the mitochondrial DNA, all being parsimony informative. Three different haplotypes (h) were observed, with haplotype 1 being predominant, represented by 85% of the sequences.

Hap_1: 17 sequences [Nad-Cox1-Fw-CT001 Nad-Cox1-Fw-CT002 Nad-Cox1-Fw-CT003 Nad-Cox1-Fw-CT004 Nad-Cox1-Fw-CT005 Nad-Cox1-Fw-CT007 Nad-Cox1-Fw-CT008 Nad-Cox1-Fw-CT009 Nad-Cox1-Fw-CT011 Nad-Cox1-Fw-CT012 Nad-Cox1-Fw-CT013 Nad-Cox1-Fw-CT014 Nad-Cox1-Fw-CT015 Nad-Cox1-Fw-CT016 Nad-Cox1-Fw-CT017 Nad-Cox1-Fw-CT018 Nad-Cox1-Fw-CT019]

Hap_2: 2 sequences [Nad-Cox1-Fw-CT006 Nad-Cox1-Fw-CT020]

Hap_3: 1 sequence [Nad-Cox1-Fw-CT010]

The effectiveness of molecular tools in identifying and characterizing different species and genotypes of tapeworms has been proposed by authors such as [3-6], with results based on the existence of interspecific variations for molecular characters [8].

Table 4 shows values of (0.279) in reference to haplotype diversity and (0.00101) for nucleotide diversity when evaluating the 454 bp fragment of the Nad1 gene from the mitochondrial DNA of Cysticercus tenuicollis.

Table 4: Variability parameters in 641 bp of the Nad1-Cox1 gene, from Cysticercus tenuicollis sequences.

Characteristic	Total
Number of sequences	20
Number of mutations	2
Haplotypes	3
Haplotype Diversity	0.279
Standard deviation	0.123
Significance	0.00101
Nucleotide Diversity	0.00046

These results coincide with those reported for Taenia hydatigena by [7] who conducted a molecular description, phylogeny, and genetic variation of Taenia hydatigena from sheep and goats in Nigeria, based on the mitochondrial Nad1 gene, and are very close to those described by others who developed the molecular identification of Taenia hydatigena from sheep in Khartoum, Sudan [9].

5 Discussion

Choosing a good protocol for DNA extraction is crucial for DNA amplification by PCR. The most commonly used technique is organic solvents like phenol/chloroform; however, this method has drawbacks in terms of time and toxicity, making it not entirely feasible since extraction protocols should be simple, quick, and reproducible while maintaining the safety and integrity of the sample. For this reason, the DNA extraction protocol based on organic solvent phenol/chloroform was used.

All 20 samples were analyzed using the selected protocol, achieving a 100% yield, meaning 20 samples were identified using specific primers for Cysticercus Tenuicollis; thus, the chosen DNA extraction protocol was ideal. However, a study by [16] determined that this classic protocol had a low yield, which could be attributed to possible oxidation of the used phenol; it also involves many steps, making it lengthy and laborious, potentially leading to DNA loss during the various steps of the technique, including the use of polluting and toxic solvents. Nevertheless, the integrity of the DNA was maintained in optimal conditions, showing a single large band as in the processed samples of Cysticercus Tenuicollis

This research generates new topics for further study to identify the gender of Cysticercus Tenuicollis that takes sheep as hosts and the possible damage they can cause to them. The impact of this project affects animal health since contracting gastrointestinal parasites decreases the vitality and well-being of canines, besides becoming a

zoonotic factor related to public health issues since sheep may ingest their eggs expelled in the feces of infected canines The environmental impact occurs when the parasite's eggs are expelled through feces and spread in the environment, causing reinfection in livestock (such as sheep, cattle, and pigs) and humans, thus continuing the biological cycle of the parasite. As a disease that affects the animal and deteriorates its health, it also has an economic impact as the diagnosis and treatment of canines reduce the owner's finances, and in most cases, the animals are not treated The technical impact is evidenced in the performance of molecular diagnosis of the parasite to generate specific data on the identification of the genus Cysticercus Tenuicollis in slaughtered sheep. Finally, the social impact must be considered in informing pet owners about the results found in the laboratory diagnosis and molecular identification and subsequently informing them about the symptoms and lesions that cause the deterioration of the animal's health and the potential contagion through the spread of zoonosis.

6 Conclusion

The clusters formed indicate a clear differentiation based on the initial DNA concentration. Samples with high DNA concentration tend to cluster together, suggesting that the quality and quantity of DNA can be critical indicators of the sample's condition and possibly the presence and state of the parasite.

The neutrosophic values (T, I, F) provide deeper insights into the uncertainty and quality of the samples:

- High Truth (T): Samples with high DNA concentrations tend to have a high T value, indicating better sample quality for future analysis.
- Low Falsehood (F): Samples with less water volume (less dilution) exhibit lower F values, which could be interpreted as less contamination or interference in the sample.
- Indetermination (I): Samples with medium DNA volume values show variability in I, reflecting the uncertainty in the precision of DNA measurement.

The use of neutrosophic numbers has allowed for the evaluation of multiple aspects of the samples simultaneously, highlighting the utility of this approach in contexts where uncertainty and variability are significant. This methodology could be adapted to other types of biological or medical analysis where data interpretation is complex. The evaluation of cluster stability through bootstrapping techniques showed that the clusters are consistently reproducible, indicating that the groupings found are robust and reliable.

Clustering samples based on neutrosophic values and their DNA concentration can have important implications for epidemiological studies of Cysticercus tenuicollis in sheep, helping to identify potential infection foci or differences in parasitic load.

Recommendations for future analyses should explore other factors such as the geographic location of the sheep, age, and other clinical data to see how these influence the clusters. Validation of these findings with alternative analysis methods and in a larger dataset is necessary to generalize the results. Practical applications of using these findings include improving sampling and diagnostic strategies for cysticercosis in sheep, thus optimizing control and prevention efforts.

This neutrosophic analysis provides a solid foundation for better understanding the variability and quality of DNA samples, offering a valuable tool for scientific research and practical applications in veterinary and public health.

7 References

[1] S. G Kilinc, H.K Kesik, and S Simsek. Molecular characterization and haplotypes of sheep and goat isolates of Cysticercus tenuicollis in Turkey. Parasitology, 146(8), 2019, 1047-1054.

[2] A.A Hama, R. H Zorab, F. Ali, A.M Salih, and A.A Hassan. Prevalence and molecular characterization of Cysticercus tenuicollis isolated from some intermediate host in Kurdistan-Iraq. Kurdistan Journal of Applied Research, 2018, 177-182.

[3] T. A Hama-Soor, A.A Mohammed, and S. Weli. Novel isolate of Cysticercus tenuicollis-Kalar isolate has been revealed in Kalar, Iraq. The Journal of Infection in Developing Countries, 15(10), 2021, 1532-1538.

[4] I. Abbas, E. El-Alfy, E. Janecek-Erfurth, and C. Strube. Molecular characterization of Cysticercus tenuicollis isolates from sheep in the Nile Delta, Egypt, and a review on Taenia hydatigena infections worldwide. Parasitology, 148(8), 2021, 913-933.

[5] M. H Radfar, S. Tajalli, and M. Jalalzadeh, Prevalence and morphological characterization of Cysticercus tenuicollis (Taenia hydatigena cysticerci) from sheep and goats in Iran. Veterinarski arhiv, 75(6), 2005, 469.

[6] H.S Farhan, S. A Al-Azizz, and S.A Al-Jadaan, Molecular Characterization of Cysticercus tenicollis from Slaughtered Sheep and Goats in Basrah Province, Iraq. Revista Electronica de Veterinaria, 23(4), 2022, 112-120.

[7] G Cengiz, G. Yucel Tenekeci, and N. Bilgen. Molecular and morphological characterization of Cysticercus tenuicollis in red deer (Cervus elaphus) from Turkey. Acta Parasitologica, 64, 2019, 652-657.

[8] Y. Zhang, W Zhao, D. Yang, Y Tian, W Zhang, and A. Liu, A. Genetic characterization of three mitochondrial gene sequences of goat/sheep-derived Coenurus cerebralis and Cysticercus tenuicollis isolates in Inner Mongolia, China. Parasite, 25, 2018.

[9] A. A Mohammed. Molecular study of from slaughtered sheep in Sulaymaniyah province, Iraq. Journal of Veterinary Research, 64(2), 2020, 275-280.
[10] F. Celik, S.G Kilinc, H.K Kesik, H. Ahmed, and S. Simsek. In-silico analysis of mt-CO1 gene of sheep isolates. Helminthologia, 58(1), 2021, 59-67.
[11] K. Adwan, A. Jayousi, S. Abuseir, I. Abbasi, G. Adwan, and N. Jarrar, N. Genetic diversity of Taenia hydatigena in the northern part of the West Bank, Palestine as determined by mitochondrial DNA sequences. Acta Parasitologica, 63(2), 2018, 299-303.
[12] H. Luo, H. Zhang, K. Li, M. Rehman, K. Mehmood, Y. Lan, Y. and J. Li. Epidemiological survey and phylogenetic characterization of Cysticercus tenuicollis isolated from Tibetan pigs in Tibet, China. BioMed Research International, 2017.
[13] A. Scala, A. Pipia, F. Dore, G. Sanna, C. Tamponi, R. Marrosu, and A. Varcasia, Epidemiological updates, and economic losses due to Taenia hydatigena in sheep from Sardinia, Italy. Parasitology Research, 114, 2015, 3137-3143.
[14] Taif Khalid Shakir,Ahmed N. Al Masri, Single Valued Neutrosophic Set for Selection of Water Supply in Intelligent Farming, Journal of International Journal of Advances in Applied Computational Intelligence, Vol. 2 , No. 2 , (2022) : 37-44 (Doi : https://doi.org/10.54216/IJAACI.020204)
[15] Jdid, M., & Smarandache, F. (2024). Finding a Basic Feasible Solution for Neutrosophic Linear Programming Models: Case Studies, Analysis, and Improvements. Neutrosophic Systems With Applications, 14, 30-37. https://doi.org/10.61356/j.nswa.2024.130
[16] Abouhawwash, M. and Jameel, M. (2023) "Evaluation Factors of Solar Power Plants to Reduce Cost Under Neutrosophic Multi-Criteria Decision Making Model", Sustainable Machine Intelligence Journal, 2, pp. (1):1–11. doi:10.61185/SMIJ.2023.22101.
[17] Estupiñan, J. Leyva, M. (2022). Neutrosophic Multicriteria Methods for the Selection of Sustainable Alternative Materials in Concrete Design. Journal of American Journal of Business and Operations Research, 6(2), 28-38.
[18] Guijarro-Rodríguez, A. A., Cevallos-Torres, L. J., Botto-Tobar, M., Leyva-Vazquez, M., & Holguin, J. Y. (2017). Clinical assessment using an algorithm based on fuzzy C-means clustering. In Technologies and Innovation: Third International Conference, CITI 2017, Guayaquil, Ecuador, October 24-27, 2017, Proceedings 3 (pp. 181-193). Springer International Publishing.
[19] Granja Huacon, S. H., Nivela Ortega, E. S., & Enríquez Palma, F. R. (2023). Análisis de clustering neutrosófico sobre el impacto de la igualdad y no discriminación en la contratación laboral en Babahoyo durante el año 2022. Neutrosophic Computing and Machine Learning, 27, 33-38.
[20] Ahmed N. Al-Masri , Manal Nasir, A Novel Fuzzy Clustering with Metaheuristic based Resource Provisioning Technique in Cloud Environment, Fusion: Practice and Applications, Vol. 6 , No. 1 , (2021) : 08-16 (Doi : https://doi.org/10.54216/FPA.060102)
[21] Liu, Z., Qiu, H., & Letchmunan, S. (2024). Self-adaptive attribute weighted neutrosophic c-means clustering for biomedical applications. Alexandria Engineering Journal, 96, 42-57.
[22] Ricardo, J. E., Domínguez Menéndez, J. J., Barcos Arias, I. F., Macías Bermúdez, J. M., & Lemus, N. M. (2021). Neutrosophic K-means for the analysis of earthquake data in Ecuador. Neutrosophic Sets & Systems, 44.
[23] Sakyi, S. A., Effah, A., Naturinda, E., Senu, E., Opoku, S., Amoani, B., ... & Danquah, K. O. (2023). Comparison of Modified Manual Acid-Phenol Chloroform Method and Commercial RNA Extraction Kits for Resource Limited Laboratories. International Journal of Clinical Practice, 2023(1), 9593796.
[24] Shumuye, N. A., Li, L., Li, W. H., Zhang, N. Z., Wu, Y. T., Wu, Y. D., ... & Jia, W. Z. (2023). Infection of sheep by Echinococcus multilocularis in Gansu, China: evidence from mitochondrial and nuclear DNA analysis. Infectious Diseases of Poverty, 12(1), 72.

Received: February 18, 2024. **Accepted:** June 10, 2024