5-2010

A Predictive Model for Secondary RNA Structure Using Graph Theory and a Neural Network.

Denise Renee Koessler
East Tennessee State University

A Predictive Model for Secondary RNA Structure Using Graph Theory and a Neural Network

---

A thesis

presented to

the faculty of the Department of Mathematics

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Master of Science in Mathematical Sciences

---

by

Denise Koessler

May 2010

Teresa Haynes, Ph.D., Co-Chair

Debra Knisley, Ph.D., Co-Chair

Robert A. Beeler, Ph.D.

Keywords: Graph Theory, Vertex Identification, Neural Network, RNA

ABSTRACT

A Predictive Model for Secondary RNA Structure

Using Graph Theory and a Neural Network

by

Denise Koessler

In this work we use a graph-theoretic representation of secondary RNA structure found in the database RAG: RNA-As-Graphs. We model the bonding of two RNA secondary structures to form a larger structure with a graph operation called merge. The resulting data from each tree merge operation is summarized and represented by a vector. We use these vectors as input values for a neural network and train the network to recognize a tree as RNA-like or not based on the merge data vector.

The network correctly assigned a high probability of RNA-likeness to trees identified as RNA-like in the RAG database, and a low probability of RNA-likeness to those classified as not RNA-like in the RAG database. We then used the neural network to predict the RNA-likeness of all the trees of order 9. The use of a graph operation to theoretically describe the bonding of secondary RNA is novel.

ACKNOWLEDGMENTS

I would like to first thank my committee co-chairs, Dr. Teresa Haynes and Dr. Debra Knisley. My thanks go to Dr. Haynes for her infectious love of graph theory, and for her motivating approach towards mathematical research. I wish to thank Dr. Debra Knisley for conceiving this project and participating in its coordination and execution. I am grateful for Dr. Robert A. Beeler for serving on my committee, and for his insightful critique of the first draft. Also, this work would not have been possible without the assistance from Dr. Jeff Knisley and his expertise with artificial neural networks.

Further, I am grateful to the National Science Foundation, Division of Graduate Education, grant DGE-0742364 awarded to G. Anderson, for fellowship support. We are indebted to Tamar Schlick and her research group at NYU for the creation of the RAG database. I appreciate the efforts of undergraduate research participants Cade Herron and Jordan Shipley.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 BACKGROUND

Our understanding of the role of RNA has changed and continues to be redefined. The role of non-coding RNA in gene regulatory networks places the study of RNA in the forefront of efforts to understand the complexities in Systems Biology.

1.1 Modeling RNA

RNA structure is divided into three classes: primary, secondary and tertiary. The RNA sequence is the primary structure. The primary structure forms the secondary by folding back onto itself. When this folding occurs, it forms Watson-Crick base pairs with intervening unpaired regions [20]. Unlike secondary protein structures where the elements of the primary sequence (amino acids) form bonds with nearby amino acids, the secondary structure of RNA includes the bonding of nucleotides at opposite ends of the sequence. For example, given the primary RNA sequence AGCGUCACAC-CCGCGGGGUAAACGCU, its secondary structure will include the Watson-Crick pairing A-U, G-C, and C-G between the first three nucleotide bases AGC and the final three bases GCU [20]. These regions occur in four types of structures known as hairpins, bulges, internal loops and junctions. Paired regions connecting these are usually referred to as stems. Secondary RNA structure can be represented by a two-dimensional drawing [20].

In this paper, we begin with theoretical approaches to describe secondary RNA structure to aid in the prediction of novel RNA structures. Because secondary RNA is represented by a two-dimensional schematic, graph theory nicely lends itself as a modeling tool for secondary RNA structure. The basic skeletal structure of secondary RNA is captured by representing the stems as edges of the graph and the regions with

unpaired bases as vertices. The resulting graph is a graph known as tree graph, or simply a tree. The RNA trees used in this work were first developed by Le et al. in [11] and Morosetti [1] to determine structural similarities in RNA.

The RAG database (RNA-As-Graphs) uses this representation and catalogs all tree structures for trees up to order 11. Every novel RNA structure is mapped onto a 2D tree model, and then cataloged according to two numerical values: the number of vertices (n) in the tree and a numerical representation (z) of its topological complexity. To quantitatively organize and archive all possible RNA tree graph representations, it is necessary to first generate the collection of all possible trees for a given number of vertices (n). For example, for the set $n = \{2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9\}$, there are a total of $\{1,\ 1,\ 2,\ 3,\ 6,\ 11,\ 23,\ 47\}$ distinct trees, respectively. Once all possible tree motifs were generated, Schlick et al. [14] generated the Laplacian matrix representation and calculated the eigenvalue spectrum $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ for each tree. According to [14], the second eigenvalue λ_2 measures a motif's topological complexity. For example, a more linear tree graph has a lower λ_2 value while a highly-branched tree graph has a high λ_2 value. This work catalogs each potential RNA motif by (n, λ_2). For easy reference, each RNA motif has a specific index $(n.z)$, where z represents an integer corresponding to the λ_2 ranking.

The trees with 2 through 8 vertices have been fully classified as known (verified), candidate, or non-candidate. The research compiled in [3] organizes all known, candidate, and non-candidate RNA trees of order 8 or less by a color coding scheme. Red trees represent known RNA, blue trees are candidate RNA, and black trees are non-candidate trees. A tree that is either a known tree or candidate tree is referred to as an RNA-like tree and a non-candidate tree is referred to as not-RNA-like. The

remaining trees on 9 or higher vertices have not been grouped into these three categories. This catalog of RNA trees is intended as a tool for searching existing RNAs and to stimulate the search for candidate RNA motifs not yet discovered in nature or a laboratory.

There are a number of secondary RNA structure prediction algorithms available online such as Zuckers MFold [12] and Vienna RFold [8]. Given the primary RNA sequence, the web server will return a list of predicted secondary folds. These programs are based on finding the secondary structure with the total lowest free energy by calculating the free energy of a number of base-pairing schemes and returning the lowest energy potential secondary structure as the most probable [8, 12]. In the majority of cases, even for long sequences, the predicted structure is a structure whose tree representation is a small ordered tree (a tree with fewer than 10 vertices). However, there are secondary RNA structures whose tree representation is a tree with 10 or more vertices. For example, the 5S ribosomal RNA *Clavibacter michiganensis* (RNA Database ID S73542) has a 10 vertex tree representation.

1.2 Research Description

In this work we consider the possibility that a larger secondary RNA structure is formed by the bonding of two smaller secondary RNA structures. We model this bonding process by defining a graph merge that occurs on the vertices of the trees. If our hypothesis is valid, then larger secondary RNA structures should arise from trees that are unique to secondary RNA structure, and not from arbitrary trees. That is, only trees that represent RNA, and hence are thermodynamically stable structures, can be used to produce a tree which is still stable. We test this hypothesis and find,

under specified constraints, stable trees are produced by merging two stable trees. Furthermore, by applying a predictive model, we find that some of the trees in the RAG database that are listed as candidate RNA structures are not clearly RNA-like in structure by our method. Our approach is novel, and may be considered as a valuable tool for refining prediction algorithms. It also illustrates the applicability of graphs as models, not only for secondary RNA, but for biomolecules in general. In order to formalize this idea, we introduce the graph-theoretic terminology and concepts.

1.3 Basic Terminology of Graph Theory

In graph theory, trees have been heavily studied both for application purposes and for theoretical investigations. As defined in [3], a graph $G = (V(G), E(G))$ is a nonempty, finite set of elements called *vertices* together with a (possibly empty) set of unordered pairs of distinct vertices of G called *edges*. The *vertex set* of G is denoted by $V(G)$, and the edge set of G is denoted by $E(G)$. Here we consider only simple graphs, that is, graphs with no loops or multiple edges. A *tree* is commonly defined as a connected graph with the property that no two vertices lie on a cycle. These two properties, connected and acyclic, completely characterize a tree since the removal of any edge will disconnect the graph, and the addition of any edge will create a cycle. Further, this implies that any tree with n vertices contains $n - 1$ edges.

An *isomorphism of graphs G and H* is a bijection between the vertex sets of G and H, $f : V(G) \to V(H)$, such that any two vertices v and w of G are adjacent in G if and only if $f(v)$ and $f(w)$ are adjacent in H. This kind of bijection is commonly called an *edge-preserving bijection* or a *structure-preserving bijection*. If an isomorphism exists

between two graphs, then we say the graphs are *isomorphic* and we write $G \simeq H$. To illustrate these terms, Figure 1 displays two isomorphic trees. Figure 2 shows the six non-isomorphic trees of order 6. Figure 3 shows the index value and color codes of the six trees on 6 vertices as shown in [15].

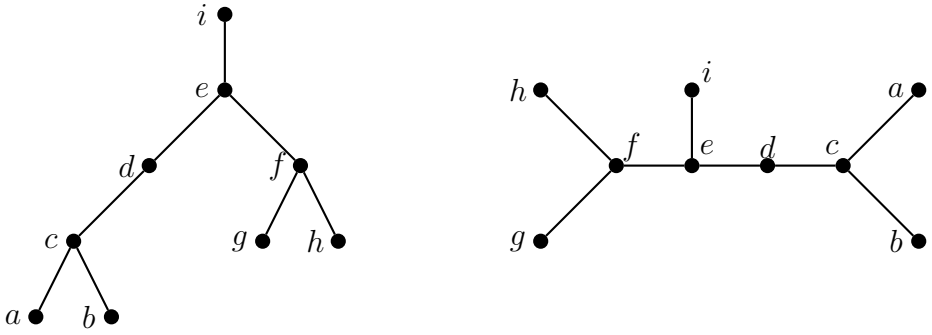

Figure 1: An Example of Two Isomorphic Trees

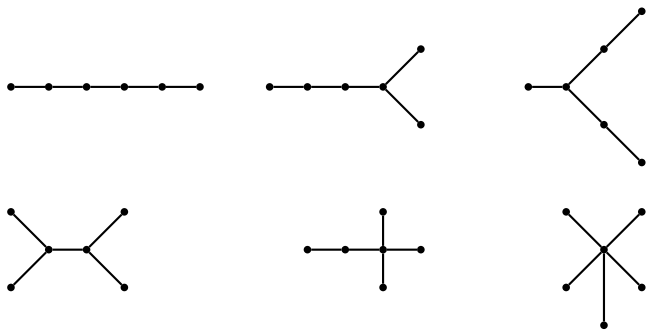

Figure 2: The Six Non-Isomorphic Trees of Order 6

Two vertices joined by an edge are said to be neighbors and the degree of a vertex v in a graph G, denoted by $deg_G(v)$, is the number of neighbors of v in G. A vertex of degree one is called a *leaf*, and its neighbor is called a *support vertex*. For use in this paper, a vertex v in a tree T is an *internal vertex* if it is neither a leaf or support vertex.

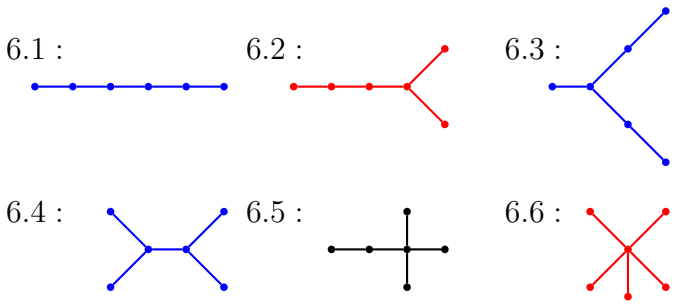

Figure 3: The Index Value and Color Category for Trees of Order 6 from RAG

Two vertices u and v are said to be *identified* if they are combined into a single vertex whose neighborhood is the union of the neighborhoods of u and v. The binary operation *merge* of two graphs G_1 and G_2 forms a new graph G_{uv} by identifying a vertex u in G_1 with a vertex v in G_2. Figure4 demonstrates vertex identification at the colored vertices for the pictured trees.

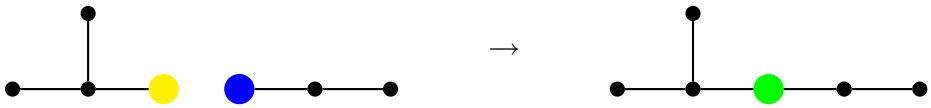

Figure 4: An Example of a Merge of Two Trees

2 RESULTS AND DISCUSSION

We considered the possibility that a larger secondary RNA structure could be formed by the bonding of two smaller secondary RNA structures. We modeled this bonding process by defining a merge operation on two trees. In this research, we determined all possible tree merges which result in a tree with 9 or fewer vertices. We used the RNA online-database RAG and the tree color code developed by Schlick et al. in [14] and discussed in the introduction. Recall that red trees are RNA-like (known), blue trees are RNA-like (candidates) and black trees are not RNA-like (non-candidates). Note that in a tree model of a secondary RNA structure, a hairpin corresponds to a vertex of degree one, internal loops and bulges are vertices of degree two, and junctions correspond to vertices of degree three or more.

Initially, we hypothesized that the color of the merging trees would be indicative of the color of the result tree. However, we found that this is not necessarily the case. Our hypothesis held when merging RNA-like tree motifs at vertices of degree one (hairpins) or degree two (bulges or internal loops). That is, when identifying vertices of degree two or less, almost all red to red or red to blue tree merges produced a red tree. However, this was not always the case when the vertices being identified included a vertex of high degree (junction).

Using these findings, we trained a neural network to recognize the known classification of a tree as RNA-like or not RNA-like in structure. The network assigned a value between 0 and 1 to classify these trees. Table 1 shows the interval values used to classify the trees as RNA-like or not-RNA-like. Paralleling the work completed in [10], our artificial neural network was trained on two classes of trees: the known RNA (red) trees and the non-candidate (black) trees. There are 15 red trees of order 7, 8

and 9 along with 11 black trees of order 7 and 8.

Table 1: The Key for Categorizing the Neural Network Prediction Values

ANN Value	Resulting Category
1.0 - 0.80:	Highly-RNA-Like
$0.7\overline{9}$ - 0.60	RNA-Like
$0.5\overline{9}$ - 0.40	Unclassifiable
$0.3\overline{9}$ - 0.20	Not-RNA-Like
$0.1\overline{9}$ - 0.0	Highly Not-RNA-Like

2.1 Predictions for the Classified RAG Trees

The MLP artificial neural network correctly predicted 100% of the known RAG trees to have a value greater than 0.50. Further, the network correctly calculated a prediction value below 0.50 for nine of the eleven black trees. However, two non-candidate RAG trees, indexed as 7.10 and 8.14, received an MLP prediction value between 0.60 and 0.50. Therefore, we labeled all trees with an MLP prediction value within the range 0.59 to .40 as "Unclassifiable". Table 2 displays the RAG classification and corresponding predicted class for each of the classified 26 trees on 7, 8 or 9 vertices.

2.2 Predictions for the Unclassified RAG Trees

After using the MLP to predict the classified RAG trees, we calculated the prediction value for the 43 unclassified trees on 9 vertices from the RAG online web database. For these 43 trees, the MLP predicted a total of 22 trees to represent RNA motifs: 18 trees were highly-RNA-like and four were only RNA-like. Further, there were 14 trees which the artificial neural network predicted to not represent RNA sec-

Table 2: The Prediction Values for the Classified RAG Trees

RAG Index	Color Class	ANN Prediction	ANN Result	RAG Index	Color Class	ANN Prediction	ANN Result
7.1	Red	1.00000	Highly RNA-Like	8.19	Black	0.33309	Not-RNA-Like
7.10	Black	0.59091	Unclassifiable	8.20	Red	0.69128	RNA-Like
7.11	Black	0.00045	Highly Not-RNA-Like	8.21	Black	0.00747	Highly Not-RNA-Like
7.2	Red	0.99860	Highly RNA-Like	8.22	Black	0.00437	Highly Not-RNA-Like
7.3	Red	0.99990	Highly RNA-Like	8.23	Black	0.00001	Highly Not-RNA-Like
7.6	Red	0.99721	Highly RNA-Like	8.3	Red	0.99796	Highly RNA-Like
7.9	Black	0.43130	Unclassifiable	8.5	Red	0.99991	Highly RNA-Like
8.10	Red	0.99994	Highly RNA-Like	8.7	Red	0.99815	Highly RNA-Like
8.11	Red	0.99682	Highly RNA-Like	8.9	Black	0.75935	RNA-Like
8.14	Black	0.52998	Unclassifiable	9.6	Red	0.99991	Highly RNA-Like
8.15	Red	0.56343	Unclassifiable	9.11	Red	0.99993	Highly RNA-Like
8.17	Black	0.36524	Not-RNA-Like	9.13	Red	0.99740	Highly RNA-Like
8.18	Black	0.00595	Highly Not-RNA-Like	9.27	Red	0.99795	Highly RNA-Like

ondary structure: 10 trees were highly not-RNA-like with four trees grouping into the not-RNA-like category. Overall, the MLP calculated an unclassifiable value for seven of the 43 trees. These values are listed in Table 3.

Table 3: The Prediction Values for the Unclassified RAG Trees

RAG Index	Color Class	ANN Prediction	ANN Result	RAG Index	Color Class	ANN Prediction	ANN Result
9.1	Unknown	1.00000	Highly RNA-Like	9.32	Unknown	0.00464	Highly Not-RNA-Like
9.10	Unknown	0.99874	Highly RNA-Like	9.33	Unknown	0.99897	Highly RNA-Like
9.12	Unknown	0.99724	Highly RNA-Like	9.34	Unknown	0.81259	Highly RNA-Like
9.14	Unknown	0.99722	Highly RNA-Like	9.35	Unknown	0.25818	Not-RNA-Like
9.15	Unknown	0.48019	Unclassifiable	9.36	Unknown	0.39903	Not-RNA-Like
9.16	Unknown	0.58718	Unclassifiable	9.37	Unknown	0.11790	Highly Not-RNA-Like
9.17	Unknown	0.99660	Highly RNA-Like	9.38	Unknown	0.57340	Unclassifiable
9.18	Unknown	0.99993	Highly RNA-Like	9.39	Unknown	0.00576	Highly Not-RNA-Like
9.19	Unknown	0.81428	Highly RNA-Like	9.4	Unknown	0.99840	Highly RNA-Like
9.2	Unknown	0.99888	Highly RNA-Like	9.40	Unknown	0.00238	Highly Not-RNA-Like
9.20	Unknown	0.58342	Unclassifiable	9.41	Unknown	0.00025	Highly Not-RNA-Like
9.21	Unknown	0.64473	RNA-Like	9.42	Unknown	0.69128	RNA-Like
9.22	Unknown	0.28250	Not-RNA-Like	9.43	Unknown	0.00756	Highly Not-RNA-Like
9.23	Unknown	0.84883	Highly RNA-Like	9.44	Unknown	0.40451	Unclassifiable
9.24	Unknown	0.99705	Highly RNA-Like	9.45	Unknown	0.00441	Highly Not-RNA-Like
9.25	Unknown	0.62696	RNA-Like	9.46	Unknown	0.00434	Highly Not-RNA-Like
9.26	Unknown	0.99942	Highly RNA-Like	9.47	Unknown	0.00002	Highly Not-RNA-Like
9.28	Unknown	0.36067	Not-RNA-Like	9.5	Unknown	0.72341	RNA-Like
9.29	Unknown	0.83375	Highly RNA-Like	9.7	Unknown	0.99875	Highly RNA-Like
9.3	Unknown	0.99840	Highly RNA-Like	9.8	Unknown	0.99909	Highly RNA-Like
9.30	Unknown	0.52189	Unclassifiable	9.9	Unknown	0.55979	Unclassifiable
9.31	Unknown	0.00769	Highly Not-RNA-Like				

2.3 A Comparative Analysis

A predictive tool based on domination parameters was used in [10] to classify all the trees on 7, 8 and 9 vertices. Here we compare our results to the original tree categories determined in [14] to results found in [10] and [15]. Our comparison is summarized in Table 4.

Table 4: A Comparative Analysis of the Results from [10], [15] and this Paper

RAG Index	RAG Color Class	Domination Predicted Class	Bonding Predicted Class	RAG Index	RAG Color Class	Domination Predicted Class	Bonding Predicted Class
7.1	Red	RNA-Like	Highly-RNA-Like	8.17	Black	Not-RNA-Like	Not-RNA-Like
7.10	Black	Not-RNA-Like	Unclassifiable	8.18	Black	Not-RNA-Like	Highly-Not-RNA-Like
7.11	Black	Not-RNA-Like	Highly-Not-RNA-Like	8.19	Black	Not-RNA-Like	Not-RNA-Like
7.2	Red	RNA-Like	Highly-RNA-Like	8.20	Red	RNA-Like	RNA-Like
7.3	Red	RNA-Like	Highly-RNA-Like	8.21	Black	Not-RNA-Like	Highly-Not-RNA-Like
7.6	Red	RNA-Like	Highly-RNA-Like	8.22	Black	Not-RNA-Like	Highly-Not-RNA-Like
7.9	Black	Not-RNA-Like	Unclassifiable	8.23	Black	Not-RNA-Like	Highly-Not-RNA-Like
8.10	Red	RNA-Like	Highly-RNA-Like	8.3	Red	RNA-Like	Highly-RNA-Like
8.11	Red	RNA-Like	Highly-RNA-Like	8.5	Red	RNA-Like	Highly-RNA-Like
8.14	Black	Not-RNA-Like	Unclassifiable	8.7	Red	RNA-Like	Highly-RNA-Like
8.15	Red	RNA-Like	Unclassifiable	8.9	Black	Not-RNA-Like	RNA-Like
9.1	Unkwn	RNA-Like	Highly-RNA-Like	9.31	Unkwn	RNA-Like	Highly-Not-RNA-Like
9.10	Unkwn	RNA-Like	Highly-RNA-Like	9.32	Unkwn	Not-RNA-Like	Highly-Not-RNA-Like
9.11	Red	RNA-Like	Highly-RNA-Like	9.33	Unkwn	RNA-Like	Highly-RNA-Like
9.12	Unkwn	RNA-Like	Highly-RNA-Like	9.34	Unkwn	RNA-Like	Highly-RNA-Like
9.13	Red	RNA-Like	Highly-RNA-Like	9.35	Unkwn	Not-RNA-Like	Not-RNA-Like
9.14	Unkwn	RNA-Like	Highly-RNA-Like	9.36	Unkwn	Not-RNA-Like	Not-RNA-Like
9.15	Unkwn	Not-RNA-Like	Unclassifiable	9.37	Unkwn	Not-RNA-Like	Highly-Not-RNA-Like
9.16	Unkwn	RNA-Like	Unclassifiable	9.38	Unkwn	RNA-Like	Unclassifiable
9.17	Unkwn	RNA-Like	Highly-RNA-Like	9.39	Unkwn	Not-RNA-Like	Highly-Not-RNA-Like
9.18	Unkwn	RNA-Like	Highly-RNA-Like	9.4	Unkwn	RNA-Like	Highly-RNA-Like
9.19	Unkwn	RNA-Like	Highly-RNA-Like	9.40	Unkwn	Not-RNA-Like	Highly-Not-RNA-Like
9.2	Unkwn	RNA-Like	Highly-RNA-Like	9.41	Unkwn	Not-RNA-Like	Highly-Not-RNA-Like
9.20	Unkwn	RNA-Like	Unclassifiable	9.42	Unkwn	RNA-Like	RNA-Like
9.21	Unkwn	RNA-Like	RNA-Like	9.43	Unkwn	RNA-Like	Highly-Not-RNA-Like
9.22	Unkwn	RNA-Like	Not-RNA-Like	9.44	Unkwn	RNA-Like	Unclassifiable
9.23	Unkwn	Not-RNA-Like	Highly-RNA-Like	9.45	Unkwn	Not-RNA-Like	Highly-Not-RNA-Like
9.24	Unkwn	RNA-Like	Highly-RNA-Like	9.46	Unkwn	Not-RNA-Like	Highly-Not-RNA-Like
9.25	Unkwn	Not-RNA-Like	RNA-Like	9.47	Unkwn	Not-RNA-Like	Highly-Not-RNA-Like
9.26	Unkwn	RNA-Like	Highly-RNA-Like	9.5	Unkwn	RNA-Like	RNA-Like
9.27	Red	RNA-Like	Highly-RNA-Like	9.6	Red	RNA-Like	Highly-RNA-Like
9.28	Unkwn	Not-RNA-Like	Not-RNA-Like	9.7	Unkwn	RNA-Like	Highly-RNA-Like
9.29	Unkwn	RNA-Like	Highly-RNA-Like	9.8	Unkwn	RNA-Like	Highly-RNA-Like
9.3	Unkwn	RNA-Like	Highly-RNA-Like	9.9	Unkwn	Not-RNA-Like	Unclassifiable
9.30	Unkwn	RNA-Like	Unclassifiable				

All three studies agreed on the classification of nine of the eleven non-candidate

(black) tree graphs. The two exceptions are graphs 7.9 and 7.10, which our study found to be unclassifiable. Further, with the exception of tree 8.15, all three research studies concluded that all known (red) tree graphs on 7, 8 and 9 vertices were RNA like based on their respective structural calculations. The model in [10] predicted tree 8.15 to be RNA-like in structure, however, their predictive model reported the highest amount of error for the classifications of this tree. We calculated a 0.56 likelihood that tree 8.15 contains RNA-like structure. As a result, both predictive models were unable to confidently classify tree 8.15.

Most notably, the predictive model used in previous RNA motif research supports the major results of this paper. As seen in Table 4, the predictive model in [10] classified 29 of the 43 unknown RNA trees to be RNA-like. When examining these results, the authors of [10] felt their model over-predicted the class of RNA-like tree graphs. Accordingly, we found a total of 18 trees to be highly-RNA-like in structure. Of those, 17 of the 18 trees in the highly-RNA-like category from this study were included to be RNA-like from the results compiled by [10]. Consequently, the predictive model of our study narrowed the class of RNA-like motifs from previous findings.

Additionally, of the 12 trees on 9 vertices that we predicted to be not-RNA-like in structure from our model, previous findings agreed with 10 of those classifications. The model in [10] predicted trees 9.31 and 9.43 to be RNA-like, whereas we found both motifs to be highly not-RNA-like. From the other direction, of the 14 trees predicted in [10] to be not-RNA-like in structure, our prediction agreed with 12 of the 14. Trees 9.23 and 9.25 are both predicted to be not RNA-like in [10], but are classified as potential RNA structures in our study. Hence, our predictive model provided more descriptive information about the structural classification of the unknown RAG tree

motifs on 9 vertices than the findings from [10].

In summary, when we compared our results with those in [10], we found two improvements. First, our neural network outcomes were not solely RNA-like or not RNA-like. Rather, our model assigned a probability, which is a measurement of a tree's RNA likeness. Second, our model predicted fewer of the trees on 9 vertices to be RNA-like, and thus seemed to be a more discriminating predictive tool.

3 METHODS

We use graph theory to model the bonding of secondary RNA structures and a predictive neural network to quantify our results.

3.1 Graph Theoretic Model

The binary operation *merge* of two trees T_1 and T_2 forms a new larger tree T_{uv} by identifying a vertex u in T_1 with a vertex v in T_2. Merging two trees of n and m vertices produces nm total trees, some of which can be isomorphic, and each resulting tree has a total of $n + m - 1$ vertices.

To accurately model RNA bonding, we must consider all possible vertex identifications between two RNA tree models. For example, there are 12 possible vertex identifications for merging trees 3.1 and 4.2. Of these 12 merges, the four non-isomorphic trees are shown in Figure 5. Figure 6 displays the official RAG identification and color classes for the trees from Figure 5.

Our research determined all possible merges forming trees on 9 and fewer vertices. When tracking the information from all possible vertex identifications between two trees, the resulting trees were noted and their frequencies counted. For example, in Figure 6, the merge of tree 3.1 and 4.2 resulted in the following tree set: $3.1 + 4.2 = \{6.2,\ 6.4,\ 6.5,\ 6.6\}$. Trees 6.2, 6.4, 6.5 and 6.6 occurred with frequencies 6, 3, 2, and 1 respectively. Additionally, we noted the type and degree of the vertices at each merge. Table 5 displays all the information for the vertex identifications of the merges between trees 3.1 and 4.2.

For all 94 graphs on 2 through 9 vertices, every possible vertex identification resulting in a graph on 9 or fewer vertices was calculated and recorded. Table A

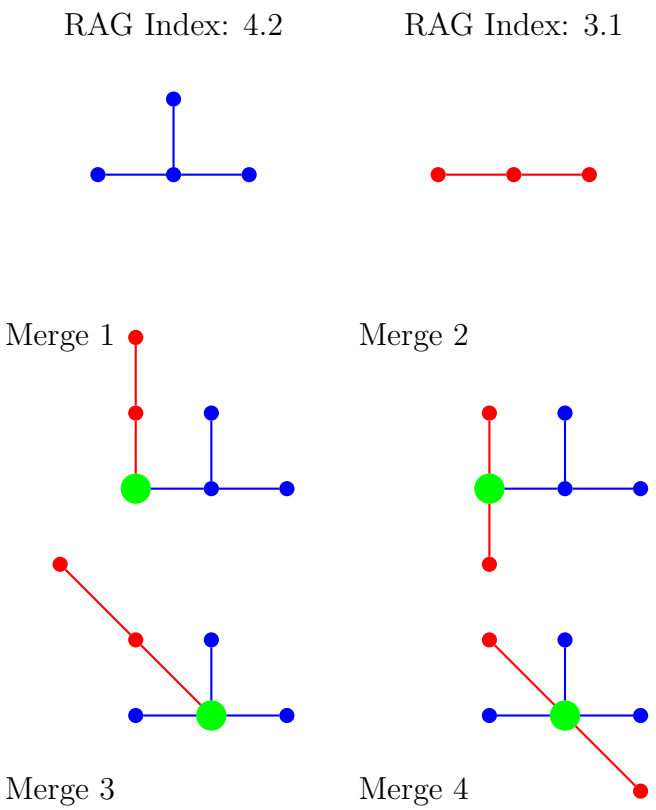

Figure 5: The Four Non-Isomorphic Resulting Trees when Merging Trees 3.1 and 4.2

Table 5: The Data Table Produced from the Tree Merge Between RAG Trees 3.1 and 4.2

3.1 and 4.2	$v \in V(3.1)$		$v \in V(4.2)$		Results		Total
Merge:	$deg(v)$	Type	$deg(u)$	Type	Graph	Color	Graphs
1	1	Leaf	1	Leaf	6.2	Red	6
2	2	Support	1	Leaf	6.4	Blue	3
3	1	Leaf	3	Support	6.5	Black	2
4	2	Support	3	Support	6.6	Red	1

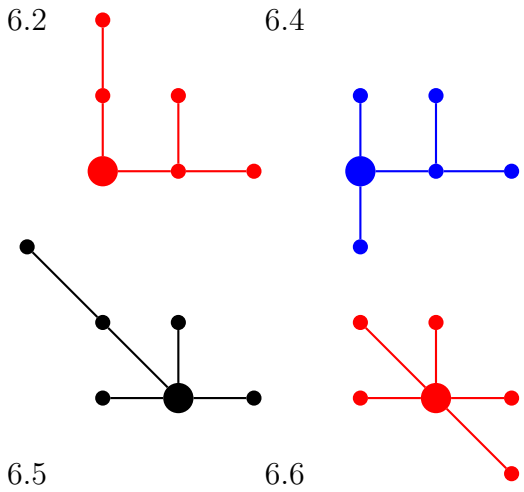

Figure 6: The RAG Identification and Color Classes for the Trees from Figure 5

(see appendix) displays the vertex identification results for all tree merges. Information from Table A was translated into data vectors. Each data vector displayed the composition information for the result tree in the following manner:

$$[\langle c_1, c_2, deg(v_1), deg(v_2)\rangle, \langle y_1, y_2\rangle], \text{ where for } i \in \{1, 2\},$$

$$c_i = \begin{cases} 1, & \text{if } T_i \text{ is red or blue} \\ 0, & T_i \text{ is black,} \end{cases}$$

$deg(v_i)$ is the degree of the identified vertex of T_i, and

$y_1 = 1$ and $y_2 = 0$ if the result tree is an RNA-like tree, and

$y_1 = 0$ and $y_2 = 1$ if the result tree is not RNA-like.

3.2 An Artificial Neural Network

In their numerical form as data vectors, the vertex identification results were used to predict the RNA-like status of the 43 unclassified trees on 9 vertices. The data

vectors from the 15 known (red) tree graphs on 7, 8 and 9 vertices along with the data vectors of the 11 non-candidate (black) tree graphs on 7 and 8 vertices made up the training data. To check the validity of our model, we then predicted the status of the 26 known tree classifications on 7, 8 and 9 vertices. Then, the model was used as a predictive tool for the 43 unclassified trees of order nine. This research paralleled previous work by the authors in [10]. In this section, we discuss the details of designing, training and using an artificial neural network as a prediction tool.

3.2.1 Description

Following the network created in [10], our approach was to train a multi-layer perceptron (MLP) artificial neural network using a standard back-propagation algorithm. Results from a back-propagation MLP can be independently reproduced by other researchers and can also provide information beyond simple predictions [10].

The 3-layer MLP was used to predict the RNA-like status of the trees. The first layer, or *input layer*, contained four perceptrons corresponding to the data vector from one vertex identification of the complete merge between two trees. The last layer, or *output layer*, consisted of two perceptrons with activations y_1 and y_2, where $y_1 = 1$ and $y_2 = 0$ if the result tree, which corresponded to the input data vector, was predicted to be an RNA tree and where $y_1 = 0$ and $y_2 = 1$ if the result tree was not RNA-like. The middle layer, or *hidden layer*, contained 24 perceptrons. The weights between the input and hidden layers were denoted by w_{jk} and the weights between the hidden and output layers were denoted by α_{ij}.

3.2.2 Implementation

The data vectors from the vertex identifications of the 26 trees on 7, 8 or 9 vertices that either are an RNA tree or not an RNA-like tree determined the training set

$$TS = \{(p^i, q^i)\}_{i=1}^{26}$$

where $p^i = \langle p_1^i,\ p_2^i,\ p_3^i,\ p_4^i \rangle$ is the data vector, $q^i = \langle 1, 0 \rangle$ if the tree is known or predicted to be an RNA tree, and $q^i = \langle 0, 1 \rangle$ if the tree is not RNA-like. The back-propagation algorithm is used to implement a gradient following minimization of the total squared error

$$E = \frac{1}{2} \sum_{i=1}^{26} ||y(p^i) - q^i||^2$$

where $y(p^i) = \langle y_1(p^i), y_2(p^i) \rangle$ is the output due to an input of p^i and the norm is generated by the corresponding dot product.

The weights were initially assigned random values close to 0. Then, for each pair (p^i, q^i), the weights α_{jk} were adjusted using

$$\alpha_{jk} \rightarrow \alpha_{jk} + \lambda \delta_j \xi_k$$

where $\xi_k = \sigma\left(\sum \omega_{kj} p_j^i - \theta_k \right)$, where $\lambda > 0$ is a fixed parameter called the *learning rate,* and where

$$\delta_j = y_j(1 - y_1)(q_j^i - y_j).$$

The weights ω_{kr} were adjusted using

$$\omega_{kl} \rightarrow \omega_{kl} + \lambda p_i^l \xi_k (1 - \xi_k) \sum_{j=1}^{2} \alpha jk \delta_j$$

In each training session, the patterns were randomly permuted to avoid bias, and training continued until E was sufficiently close to 0 [2].

The MLP artificial neural network was trained and tested by *predicting complements*. During this procedure, the vertex identification data vectors of the 26 classified tree motifs were randomly partitioned into a training set and a complement set. Predicting complements was performed with 15% of the data vectors in the complement set for each trial.

The network was trained using the data *not in* the complement until the total squared error was close to 0 (approximately 10,000 iterations for each of the 5 repetitions of the classifier experiment). Once the network had been trained, it was used to predict the classification of the data in the complement set. This is known as leave-v-out cross-validation. According to [18], cross-validation is a reliable measure of the generalization error of the network when the training set is not too large. In each of the 5 repetitions of the classifier experiment, the root-mean-square error for the complement predictions was less than 5% of the class value of 1 (i.e., below 0.05).

In order to most accurately utilize our data, each tree's final classification was calculated as an average of a linear combination of prediction values from the vertex identifications. To do so, we began this procedure by using the MLP to predict the value for each vertex identification for a given tree. Then, this value was multiplied by a weight which refers to the total number of graph isomorphisms for the vertex identification. This weight was noted for each identification and can be referenced in the "Total Graphs" column of Table A (see appendix). To normalize the result, the linear combination of all the vertex identification values was divided by the sum of the weights. This final average determined the prediction value for the tree. Table 6 outlines this procedure for tree 7.9.

Table 6: An Example of the Algorithm used to Determine the Prediction Value for Tree 7.9

$a[7.9] := 4 * RNANet : -Classify(\langle 1,1,1,3\rangle);$	$a_{[7.9]} := \langle 0.95945,\ 3.52754\rangle$
$b[7.9] := 2 * RNANet : -Classify(\langle 1,0,1,2\rangle);$	$b_{[7.9]} := \langle 0.00030,\ 1.99985\rangle$
$c[7.9] := 1 * RNANet : -Classify(\langle 1,1,2,2\rangle);$	$c_{[7.9]} := \langle 0.97666,\ 0.02253\rangle$
$d[7.9] := 4 * RNANet : -Classify(\langle 1,1,2,1\rangle);$	$d_{[7.9]} := \langle 3.95652,\ 0.05185\rangle$
$e[7.9] := 6 * RNANet : -Classify(\langle 1,1,1,3\rangle);$	$e_{[7.9]} := \langle 1.43917,\ 5.29130\rangle$
$Class[7.9] := \frac{a[7.9]+b[7.9]+c[7.9]+d[7.9]+e[7.9]}{17}$	$Class_{[7.9]} := \langle 0.43130,\ 0.64077\rangle$

Once the MLP was fully trained, the network was used to predict the classification of the 26 red or black RAG trees on 7, 8 and 9 vertices. For these trees, the final MLP prediction values ranged from 1.0 to 0.0. As a result, the key in Table 1 uses a range to classify the final values.

4 CONCLUSIONS

Using a tree representation of secondary RNA structure, we modeled the creation of a larger structure from the bonding of two smaller structures by considering all combinatorial possibilities. We modeled the bonding with a graph operation called (vertex) merge. Data from this process included information on the degrees of the vertices identified in the merge and the classification of substructures. We created data vectors and then utilized these data vectors from known RNA trees on 7, 8 and 9 vertices together with the data vectors from the non-candidate RNA trees on 7 and 8 vertices to create and train a neural network to recognize a tree as RNA-like or not-RNA-like in structure. We applied this predictive tool to categorize known RNA classifications and to predict unknown RNA trees.

The results for the 15 red trees of orders seven, eight, and nine agreed with the classifications from the RAG database and previous research in [10]. Further, our neural network correctly classified 9 of the 11 black, or non-candidate, trees on 7 and 8 vertices to agree with previous research. However, the authors of [10] felt their model over-predicted the class of RNA-like trees for those 43 unclassified trees on 9 vertices. Their results classified 29 of the unknown trees as RNA-like in structure, and 14 as not-RNA-like. As a result, our study narrowed down the class of highly-RNA-like tree structures on 9 vertices from the 29 predicted in [10] to 18 according to the values calculated by the MLP artificial neural network.

We revealed that graphical operations from the field of mathematical graph theory can successfully model secondary RNA motifs. Further, we demonstrated that these numerical values from these operations can enable the training of an artificial neural network to recognize the difference between likely and unlikely RNA struc-

tures. These findings, along with those from previous predictive models, exhibited the power of mathematical graph theory as an effective modeling method. By representing bimolecular structures with graph theory, modern researchers can enter an extensive and unexplored field of quantitative biology.

Although trees have previously been used to model secondary RNA structure, the applications of techniques from graph theory have been limited. There are numerous binary operations on graphs, such as the Cartesian product and graph join. As far as we know, our application of graph merge to model RNA binding is novel and has proven to be a valuable tool.

As a follow up to this study, future research could combine the data from [10, 14] and this paper to create a more powerful predictive model. A more intelligent artificial neural network, or another predictive tool, could utilize all three sets of data to predict the classifications for all the RAG trees. Additionally, future projects could examine the effect of other graphical invariants and operations on the structural properties of the RAG motifs. Another potential research project could be to use the ideas of our research, and those from [10], to examine the structural components of the unclassified RAG trees on 10 vertices from [15].

BIBLIOGRAPHY

[1] G. Benedetti and S. Morosetti, A Graph-Topological Approach to Recognition of Pattern and Similarity in RNA Secondary Structures, *Biol Chem.* **22** (1996) 179-184.

[2] N. Bose and P. Liang, *Neural Network Fundamentals with Graphs, Algorithms, and Applications,* New York: McGraw-Hill (1996).

[3] G. Chartrand, *Introductory Graph Theory*, Boston: Prindle, Weber & Schmitt (1977).

[4] H. Gan, S. Pasquali, and T. Schlick, Exploring the Repertoire of RNA Secondary Motifs Using Graph Theory: Implications for RNA Design, *Nucleic Acids Research,* **31** (2003) 2926-2943.

[5] F. Harary and G. Prings, The Number of Homeomorphically Irreducible Trees and Other Species, *Acta Math,* **101** (1959) 141-162.

[6] T. Haynes, S. Hedetniemi, and P. Slater, *Fundamentals of Domination in Graphs*, New York: Marcel Dekker, Inc. (1998).

[7] D. Koessler, D. Knisley, J. Knisley, and T. Haynes, A predictive model for secondary RNA structure using graph theory and a neural network, *BMC Bioinformatics* (Submitted March 2010.)

[8] I. Hofacker, Vienna RNA Package: RNA Secondary Structure Prediction and Comparison, [http://www.tbi.univie.ac.at/RNA/], (2007). February 2010.

[9] D. Knisley, T. Haynes T, E. Seier, and Y. Zou, A Quantitative Analysis of Secondary RNA Structure Using Domination Based Parameters on Trees, *BMC Bioinformatics* **7** (2006) 108.

[10] D. Knisley D, T. Haynes T, and J. Knisley, Using a Neural Network to Identify Secondary RNA Structures Quantified by Graphical Invariants, *Communications in Mathematical and in Computer Chemistry / MATCH,* **60** (2008) 277-290.

[11] S. Le, R. Nussinov, and J. Maziel, Tree Graphs of RNA Secondary Structures and their Comparison, *Comp. Biomed. Res.,* **22** (1989) 461-473.

[12] N. Markham, The Rensselaer bioinformatics web server, [http://mfold.bioinfo.rpi.edu/], (2005). February 2010.

[13] M. Nelson and S. Istrail, RNA Structure and Prediction Computational Molecular Biology, [http://tuvalu.santafe.edu/ pth/rna.html], (1996). August 2009.

[14] T. Schlick, D. Fera, N. Kim, N. Shiffeldrim, J. Zorn, U. Laserson, and H. Gan, RAG: RNA-As-Graphs Web Resource, *BMC Bioinformatics,* **5** (2004) 88.

[15] T. Schlick, D. Fera, N. Kim, N. Shiffeldrim, J. Zorn, U. Laserson, and H. Gan, RAG: RNA as Graphs Web Resource, [http://monod.biomath.nyu.edu/rna/rna.php], (2007). February 2010.

[16] T. Schlick, D. Fera, N. Kim, N. Shiffeldrim, J. Zorn, U. Laserson, H. Gan, and M. Tang, RAG: RNA-As-Graphs Database-Concepts, Analysis, and Features, *Bioinformatics,* **2**0 (2004) 1285-1291.

[17] T. Schlick, N. Kim, N. Shiffeldrim, and H. Gan, Candidates for Novel RNA Topologies, *J Mol Biol,* **341** (2004) 1129-1144.

[18] J. Shao, Linear model selection by cross-validation, *J. Am. Statistical Association,* **88** (1993) 486-494.

[19] S. Vishveshwara, K. V. Brinda, and N. Kannan, Protein Structure: Insights from Graph Theory, *Journal of Theoretical and Computational Chemistry,* **1** (2002) 187-211.

[20] S. Waterman, *Introduction to Computational Biology*, Boca Raton: Chapman & Hall (1995).

APPENDICES

Appendix A: Literature Review

RAG: RNA-As-Graphs Web Resource

Unlike the previous RNA databases which have focused on archiving known RNA primary, secondary and tertiary structures, T. Schlick, D. Fera, N. Kim, N. Shiffeldrim, J. Zorn, U. Laserson, and H. Gan developed an approach in [14] for cataloging and classifying all possible RNA structures. Their methods are based on the topological properties of RNA secondary motifs. The secondary structures of RNA, such as bulges, loops, junctions, and stems, strongly correlate with their functional properties. As a result, classifying RNA by their secondary topologies aids in the identification of new RNAs and stimulates the search for candidate RNA motifs not yet discovered in nature or a laboratory.

The RNA-As-Graphs (RAG) database models existing, candidate, and non-candidate RNA structures using graph theory. As shown in Figure 7, when converting existing RNA secondary representations to tree graphs, Schlick et al. [14] followed the following rules: (1) A bulge, hairpin loop, or internal loop is considered a vertex ($\bullet$) when there is more than one unmatched nucleotide or non-complementary base pair; (2) A junction, the location where three or more stems meet, is a vertex; (3) The $3'$ and $5'$ ends of a helical stem are considered a vertex; (4) An RNA stem with more than one complementary base pair is represented as an edge ($-$); (5) the complementary base pairs are AU, GC and GU. It is necessary to note that these tree graph representations do not specify the exact sequence or the length of an RNA molecule, even though the length can be approximated. Furthermore, these graphs do not specify geometric

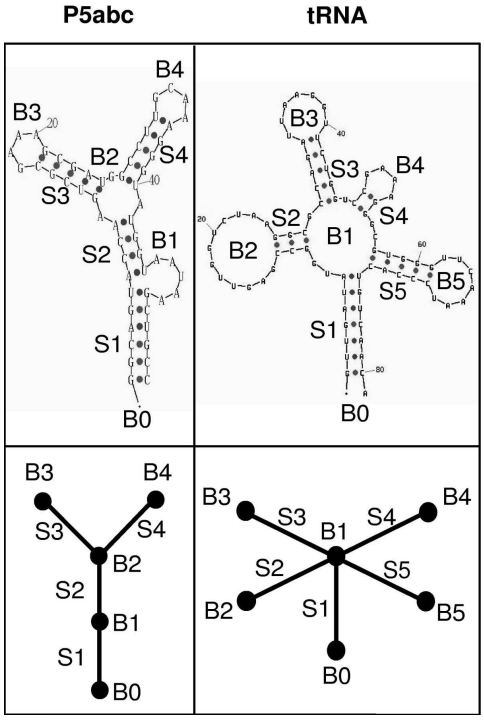

Figure 7: The Modeling Method Used by Schlick et al. for the RAG Web Resource This figure was copied from [14] with the author's permission.

aspects of the secondary structure. These models solely represent the connectivity of an RNA motif.

Every novel RNA structure is mapped onto a 2D tree model and then cataloged according to two numerical values: the number of vertices (n) in each graph and a numerical representation (z) of its topological complexity. To quantitatively organize and archive all possible RNA tree graph representations, it is necessary to first generate the collection of all possible tree graphs for a given number of vertices (n). For example, for the set $n = \{2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9\}$, there are a total of $\{1,\ 1,\ 2,\ 3,\ 6,\ 11,\ 23,\ 47\}$ distinct tree graphs, respectively. Once all possible tree motifs were generated, Schlick et al. in [14] generated the Laplacian matrix representation and calculated the eigenvalue spectrum $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ for each tree.

According to [14], the second eigenvalue λ_2 measures a motif's topological complexity. For example, a more linear tree graph has a lower λ_2 value while a highly-branched tree graph has a high λ_2 value. Then, this work catalogs each potential RNA motif by (n, λ_2). For easy reference, each RNA motif then has a specific index $(n.z)$, where z represents an integer corresponding to the λ_2 ranking.

There are other aspects of the RAG database described in [14]. However, for the intentions of this thesis, it is necessary to close this review with a description of Schlick's clustering analysis of the RNA tree graphs. As previously stated, the RNA's topological complexity is described by Laplacian eigenvalues. However, these numbers can also be used to predict the RNA topologies that are likely to occur naturally. Schlick et al. [14] applied the method of Partitioning Around Medoids (PAM) to assemble the enumerated RNA tree graphs into RNA-like and non-RNA clusters or groups. The RNA-like cluster must contain predominately existing RNA topologies and the non-RNA cluster must contain few or no natural RNA motifs in it. Then, those clustered into the RNA-like category which have not been identified naturally became known as *candidate* trees. Those grouped into the non-RNA category were labeled as *non-candidate* trees.

The trees with 2 through 8 vertices have been fully classified as existing, candidate or non-candidate. However, there are only four trees on greater than 8 vertices which have been found to be existing RNA motifs. As a result, the remaining trees on 9 or higher vertices have not been grouped into the existing, candidate, or non-candidate categories. Finally, the online web database [15] takes one additional step to catalog these secondary motifs. This database pictures all existing RNA trees as as a red graphs, the candidate trees are colored blue, and the non-candidate trees are black.

The research compiled in [14] uniquely organizes all known, candidate, and non-candidate RNA motifs by schematic graphical representations. This catalog of RNA tree graphs is intended as a tool for searching existing RNAs and for discovering additional RNA molecules. According to [14], their main goals were to systematically catalog all possible RNA motif libraries, rank RNA motifs with different degrees of topological complexity and stimulate the search for candidate RNA motifs not yet discovered in nature or a laboratory.

A Quantitative Analysis of Secondary RNA Structures Using Domination Based Parameters on Trees

Based on the work compiled in [14], D. Knisley et al. [9] used the RAG online database to classify the RNA motifs based on graphical invariants. According to Knisley et al. [9], using tree graphs as models of RNA, proteins and nucleic acids is fertile ground for the discovery of new and innovative methods for the numerical characterization of these biological molecules. Furthermore, the information gathered from the analysis of graph theoretic models plays a vital role in assisting protein structure prediction algorithms. As a result, this work centralized on identifying structural characteristics of secondary RNA models.

Given the library of trees generated in [14], the work in [9] analyzed the structural properties of the RNA tree motifs to classify them as RNA-like or non-RNA models. In particular, they used domination based invariants, properties of graphs that are fixed under graph isomorphisms, whose definitions can be found in [6]. In calculating and tabulating a multitude of graphical invariants for each tree from the RAG database, Knisley et al. [9] noticed that the domination parameters behaved in two distinct ways with respect to the Laplacian eigenvalues discussed in [14]. The domination (γ), total domination(γ_t), and global alliance (γ_a) numbers tended to decrease as the eigenvalues increased. Further, the locating-domination (γ_L) and differentiating-domination (γ_D) numbers tended to increase as the eigenvalues increased. As a result, they grouped these invariants into two classes, namely P_1 and P_2. These parameters are defined as follows:

$$P_1 = \frac{\gamma+\gamma_t+\gamma_a}{n} \text{ and } P_2 = \frac{\gamma_L+\gamma_D}{n}.$$

Dividing each parameter by the total number of vertices in the tree (n) normalized

the results. A third parameter, P_2^* was defined as follows:

$$P_2^* = \gamma_L + \gamma_D + n \cdot \lambda_2$$

The work completed in [9] then applied logistic models to predict the probability that a tree is a native RNA structure based on the three parameters P_1, P_2, and P_2^* described above. Two different logistic models were estimated using SAS, one based on P_1 and P_2 and another model based on P_1 and P_2^*. These logistic models correctly classified 100% of existing and non-candidate RNA tree models. When considering trees on 7 or 8 vertices, the logistic model classified ten of the twelve candidate RNA trees to be RNA-like and two of the twelve candidate trees to be non-RNA models. These results can be seen in Table 7.

The research in [9] utilized logistic models to predict the probability that a tree from the RAG database is an existing or non-candidate RNA structure. These logistic models were created with domination based invariants from graph theory and correctly classified all existing and non-candidate RNA tree motifs on 7 and 8 vertices.

Table 7: The Status and Prediction for Trees with Seven and Eight Vertices from [9]

Vertices	ID	P(Native) Model 1	P(Native) Model 2	RAG Status	Domination Predicted Status
7	1	1.00000	1.00000	known	RNA-Like
7	2	0.99898	0.99991	known	RNA-Like
7	3	1.00000	1.00000	known	RNA-Like
7	4	0.00040	0.00392	candidate	Not-RNA-Like
7	5	0.99951	0.99991	candidate	RNA-Like
7	6	0.99834	0.99908	known	RNA-Like
7	7	0.99911	0.99908	candidate	RNA-Like
7	8	1.00000	1.00000	candidate	RNA-Like
7	9	0.00000	0.00000	non-candidate	Not-RNA-Like
7	10	0.00000	0.00000	non-candidate	Not-RNA-Like
7	11	0.00002	0.00000	non-candidate	Not-RNA-Like
8	1	1.00000	1.00000	candidate	RNA-Like
8	2	1.00000	1.00000	candidate	RNA-Like
8	3	1.00000	1.00000	known	RNA-Like
8	4	0.98853	0.99359	candidate	RNA-Like
8	5	1.00000	1.00000	known	RNA-Like
8	6	0.99049	0.99359	candidate	RNA-Like
8	7	1.00000	1.00000	known	RNA-Like
8	8	0.96824	0.95104	candidate	RNA-Like
8	9	0.00124	0.00269	non-candidate	Not-RNA-Like
8	10	1.00000	1.00000	known	RNA-Like
8	11	1.00000	1.00000	known	RNA-Like
8	12	1.00000	1.00000	candidate	RNA-Like
8	13	0.00040	0.00034	candidate	Not-RNA-Like
8	14	0.00196	0.00269	non-candidate	Not-RNA-Like
8	15	0.99659	0.99359	known	RNA-Like
8	16	1.00000	1.00000	candidate	RNA-Like
8	17	0.00073	0.00034	non-candidate	Not-RNA-Like
8	18	0.00000	0.00000	non-candidate	Not-RNA-Like
8	19	0.00000	0.00000	non-candidate	Not-RNA-Like
8	20	1.00000	1.00000	known	RNA-Like
8	21	0.00154	0.00034	non-candidate	Not-RNA-Like
8	22	0.00000	0.00000	non-candidate	Not-RNA-Like
8	23	0.00000	0.00000	non-candidate	Not-RNA-Like

Using a Neural Network to Identify Secondary RNA Structures Quantified by Graphical Invariants

As a follow up to the research presented in [9], D. Knisley et al. employed a predictive model in [10] to analyze the RNA motifs from the RAG online database of [15]. Their work also provides additional information about RNA secondary structures. According to their research, the class of non-coding RNAs is rapidly expanding, and evidence suggests that half of the human RNAs are for non-coding purposes [10]. This motivates the creation of a comprehensive database of RNA motifs with both structural and sequential information. Accordingly, their research parallels previous work described in [9]. However, it expands upon these previous findings by applying a predictive model to recognize RNA-like or non-RNA-like secondary structures.

The key component of the research in [10] derives from the application of an artificial neural network to classify the unknown tree structures on 9 vertices from the RAG online database. Knisley et al. [10] trained an artificial neural network using the same parameters P_1 and P_2, from [9] along with some additional invariants. The definitions of the following invariants can be found in [10]. First, Knisley et al. [10] created the line graph, denoted $L(T)$, for each tree in the RAG online database and then calculated the radius ($rad(L)$), diameter ($diam(L)$), and the block (B) for each line graph. These invariants were normalized into a third parameter P_3 where $\|B\|$ is the number of blocks in the line graph and n is the number of vertices in the line graph:

$$P_3 = \frac{diam(L(T))+rad(L(T))+\|B\|}{n}$$

Using P_1, P_2, and P_3 as their graph theoretic parameters, Knisley et al. [10] trained a multi-layer perceptron (MLP) artificial neural network using a standard

back propagation algorithm. There are twenty-two trees of order 7 or 8 which are verified to be existing RNA motifs or classified as not-RNA-like (non-candidate) tree structures. There are four trees of order 9 which represent existing RNA motifs. With the parameter information of these twenty-six classified trees, an artificial neural network was trained and tested using *leave one out (LOO)* cross-validation. For these twenty-six motifs, Table 8 (from [10]) displays each known tree's classification and corresponding prediction error from the artificial neural network.

Table 8: Error in Predicting the Class of the Given Tree using Leave One Out Cross Validation: Class = 1 if Tree is RNA-Like, Class = 0 if not RNA-Like

Tree	Class	Error	Tree	Class	Error
7.1	1	3.22636E-08	8.14	0	2.86665E-05
7.2	1	6.06176E-05	8.15	1	0.210246175
7.3	1	5.29676E-08	8.17	0	0.020610822
7.6	1	4.35718E-05	8.18	0	7.11512E-09
7.9	0	3.67912E-08	8.19	0	3.0159E-08
7.10	0	6.13133E-09	8.20	1	2.60021E-07
7.11	0	3.136E-09	8.21	0	0.020787623
8.3	1	3.19055E-08	8.22	0	2.04457E-07
8.5	1	2.68612E-08	8.23	0	5.2509E-09
8.7	1	4.32724E-06	9.6	1	3.14169E-08
8.9	0	3.72807E-05	9.11	1	1.10466E-06
8.10	1	5.86222E-07	9.13	1	0.002645342
8.11	1	3.81045E-08	9.27	1	5.74173E-07

Following its training, the MLP was used to predict whether or not the candidate trees of orders 7 or 8 and the unclassified trees of order 9 were RNA-like or not RNA-like based on their respective three parameter values. These results are shown in Table 9.

The research presented in [10] demonstrates that when secondary RNA motifs are represented by tree models and quantified with graphical invariants, their mathemat-

Table 9: Predictions for the 55 Unclassified Trees from [10]

Tree	Class	StDev	Tree	Class	StDev
7.4	0.30894009	0.059897352	9.2	0.977981363	0.03237958
7.5	0.963595271	0.044234638	9.21	0.999997116	7.81E-06
7.7	0.999900712	0.000222055	9.22	0.030237041	0.007916403
7.8	0.999999663	4.30E-07	9.23	4.64E-07	1.23E-06
8.1	0.999999775	3.32E-07	9.24	0.999971294	5.19E-05
8.2	0.999999552	7.27E-07	9.25	9.32E-08	1.99E-07
8.4	0.998541188	0.004658326	9.26	0.999860004	0.000365012
8.6	0.998541188	0.004658326	9.28	1.69E-05	2.10E-05
8.8	0.999995579	8.72E-06	9.29	0.999999319	8.96E-07
8.12	0.999995579	8.72E-06	9.3	0.999971294	5.19E-05
8.13	0.014562009	0.001888189	9.31	0.977981363	0.03237958
8.16	0.999995579	8.72E-06	9.32	1.27E-06	2.21E-06
9.1	0.999999871	1.99E-07	9.33	0.58286479	0.056754084
9.2	0.999996962	1.02E-05	9.34	0.973029577	0.013048741
9.3	0.977981363	0.03237958	9.35	3.38E-06	4.69E-06
9.4	0.999999809	2.68E-07	9.36	0.047779882	0.029252496
9.5	0.999993215	2.02E-05	9.37	1.69E-05	2.10E-05
9.7	0.999997116	7.81E-06	9.38	0.998666897	0.00214774
9.8	0.999997116	7.81E-06	9.39	2.24E-07	5.27E-07
9.9	0.00014299	0.000150072	9.4	4.13E-08	7.92E-08
9.1	0.999971294	5.19E-05	9.41	3.68E-05	7.86E-05
9.12	0.999999566	6.33E-07	9.42	0.999997632	2.87E-06
9.14	0.999999764	3.16E-07	9.43	0.973029577	0.013048741
9.15	0.004035487	0.000863529	9.44	0.095946916	0.014793406
9.16	0.477586547	0.074602833	9.45	2.76E-07	5.15E-07
9.17	0.999802708	0.000223001	9.46	2.37E-08	4.47E-08
9.18	0.999999597	5.96E-07	9.47	3.60E-07	1.01E-06
9.19	0.999998844	2.27E-06			

ical properties can be indicative of novel RNA structure. Further, this work shows that these numerical parameters classify RNA secondary structure well enough so that an artificial neural network can be trained to recognize the difference.

Appendix B: Table A

The Complete Table of all Vertex Identifications from all Tree Merges

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
2.1+2.1	Red	Leaf	1	Red	Leaf	1	3.1	Red	4
2.1+3.1	Red	Leaf	1	Red	Leaf	1	4.1	Red	4
2.1+3.1	Red	Leaf	1	Red	Support	2	4.2	Blue	2
2.1+4.1	Red	Leaf	1	Red	Leaf	1	5.1	Red	4
2.1+4.1	Red	Leaf	1	Red	Support	2	5.2	Red	4
2.1+4.2	Red	Leaf	1	Blue	Leaf	1	5.2	Red	6
2.1+4.2	Red	Leaf	1	Blue	Support	3	5.3	Red	2
2.1+5.1	Red	Leaf	1	Red	Leaf	1	6.1	Blue	4
2.1+5.1	Red	Leaf	1	Red	Support	2	6.2	Red	4
2.1+5.1	Red	Leaf	1	Red	Internal	2	6.3	Blue	2
2.1+5.2	Red	Leaf	1	Red	Leaf	1	6.2	Red	2
2.1+5.2	Red	Leaf	1	Red	Leaf	1	6.3	Blue	4
2.1+5.2	Red	Leaf	1	Red	Support	2	6.4	Blue	2
2.1+5.2	Red	Leaf	1	Red	Support	3	6.5	Black	2
2.1+5.3	Red	Leaf	1	Red	Leaf	1	6.5	Black	8
2.1+5.3	Red	Leaf	1	Red	Support	4	6.6	Red	2
2.1+6.1	Red	Leaf	1	Blue	Leaf	1	7.1	Red	4
2.1+6.1	Red	Leaf	1	Blue	Support	2	7.2	Red	4
2.1+6.1	Red	Leaf	1	Blue	Internal	2	7.3	Red	4
2.1+6.2	Red	Leaf	1	Red	Leaf	1	7.2	Red	2
2.1+6.2	Red	Leaf	1	Red	Leaf	1	7.3	Red	4
2.1+6.2	Red	Leaf	1	Red	Support	2	7.4	Blue	2
2.1+6.2	Red	Leaf	1	Red	Support	3	7.5	Blue	2
2.1+6.2	Red	Leaf	1	Red	Internal	2	7.6	Red	2
2.1+6.3	Red	Leaf	1	Blue	Leaf	1	7.3	Red	4
2.1+6.3	Red	Leaf	1	Blue	Support	2	7.6	Red	4
2.1+6.3	Red	Leaf	1	Blue	Support	3	7.7	Blue	2
2.1+6.3	Red	Leaf	1	Blue	Leaf	1	7.8	Blue	2
2.1+6.4	Red	Leaf	1	Blue	Leaf	1	7.6	Red	8
2.1+6.4	Red	Leaf	1	Blue	Support	3	7.9	Black	4
2.1+6.5	Red	Leaf	1	Black	Support	4	7.10	Black	2
2.1+6.5	Red	Leaf	1	Black	Leaf	1	7.5	Blue	2
2.1+6.5	Red	Leaf	1	Black	Leaf	1	7.7	Blue	6
2.1+6.5	Red	Leaf	1	Black	Support	2	7.9	Black	2
2.1+6.6	Red	Leaf	1	Red	Leaf	1	7.10	Black	10
2.1+6.6	Red	Leaf	1	Red	Support	5	7.11	Black	2
2.1+7.1	Red	Leaf	1	Red	Leaf	1	8.1	Blue	4
2.1+7.1	Red	Leaf	1	Red	Support	2	8.2	Blue	4
2.1+7.1	Red	Leaf	1	Red	Internal	2	8.3	Red	4
2.1+7.1	Red	Leaf	1	Red	Internal	2	8.5	Red	2

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	*deg* (v)	Tree Color	Vert. Type	*deg* (v)	RAG Graph	Tree Color	Total Graphs
2.1+7.10	Red	Leaf	1	Black	Leaf	1	8.14	Black	2
2.1+7.10	Red	Leaf	1	Black	Support	2	8.19	Black	2
2.1+7.10	Red	Leaf	1	Black	Leaf	1	8.21	Black	8
2.1+7.10	Red	Leaf	1	Black	Support	5	8.22	Black	2
2.1+7.11	Red	Leaf	1	Black	Leaf	1	8.22	Black	12
2.1+7.11	Red	Leaf	1	Black	Support	6	8.23	Black	2
2.1+7.2	Red	Leaf	1	Red	Leaf	1	8.2	Blue	2
2.1+7.2	Red	Leaf	1	Red	Leaf	1	8.3	Red	4
2.1+7.2	Red	Leaf	1	Red	Support	2	8.4	Blue	2
2.1+7.2	Red	Leaf	1	Red	Support	3	8.6	Blue	2
2.1+7.2	Red	Leaf	1	Red	Internal	2	8.7	Red	2
2.1+7.2	Red	Leaf	1	Red	Internal	2	8.8	Blue	2
2.1+7.3	Red	Leaf	1	Red	Leaf	1	8.10	Red	2
2.1+7.3	Red	Leaf	1	Red	Internal	2	8.11	Red	2
2.1+7.3	Red	Leaf	1	Red	Internal	3	8.12	Blue	2
2.1+7.3	Red	Leaf	1	Red	Leaf	1	8.3	Red	2
2.1+7.3	Red	Leaf	1	Red	Leaf	1	8.5	Red	2
2.1+7.3	Red	Leaf	1	Red	Support	2	8.7	Red	2
2.1+7.3	Red	Leaf	1	Red	Support	2	8.8	Blue	2
2.1+7.4	Red	Leaf	1	Blue	Internal	2	8.13	Blue	2
2.1+7.4	Red	Leaf	1	Blue	Support	3	8.9	Black	4
2.1+7.5	Red	Leaf	1	Blue	Leaf	1	8.12	Blue	6
2.1+7.5	Red	Leaf	1	Blue	Support	4	8.14	Black	2
2.1+7.5	Red	Leaf	1	Blue	Internal	2	8.15	Red	2
2.1+7.5	Red	Leaf	1	Blue	Leaf	1	8.6	Blue	2
2.1+7.5	Red	Leaf	1	Blue	Support	2	8.9	Black	2
2.1+7.6	Red	Leaf	1	Red	Leaf	1	8.11	Red	4
2.1+7.6	Red	Leaf	1	Red	Support	2	8.13	Blue	2
2.1+7.6	Red	Leaf	1	Red	Support	3	8.15	Red	2
2.1+7.6	Red	Leaf	1	Red	Leaf	1	8.16	Blue	2
2.1+7.6	Red	Leaf	1	Red	Internal	3	8.17	Black	2
2.1+7.6	Red	Leaf	1	Red	Leaf	1	8.8	Blue	2
2.1+7.7	Red	Leaf	1	Blue	Leaf	1	8.12	Blue	4
2.1+7.7	Red	Leaf	1	Blue	Support	2	8.17	Black	4
2.1+7.7	Red	Leaf	1	Blue	Leaf	1	8.20	Red	4
2.1+7.7	Red	Leaf	1	Blue	Internal	4	8.21	Black	2
2.1+7.8	Red	Leaf	1	Blue	Leaf	1	8.10	Red	6
2.1+7.8	Red	Leaf	1	Blue	Support	2	8.16	Blue	6
2.1+7.8	Red	Leaf	1	Blue	Internal	3	8.20	Red	2
2.1+7.9	Red	Leaf	1	Black	Leaf	1	8.15	Red	4

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	*deg* (v)	Tree Color	Vert. Type	*deg* (v)	RAG Graph	Tree Color	Total Graphs
2.1+7.9	Red	Leaf	1	Black	Leaf	1	8.17	Black	6
2.1+7.9	Red	Leaf	1	Black	Support	3	8.18	Black	2
2.1+7.9	Red	Leaf	1	Black	Support	4	8.19	Black	2
2.1+8.1	Red	Leaf	1	Blue	Leaf	1	9.1	Unknown	4
2.1+8.1	Red	Leaf	1	Blue	Support	2	9.2	Unknown	4
2.1+8.1	Red	Leaf	1	Blue	Internal	2	9.4	Unknown	4
2.1+8.1	Red	Leaf	1	Blue	Internal	2	9.6	Red	4
2.1+8.10	Red	Leaf	1	Red	Leaf	1	9.11	Red	2
2.1+8.10	Red	Leaf	1	Red	Support	2	9.17	Unknown	2
2.1+8.10	Red	Leaf	1	Red	Leaf	1	9.18	Unknown	4
2.1+8.10	Red	Leaf	1	Red	Support	2	9.26	Unknown	4
2.1+8.10	Red	Leaf	1	Red	Internal	2	9.27	Red	2
2.1+8.10	Red	Leaf	1	Red	Internal	3	9.29	Unknown	2
2.1+8.11	Red	Leaf	1	Red	Leaf	1	9.14	Unknown	4
2.1+8.11	Red	Leaf	1	Red	Support	2	9.24	Unknown	4
2.1+8.11	Red	Leaf	1	Red	Leaf	1	9.27	Red	4
2.1+8.11	Red	Leaf	1	Red	Internal	3	9.30	Unknown	4
2.1+8.12	Red	Leaf	1	Blue	Leaf	1	9.19	Unknown	2
2.1+8.12	Red	Leaf	1	Blue	Leaf	1	9.21	Unknown	2
2.1+8.12	Red	Leaf	1	Blue	Support	2	9.22	Unknown	2
2.1+8.12	Red	Leaf	1	Blue	Leaf	1	9.29	Unknown	4
2.1+8.12	Red	Leaf	1	Blue	Internal	2	9.30	Unknown	2
2.1+8.12	Red	Leaf	1	Blue	Internal	4	9.31	Unknown	2
2.1+8.12	Red	Leaf	1	Blue	Support	2	9.44	Unknown	2
2.1+8.13	Red	Leaf	1	Blue	Leaf	1	9.24	Unknown	8
2.1+8.13	Red	Leaf	1	Blue	Support	3	9.28	Unknown	4
2.1+8.13	Red	Leaf	1	Blue	Leaf	1	9.33	Unknown	2
2.1+8.13	Red	Leaf	1	Blue	Support	3	9.35	Unknown	2
2.1+8.14	Red	Leaf	1	Black	Leaf	1	9.15	Unknown	2
2.1+8.14	Red	Leaf	1	Black	Support	2	9.25	Unknown	2
2.1+8.14	Red	Leaf	1	Black	Leaf	1	9.31	Unknown	8
2.1+8.14	Red	Leaf	1	Black	Support	2	9.32	Unknown	2
2.1+8.14	Red	Leaf	1	Black	Internal	2	9.36	Unknown	2
2.1+8.15	Red	Leaf	1	Red	Leaf	1	9.20	Unknown	2
2.1+8.15	Red	Leaf	1	Red	Support	2	9.28	Unknown	2
2.1+8.15	Red	Leaf	1	Red	Leaf	1	9.30	Unknown	6
2.1+8.15	Red	Leaf	1	Red	Leaf	1	9.34	Unknown	2
2.1+8.15	Red	Leaf	1	Red	Support	4	9.36	Unknown	2
2.1+8.15	Red	Leaf	1	Red	Internal	3	9.37	Unknown	2
2.1+8.16	Red	Leaf	1	Blue	Leaf	1	9.26	Unknown	4
2.1+8.16	Red	Leaf	1	Blue	Leaf	1	9.27	Red	4

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
2.1+8.16	Red	Leaf	1	Blue	Support	2	9.33	Unknown	4
2.1+8.16	Red	Leaf	1	Blue	Support	3	9.34	Unknown	2
2.1+8.16	Red	Leaf	1	Blue	Internal	3	9.38	Unknown	2
2.1+8.17	Red	Leaf	1	Black	Leaf	1	9.30	Unknown	4
2.1+8.17	Red	Leaf	1	Black	Support	2	9.35	Unknown	2
2.1+8.17	Red	Leaf	1	Black	Support	3	9.37	Unknown	2
2.1+8.17	Red	Leaf	1	Black	Leaf	1	9.38	Unknown	4
2.1+8.17	Red	Leaf	1	Black	Internal	4	9.39	Unknown	2
2.1+8.17	Red	Leaf	1	Black	Leaf	1	9.44	Unknown	2
2.1+8.18	Red	Leaf	1	Black	Leaf	1	9.37	Unknown	12
2.1+8.18	Red	Leaf	1	Black	Support	4	9.40	Unknown	4
2.1+8.19	Red	Leaf	1	Black	Leaf	1	9.36	Unknown	4
2.1+8.19	Red	Leaf	1	Black	Leaf	1	9.39	Unknown	8
2.1+8.19	Red	Leaf	1	Black	Support	3	9.40	Unknown	2
2.1+8.19	Red	Leaf	1	Black	Support	5	9.41	Unknown	2
2.1+8.2	Red	Leaf	1	Blue	Internal	2	9.10	Unknown	2
2.1+8.2	Red	Leaf	1	Blue	Leaf	1	9.2	Unknown	2
2.1+8.2	Red	Leaf	1	Blue	Support	2	9.3	Unknown	2
2.1+8.2	Red	Leaf	1	Blue	Leaf	1	9.4	Unknown	4
2.1+8.2	Red	Leaf	1	Blue	Support	3	9.5	Unknown	2
2.1+8.2	Red	Leaf	1	Blue	Internal	2	9.7	Unknown	2
2.1+8.2	Red	Leaf	1	Blue	Internal	2	9.8	Unknown	2
2.1+8.20	Red	Leaf	1	Red	Leaf	1	9.29	Unknown	6
2.1+8.20	Red	Leaf	1	Red	Support	2	9.38	Unknown	6
2.1+8.20	Red	Leaf	1	Red	Leaf	1	9.42	Unknown	2
2.1+8.20	Red	Leaf	1	Red	Internal	4	9.43	Unknown	2
2.1+8.21	Red	Leaf	1	Black	Leaf	1	9.31	Unknown	4
2.1+8.21	Red	Leaf	1	Black	Support	2	9.39	Unknown	4
2.1+8.21	Red	Leaf	1	Black	Leaf	1	9.43	Unknown	6
2.1+8.21	Red	Leaf	1	Black	Support	5	9.45	Unknown	2
2.1+8.22	Red	Leaf	1	Black	Leaf	1	9.32	Unknown	2
2.1+8.22	Red	Leaf	1	Black	Support	2	9.41	Unknown	2
2.1+8.22	Red	Leaf	1	Black	Leaf	1	9.45	Unknown	10
2.1+8.22	Red	Leaf	1	Black	Internal	5	9.46	Unknown	2
2.1+8.23	Red	Leaf	1	Black	Leaf	1	9.46	Unknown	14
2.1+8.23	Red	Leaf	1	Black	Internal	7	9.47	Unknown	2
2.1+8.3	Red	Leaf	1	Red	Leaf	1	9.11	Red	2
2.1+8.3	Red	Leaf	1	Red	Internal	2	9.12	Unknown	2
2.1+8.3	Red	Leaf	1	Red	Internal	2	9.14	Unknown	2
2.1+8.3	Red	Leaf	1	Red	Support	3	9.21	Unknown	2
2.1+8.3	Red	Leaf	1	Red	Leaf	1	9.4	Unknown	2

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
2.1+8.3	Red	Leaf	1	Red	Leaf	1	9.6	Red	2
2.1+8.3	Red	Leaf	1	Red	Support	2	9.7	Unknown	2
2.1+8.3	Red	Leaf	1	Red	Support	2	9.8	Unknown	2
2.1+8.4	Red	Leaf	1	Blue	Internal	2	9.13	Red	4
2.1+8.4	Red	Leaf	1	Blue	Leaf	1	9.7	Unknown	8
2.1+8.4	Red	Leaf	1	Blue	Support	3	9.9	Unknown	4
2.1+8.5	Red	Leaf	1	Red	Support	2	9.10	Unknown	4
2.1+8.5	Red	Leaf	1	Red	Internal	2	9.14	Unknown	4
2.1+8.5	Red	Leaf	1	Red	Leaf	1	9.18	Unknown	2
2.1+8.5	Red	Leaf	1	Red	Support	3	9.19	Unknown	2
2.1+8.5	Red	Leaf	1	Red	Leaf	1	9.6	Red	4
2.1+8.6	Red	Leaf	1	Blue	Support	4	9.15	Unknown	2
2.1+8.6	Red	Leaf	1	Blue	Internal	2	9.16	Unknown	2
2.1+8.6	Red	Leaf	1	Blue	Internal	2	9.20	Unknown	2
2.1+8.6	Red	Leaf	1	Blue	Leaf	1	9.21	Unknown	6
2.1+8.6	Red	Leaf	1	Blue	Leaf	1	9.5	Unknown	2
2.1+8.6	Red	Leaf	1	Blue	Support	2	9.9	Unknown	2
2.1+8.7	Red	Leaf	1	Red	Leaf	1	9.10	Unknown	2
2.1+8.7	Red	Leaf	1	Red	Leaf	1	9.12	Unknown	4
2.1+8.7	Red	Leaf	1	Red	Support	2	9.13	Red	2
2.1+8.7	Red	Leaf	1	Red	Support	3	9.16	Unknown	2
2.1+8.7	Red	Leaf	1	Red	Leaf	1	9.17	Unknown	2
2.1+8.7	Red	Leaf	1	Red	Internal	3	9.22	Unknown	2
2.1+8.7	Red	Leaf	1	Red	Internal	2	9.24	Unknown	2
2.1+8.8	Red	Leaf	1	Blue	Support	2	9.13	Red	2
2.1+8.8	Red	Leaf	1	Blue	Leaf	1	9.14	Unknown	4
2.1+8.8	Red	Leaf	1	Blue	Support	3	9.20	Unknown	2
2.1+8.8	Red	Leaf	1	Blue	Internal	2	9.24	Unknown	2
2.1+8.8	Red	Leaf	1	Blue	Leaf	1	9.26	Unknown	2
2.1+8.8	Red	Leaf	1	Blue	Internal	3	9.44	Unknown	2
2.1+8.8	Red	Leaf	1	Blue	Leaf	1	9.8	Unknown	2
2.1+8.9	Red	Leaf	1	Black	Leaf	1	9.16	Unknown	4
2.1+8.9	Red	Leaf	1	Black	Leaf	1	9.22	Unknown	6
2.1+8.9	Red	Leaf	1	Black	Support	3	9.23	Unknown	2
2.1+8.9	Red	Leaf	1	Black	Support	4	9.25	Unknown	2
2.1+8.9	Red	Leaf	1	Black	Internal	2	9.28	Unknown	2
3.1+3.1	Red	Leaf	1	Red	Leaf	1	5.1	Red	4
3.1+3.1	Red	Leaf	1	Red	Support	2	5.2	Red	4
3.1+3.1	Red	Support	2	Red	Support	2	5.3	Red	1
3.1+4.1	Red	Leaf	1	Red	Leaf	1	6.1	Blue	4
3.1+4.1	Red	Support	2	Red	Leaf	1	6.2	Red	2

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
3.1+4.1	Red	Leaf	1	Red	Support	2	6.3	Blue	4
3.1+4.1	Red	Support	2	Red	Support	2	6.5	Black	2
3.1+4.2	Red	Leaf	1	Blue	Leaf	1	6.2	Red	6
3.1+4.2	Red	Support	2	Blue	Leaf	1	6.4	Blue	3
3.1+4.2	Red	Leaf	1	Blue	Support	3	6.5	Black	2
3.1+4.2	Red	Support	2	Blue	Support	3	6.6	Red	1
3.1+5.1	Red	Leaf	1	Red	Leaf	1	7.1	Red	4
3.1+5.1	Red	Support	2	Red	Leaf	1	7.2	Red	2
3.1+5.1	Red	Leaf	1	Red	Support	2	7.3	Red	4
3.1+5.1	Red	Support	2	Red	Support	2	7.5	Blue	2
3.1+5.1	Red	Support	2	Red	Internal	2	7.7	Blue	1
3.1+5.1	Red	Leaf	1	Red	Internal	2	7.8	Blue	2
3.1+5.2	Red	Support	2	Red	Support	3	7.10	Black	1
3.1+5.2	Red	Leaf	1	Red	Leaf	1	7.2	Red	2
3.1+5.2	Red	Leaf	1	Red	Leaf	1	7.3	Red	4
3.1+5.2	Red	Support	2	Red	Leaf	1	7.4	Blue	1
3.1+5.2	Red	Leaf	1	Red	Support	2	7.6	Red	4
3.1+5.2	Red	Leaf	1	Red	Support	3	7.7	Blue	2
3.1+5.2	Red	Support	2	Red	Support	2	7.9	Black	1
3.1+5.3	Red	Leaf	1	Red	Support	4	7.10	Black	2
3.1+5.3	Red	Support	2	Red	Support	4	7.11	Black	1
3.1+5.3	Red	Leaf	1	Red	Leaf	1	7.5	Blue	8
3.1+5.3	Red	Support	2	Red	Leaf	1	7.9	Black	4
3.1+6.1	Red	Leaf	1	Blue	Leaf	1	8.1	Blue	4
3.1+6.1	Red	Leaf	1	Blue	Internal	2	8.10	Red	4
3.1+6.1	Red	Support	2	Blue	Internal	2	8.12	Blue	2
3.1+6.1	Red	Support	2	Blue	Leaf	1	8.2	Blue	2
3.1+6.1	Red	Leaf	1	Blue	Support	2	8.3	Red	4
3.1+6.1	Red	Support	2	Blue	Support	2	8.6	Blue	2
3.1+6.2	Red	Leaf	1	Red	Internal	3	8.12	Blue	2
3.1+6.2	Red	Support	2	Red	Internal	3	8.14	Black	1
3.1+6.2	Red	Leaf	1	Red	Internal	2	8.16	Blue	2
3.1+6.2	Red	Support	2	Red	Internal	2	8.17	Black	1
3.1+6.2	Red	Leaf	1	Red	Leaf	1	8.2	Blue	2
3.1+6.2	Red	Support	2	Red	Leaf	1	8.4	Blue	1
3.1+6.2	Red	Leaf	1	Red	Leaf	1	8.5	Red	4
3.1+6.2	Red	Leaf	1	Red	Support	2	8.7	Red	2
3.1+6.2	Red	Support	2	Red	Leaf	1	8.8	Blue	2
3.1+6.2	Red	Support	2	Red	Support	2	8.9	Black	1
3.1+6.3	Red	Leaf	1	Blue	Leaf	1	8.10	Red	2
3.1+6.3	Red	Leaf	1	Blue	Support	2	8.11	Red	4

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
3.1+6.3	Red	Leaf	1	Blue	Leaf	1	8.13	Blue	4
3.1+6.3	Red	Support	2	Blue	Support	2	8.15	Red	2
3.1+6.3	Red	Support	2	Blue	Leaf	1	8.16	Blue	1
3.1+6.3	Red	Leaf	1	Blue	Support	3	8.20	Red	2
3.1+6.3	Red	Support	2	Blue	Support	3	8.21	Black	1
3.1+6.3	Red	Support	2	Blue	Leaf	1	8.7	Red	2
3.1+6.4	Red	Support	2	Blue	Leaf	1	8.13	Blue	4
3.1+6.4	Red	Leaf	1	Blue	Support	3	8.17	Black	4
3.1+6.4	Red	Support	2	Blue	Support	3	8.19	Black	2
3.1+6.4	Red	Leaf	1	Blue	Leaf	1	8.8	Blue	8
3.1+6.5	Red	Leaf	1	Black	Leaf	1	8.12	Blue	6
3.1+6.5	Red	Leaf	1	Black	Support	2	8.15	Red	2
3.1+6.5	Red	Support	2	Black	Leaf	1	8.17	Black	3
3.1+6.5	Red	Support	2	Black	Support	2	8.18	Black	1
3.1+6.5	Red	Leaf	1	Black	Support	4	8.21	Black	2
3.1+6.5	Red	Support	2	Black	Support	4	8.22	Black	1
3.1+6.5	Red	Leaf	1	Black	Leaf	1	8.6	Blue	2
3.1+6.5	Red	Support	2	Black	Leaf	1	8.9	Black	1
3.1+6.6	Red	Leaf	1	Red	Leaf	1	8.14	Black	10
3.1+6.6	Red	Support	2	Red	Leaf	1	8.19	Black	5
3.1+6.6	Red	Leaf	1	Red	Support	5	8.22	Black	2
3.1+6.6	Red	Support	2	Red	Support	5	8.23	Black	1
3.1+7.1	Red	Leaf	1	Red	Leaf	1	9.1	Unknown	4
3.1+7.1	Red	Leaf	1	Red	Internal	2	9.11	Red	4
3.1+7.1	Red	Leaf	1	Red	Internal	2	9.18	Unknown	2
3.1+7.1	Red	Support	2	Red	Internal	2	9.19	Unknown	1
3.1+7.1	Red	Support	2	Red	Leaf	1	9.2	Unknown	2
3.1+7.1	Red	Support	2	Red	Internal	2	9.21	Unknown	2
3.1+7.1	Red	Leaf	1	Red	Support	2	9.4	Unknown	4
3.1+7.1	Red	Support	2	Red	Support	2	9.5	Unknown	2
3.1+7.10	Red	Leaf	1	Black	Leaf	1	9.15	Unknown	2
3.1+7.10	Red	Support	2	Black	Leaf	1	9.25	Unknown	1
3.1+7.10	Red	Leaf	1	Black	Leaf	1	9.31	Unknown	8
3.1+7.10	Red	Leaf	1	Black	Support	2	9.36	Unknown	2
3.1+7.10	Red	Support	2	Black	Leaf	1	9.39	Unknown	4
3.1+7.10	Red	Support	2	Black	Support	2	9.40	Unknown	1
3.1+7.10	Red	Leaf	1	Black	Support	5	9.45	Unknown	2
3.1+7.10	Red	Support	2	Black	Support	5	9.46	Unknown	1
3.1+7.11	Red	Leaf	1	Black	Leaf	1	9.32	Unknown	12
3.1+7.11	Red	Support	2	Black	Leaf	1	9.41	Unknown	6
3.1+7.11	Red	Leaf	1	Black	Support	6	9.46	Unknown	2

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
3.1+7.11	Red	Support	2	Black	Support	6	9.47	Unknown	1
3.1+7.2	Red	Support	2	Red	Support	3	9.15	Unknown	1
3.1+7.2	Red	Leaf	1	Red	Internal	2	9.17	Unknown	2
3.1+7.2	Red	Leaf	1	Red	Leaf	1	9.2	Unknown	2
3.1+7.2	Red	Leaf	1	Red	Support	3	9.21	Unknown	2
3.1+7.2	Red	Support	2	Red	Internal	2	9.22	Unknown	1
3.1+7.2	Red	Leaf	1	Red	Internal	2	9.26	Unknown	2
3.1+7.2	Red	Support	2	Red	Leaf	1	9.3	Unknown	1
3.1+7.2	Red	Support	2	Red	Internal	2	9.44	Unknown	1
3.1+7.2	Red	Leaf	1	Red	Leaf	1	9.6	Red	4
3.1+7.2	Red	Leaf	1	Red	Support	2	9.7	Unknown	2
3.1+7.2	Red	Support	2	Red	Leaf	1	9.8	Unknown	2
3.1+7.2	Red	Support	2	Red	Support	2	9.9	Unknown	1
3.1+7.3	Red	Support	2	Red	Leaf	1	9.10	Unknown	1
3.1+7.3	Red	Leaf	1	Red	Support	2	9.12	Unknown	2
3.1+7.3	Red	Leaf	1	Red	Support	2	9.14	Unknown	2
3.1+7.3	Red	Support	2	Red	Support	2	9.16	Unknown	1
3.1+7.3	Red	Leaf	1	Red	Leaf	1	9.18	Unknown	2
3.1+7.3	Red	Support	2	Red	Support	2	9.20	Unknown	1
3.1+7.3	Red	Support	2	Red	Leaf	1	9.26	Unknown	1
3.1+7.3	Red	Leaf	1	Red	Internal	2	9.27	Red	2
3.1+7.3	Red	Leaf	1	Red	Internal	3	9.29	Unknown	2
3.1+7.3	Red	Support	2	Red	Internal	2	9.30	Unknown	1
3.1+7.3	Red	Support	2	Red	Internal	3	9.31	Unknown	1
3.1+7.3	Red	Leaf	1	Red	Leaf	1	9.4	Unknown	2
3.1+7.3	Red	Leaf	1	Red	Leaf	1	9.6	Red	2
3.1+7.3	Red	Support	2	Red	Leaf	1	9.7	Unknown	1
3.1+7.4	Red	Leaf	1	Blue	Leaf	1	9.10	Unknown	8
3.1+7.4	Red	Support	2	Blue	Leaf	1	9.13	Red	4
3.1+7.4	Red	Leaf	1	Blue	Support	3	9.22	Unknown	4
3.1+7.4	Red	Support	2	Blue	Support	3	9.25	Unknown	2
3.1+7.4	Red	Leaf	1	Blue	Internal	2	9.33	Unknown	2
3.1+7.4	Red	Support	2	Blue	Internal	2	9.35	Unknown	1
3.1+7.5	Red	Leaf	1	Blue	Support	2	9.16	Unknown	2
3.1+7.5	Red	Leaf	1	Blue	Leaf	1	9.19	Unknown	6
3.1+7.5	Red	Support	2	Blue	Support	2	9.23	Unknown	1
3.1+7.5	Red	Leaf	1	Blue	Support	4	9.31	Unknown	2
3.1+7.5	Red	Support	2	Blue	Support	4	9.32	Unknown	1
3.1+7.5	Red	Leaf	1	Blue	Internal	2	9.34	Unknown	2
3.1+7.5	Red	Support	2	Blue	Internal	2	9.37	Unknown	1
3.1+7.5	Red	Support	2	Blue	Leaf	1	9.44	Unknown	3
3.1+7.5	Red	Leaf	1	Blue	Leaf	1	9.5	Unknown	2

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
3.1+7.5	Red	Support	2	Blue	Leaf	1	9.9	Unknown	1
3.1+7.6	Red	Support	2	Red	Leaf	1	9.13	Red	1
3.1+7.6	Red	Leaf	1	Red	Leaf	1	9.14	Unknown	4
3.1+7.6	Red	Leaf	1	Red	Support	2	9.24	Unknown	2
3.1+7.6	Red	Support	2	Red	Leaf	1	9.24	Unknown	2
3.1+7.6	Red	Leaf	1	Red	Leaf	1	9.26	Unknown	2
3.1+7.6	Red	Support	2	Red	Support	2	9.28	Unknown	1
3.1+7.6	Red	Leaf	1	Red	Support	3	9.30	Unknown	2
3.1+7.6	Red	Support	2	Red	Leaf	1	9.33	Unknown	1
3.1+7.6	Red	Support	2	Red	Support	3	9.36	Unknown	1
3.1+7.6	Red	Leaf	1	Red	Internal	3	9.38	Unknown	2
3.1+7.6	Red	Support	2	Red	Internal	3	9.39	Unknown	1
3.1+7.6	Red	Leaf	1	Red	Leaf	1	9.8	Unknown	2
3.1+7.7	Red	Leaf	1	Blue	Leaf	1	9.21	Unknown	4
3.1+7.7	Red	Support	2	Blue	Leaf	1	9.22	Unknown	2
3.1+7.7	Red	Leaf	1	Blue	Leaf	1	9.29	Unknown	4
3.1+7.7	Red	Leaf	1	Blue	Support	2	9.30	Unknown	4
3.1+7.7	Red	Support	2	Blue	Support	2	9.37	Unknown	2
3.1+7.7	Red	Support	2	Blue	Leaf	1	9.38	Unknown	2
3.1+7.7	Red	Leaf	1	Blue	Internal	4	9.43	Unknown	2
3.1+7.7	Red	Support	2	Blue	Internal	3	9.45	Unknown	1
3.1+7.8	Red	Leaf	1	Blue	Leaf	1	9.11	Red	6
3.1+7.8	Red	Support	2	Blue	Leaf	1	9.17	Unknown	3
3.1+7.8	Red	Leaf	1	Blue	Support	2	9.27	Red	6
3.1+7.8	Red	Support	2	Blue	Support	2	9.34	Unknown	3
3.1+7.8	Red	Leaf	1	Blue	Internal	3	9.42	Unknown	2
3.1+7.8	Red	Support	2	Blue	Internal	3	9.43	Unknown	1
3.1+7.9	Red	Leaf	1	Black	Leaf	1	9.20	Unknown	4
3.1+7.9	Red	Support	2	Black	Leaf	1	9.28	Unknown	2
3.1+7.9	Red	Support	2	Black	Leaf	1	9.35	Unknown	3
3.1+7.9	Red	Leaf	1	Black	Support	3	9.37	Unknown	2
3.1+7.9	Red	Leaf	1	Black	Support	4	9.39	Unknown	2
3.1+7.9	Red	Support	2	Black	Support	3	9.40	Unknown	1
3.1+7.9	Red	Support	2	Black	Support	3	9.41	Unknown	1
3.1+7.9	Red	Leaf	1	Black	Leaf	1	9.44	Unknown	6
4.1+4.1	Red	Leaf	1	Red	Leaf	1	7.1	Red	4
4.1+4.1	Red	Leaf	1	Red	Support	2	7.3	Red	8
4.1+4.1	Red	Support	2	Red	Support	2	7.7	Blue	4
4.1+4.2	Red	Support	2	Blue	Support	3	7.10	Black	2
4.1+4.2	Red	Leaf	1	Blue	Leaf	1	7.2	Red	6
4.1+4.2	Red	Leaf	1	Blue	Support	3	7.5	Blue	2

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
4.1+4.2	Red	Support	2	Blue	Leaf	1	7.6	Red	6
4.1+5.1	Red	Leaf	1	Red	Leaf	1	8.1	Blue	4
4.1+5.1	Red	Leaf	1	Red	Internal	2	8.10	Red	2
4.1+5.1	Red	Support	2	Red	Support	2	8.12	Blue	4
4.1+5.1	Red	Support	2	Red	Internal	2	8.20	Red	2
4.1+5.1	Red	Support	2	Red	Leaf	1	8.3	Red	4
4.1+5.1	Red	Leaf	1	Red	Support	2	8.5	Red	4
4.1+5.2	Red	Support	2	Red	Support	3	8.12	Blue	2
4.1+5.2	Red	Support	2	Red	Support	2	8.17	Black	2
4.1+5.2	Red	Leaf	1	Red	Leaf	1	8.2	Blue	2
4.1+5.2	Red	Support	2	Red	Support	3	8.21	Black	2
4.1+5.2	Red	Leaf	1	Red	Leaf	1	8.3	Red	4
4.1+5.2	Red	Support	2	Red	Leaf	1	8.7	Red	2
4.1+5.2	Red	Leaf	1	Red	Support	2	8.8	Blue	2
4.1+5.2	Red	Support	2	Red	Leaf	1	8. 11	Red	4
4.1+5.3	Red	Leaf	1	Red	Support	4	8.14	Black	2
4.1+5.3	Red	Support	2	Red	Leaf	1	8.15	Red	8
4.1+5.3	Red	Support	2	Red	Support	4	8.22	Black	2
4.1+5.3	Red	Leaf	1	Red	Leaf	1	8.6	Blue	8
4.1+6.1	Red	Leaf	1	Blue	Leaf	1	9.1	Unknown	4
4.1+6.1	Red	Leaf	1	Blue	Internal	2	9.18	Unknown	4
4.1+6.1	Red	Support	2	Blue	Support	2	9.21	Unknown	4
4.1+6.1	Red	Support	2	Blue	Internal	2	9.29	Unknown	4
4.1+6.1	Red	Support	2	Blue	Leaf	1	9.4	Unknown	4
4.1+6.1	Red	Leaf	1	Blue	Support	2	9.6	Red	4
4.1+6.2	Red	Leaf	1	Red	Support	2	9.10	Unknown	2
4.1+6.2	Red	Support	2	Red	Leaf	1	9.14	Unknown	4
4.1+6.2	Red	Leaf	1	Red	Support	3	9.19	Unknown	2
4.1+6.2	Red	Leaf	1	Red	Leaf	1	9.2	Unknown	2
4.1+6.2	Red	Support	2	Red	Support	2	9.22	Unknown	2
4.1+6.2	Red	Leaf	1	Red	Internal	2	9.26	Unknown	2
4.1+6.2	Red	Support	2	Red	Support	3	9.31	Unknown	2
4.1+6.2	Red	Support	2	Red	Internal	2	9.38	Unknown	2
4.1+6.2	Red	Leaf	1	Red	Leaf	1	9.6	Red	4
4.1+6.2	Red	Support	2	Red	Leaf	1	9.7	Unknown	2
4.1+6.3	Red	Leaf	1	Blue	Leaf	1	9.11	Red	2
4.1+6.3	Red	Support	2	Blue	Leaf	1	9.12	Unknown	4
4.1+6.3	Red	Leaf	1	Blue	Support	2	9.14	Unknown	4
4.1+6.3	Red	Support	2	Blue	Leaf	1	9.27	Red	2
4.1+6.3	Red	Leaf	1	Blue	Support	3	9.29	Unknown	2
4.1+6.3	Red	Support	2	Blue	Support	2	9.30	Unknown	4

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
4.1+6.3	Red	Leaf	1	Blue	Leaf	1	9.4	Unknown	4
4.1+6.3	Red	Support	2	Blue	Support	3	9.43	Unknown	2
4.1+6.4	Red	Support	2	Blue	Leaf	1	9.24	Unknown	8
4.1+6.4	Red	Support	2	Blue	Support	3	9.39	Unknown	4
4.1+6.4	Red	Leaf	1	Blue	Support	3	9.44	Unknown	4
4.1+6.4	Red	Leaf	1	Blue	Leaf	1	9.8	Unknown	8
4.1+6.5	Red	Support	2	Black	Leaf	1	9.16	Unknown	2
4.1+6.5	Red	Leaf	1	Black	Support	2	9.20	Unknown	2
4.1+6.5	Red	Leaf	1	Black	Leaf	1	9.21	Unknown	6
4.1+6.5	Red	Support	2	Black	Leaf	1	9.30	Unknown	6
4.1+6.5	Red	Leaf	1	Black	Support	4	9.31	Unknown	2
4.1+6.5	Red	Support	2	Black	Support	2	9.37	Unknown	2
4.1+6.5	Red	Support	2	Black	Support	4	9.45	Unknown	2
4.1+6.5	Red	Leaf	1	Black	Leaf	1	9.5	Unknown	2
4.1+6.6	Red	Leaf	1	Red	Leaf	1	9.15	Unknown	10
4.1+6.6	Red	Leaf	1	Red	Support	5	9.32	Unknown	2
4.1+6.6	Red	Support	2	Red	Leaf	1	9.36	Unknown	10
4.1+6.6	Red	Support	2	Red	Support	5	9.46	Unknown	2
4.2+4.2	Blue	Support	3	Blue	Support	3	7.11	Black	1
4.2+4.2	Blue	Leaf	1	Blue	Leaf	1	7.4	Blue	9
4.2+4.2	Blue	Leaf	1	Blue	Support	3	7.9	Black	6
4.2+5.1	Blue	Support	3	Red	Support	2	8.14	Black	2
4.2+5.1	Blue	Leaf	1	Red	Internal	2	8.16	Blue	3
4.2+5.1	Blue	Leaf	1	Red	Leaf	1	8.2	Blue	6
4.2+5.1	Blue	Support	3	Red	Internal	2	8.21	Black	1
4.2+5.1	Blue	Support	3	Red	Leaf	1	8.6	Blue	2
4.2+5.1	Blue	Leaf	1	Red	Support	2	8.8	Blue	6
4.2+5.2	Blue	Leaf	1	Red	Support	2	8.13	Blue	3
4.2+5.2	Blue	Support	3	Red	Leaf	1	8.15	Red	2
4.2+5.2	Blue	Leaf	1	Red	Support	3	8.17	Black	3
4.2+5.2	Blue	Support	3	Red	Support	2	8.19	Black	1
4.2+5.2	Blue	Support	3	Red	Support	3	8.22	Black	1
4.2+5.2	Blue	Leaf	1	Red	Leaf	1	8.4	Blue	3
4.2+5.2	Blue	Leaf	1	Red	Leaf	1	8.7	Red	6
4.2+5.2	Blue	Support	3	Red	Leaf	1	8.9	Black	1
4.2+5.3	Blue	Support	3	Red	Leaf	1	8.18	Black	4
4.2+5.3	Blue	Leaf	1	Red	Support	4	8.19	Black	3
4.2+5.3	Blue	Support	3	Red	Support	4	8.23	Black	1
4.2+5.3	Blue	Leaf	1	Red	Leaf	1	8.9	Black	12
4.2+6.1	Blue	Support	3	Blue	Support	2	9.15	Unknown	2
4.2+6.1	Blue	Leaf	1	Blue	Leaf	1	9.2	Unknown	6
4.2+6.1	Blue	Leaf	1	Blue	Internal	2	9.26	Unknown	6

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
4.2+6.1	Blue	Support	3	Blue	Internal	2	9.31	Unknown	2
4.2+6.1	Blue	Support	3	Blue	Leaf	1	9.5	Unknown	2
4.2+6.1	Blue	Leaf	1	Blue	Support	2	9.8	Unknown	6
4.2+6.2	Blue	Leaf	1	Red	Leaf	1	9.10	Unknown	6
4.2+6.2	Blue	Leaf	1	Red	Support	2	9.13	Red	3
4.2+6.2	Blue	Support	3	Red	Leaf	1	9.20	Unknown	2
4.2+6.2	Blue	Support	3	Red	Support	2	9.25	Unknown	1
4.2+6.2	Blue	Leaf	1	Red	Leaf	1	9.3	Unknown	3
4.2+6.2	Blue	Support	3	Red	Support	3	9.32	Unknown	1
4.2+6.2	Blue	Leaf	1	Red	Internal	2	9.33	Unknown	3
4.2+6.2	Blue	Support	3	Red	Internal	2	9.39	Unknown	1
4.2+6.2	Blue	Leaf	1	Red	Support	3	9.44	Unknown	3
4.2+6.2	Blue	Support	3	Red	Leaf	1	9.9	Unknown	1
4.2+6.3	Blue	Support	3	Blue	Leaf	1	9.16	Unknown	2
4.2+6.3	Blue	Leaf	1	Blue	Leaf	1	9.17	Unknown	3
4.2+6.3	Blue	Leaf	1	Blue	Support	2	9.24	Unknown	6
4.2+6.3	Blue	Support	3	Blue	Leaf	1	9.34	Unknown	1
4.2+6.3	Blue	Support	3	Blue	Support	2	9.36	Unknown	2
4.2+6.3	Blue	Leaf	1	Blue	Internal	3	9.38	Unknown	3
4.2+6.3	Blue	Support	3	Blue	Internal	3	9.45	Unknown	1
4.2+6.3	Blue	Leaf	1	Blue	Leaf	1	9.7	Unknown	6
4.2+6.4	Blue	Leaf	1	Blue	Leaf	1	9.13	Red	12
4.2+6.4	Blue	Support	3	Blue	Leaf	1	9.28	Unknown	4
4.2+6.4	Blue	Leaf	1	Blue	Support	3	9.35	Unknown	6
4.2+6.4	Blue	Support	3	Blue	Support	3	9.41	Unknown	2
4.2+6.5	Blue	leaf	1	Black	Leaf	1	9.22	Unknown	9
4.2+6.5	Blue	Support	3	Black	Leaf	1	9.23	Unknown	1
4.2+6.5	Blue	leaf	1	Black	Support	2	9.28	Unknown	3
4.2+6.5	Blue	Support	3	Black	Leaf	1	9.37	Unknown	3
4.2+6.5	Blue	Leaf	1	Black	Support	4	9.39	Unknown	3
4.2+6.5	Blue	Support	3	Black	Support	2	9.40	Unknown	1
4.2+6.5	Blue	Support	3	Black	Support	4	9.46	Unknown	1
4.2+6.5	Blue	leaf	1	Black	Leaf	1	9.9	Unknown	3
4.2+6.6	Blue	Leaf	1	Red	Leaf	1	9.25	Unknown	15
4.2+6.6	Blue	Support	3	Red	Leaf	1	9.40	Unknown	5
4.2+6.6	Blue	Leaf	1	Red	Support	5	9.41	Unknown	3
4.2+6.6	Blue	Support	3	Red	Support	5	9.47	Unknown	1
5.1+5.1	Red	leaf	1	Red	Leaf	1	9.1	Unknown	4
5.1+5.1	Red	Leaf	1	Red	Support	2	9.6	Red	8
5.1+5.1	Red	Support	2	Red	Support	2	9.19	Unknown	4

RAG Trees to be Merged	DATA FROM THE 1^{st} TREE			DATA FROM THE 2^{nd} TREE			RESULTS		
	Tree Color	Vert. Type	deg (v)	Tree Color	Vert. Type	deg (v)	RAG Graph	Tree Color	Total Graphs
5.1+5.1	Red	leaf	1	Red	Internal	2	9.11	Red	4
5.1+5.1	Red	Leaf	1	Red	Internal	2	9.29	Unknown	4
5.1+5.1	Red	Internal	2	Red	Internal	2	9.42	Unknown	1
5.1+5.2	Red	Support	2	Red	Leaf	1	9.10	Unknown	2
5.1+5.2	Red	Support	2	Red	Leaf	1	9.14	Unknown	4
5.1+5.2	Red	Internal	2	Red	Leaf	1	9.17	Unknown	1
5.1+5.2	Red	Leaf	1	Red	Leaf	1	9.2	Unknown	2
5.1+5.2	Red	Leaf	1	Red	Support	3	9.21	Unknown	2
5.1+5.2	Red	Internal	2	Red	Leaf	1	9.27	Red	2
5.1+5.2	Red	Support	2	Red	Support	3	9.31	Unknown	2
5.1+5.2	Red	Internal	2	Red	Support	2	9.38	Unknown	1
5.1+5.2	Red	Leaf	1	Red	Leaf	1	9.4	Unknown	4
5.1+5.2	Red	Internal	2	Red	Support	3	9.43	Unknown	1
5.1+5.2	Red	Support	2	Red	Support	2	9.44	Unknown	2
5.1+5.2	Red	Leaf	1	Red	Support	2	9.8	Unknown	2
5.1+5.3	Red	Leaf	1	Red	Support	4	9.15	Unknown	2
5.1+5.3	Red	Support	2	Red	Leaf	1	9.20	Unknown	8
5.1+5.3	Red	Support	2	Red	Support	4	9.32	Unknown	2
5.1+5.3	Red	Internal	2	Red	Leaf	1	9.34	Unknown	4
5.1+5.3	Red	Internal	2	Red	Support	4	9.45	Unknown	1
5.1+5.3	Red	Leaf	1	Red	Leaf	1	9.5	Unknown	8
5.2+5.2	Red	Leaf	1	Red	Leaf	1	9.12	Unknown	4
5.2+5.2	Red	Support	2	Red	Leaf	1	9.13	Red	2
5.2+5.2	Red	Support	3	Red	Leaf	1	9.22	Unknown	2
5.2+5.2	Red	Leaf	1	Red	Support	2	9.24	Unknown	4
5.2+5.2	Red	Leaf	1	Red	Leaf	1	9.3	Unknown	1
5.2+5.2	Red	Leaf	1	Red	Support	3	9.30	Unknown	4
5.2+5.2	Red	Support	2	Red	Support	2	9.35	Unknown	1
5.2+5.2	Red	Support	3	Red	Support	2	9.39	Unknown	2
5.2+5.2	Red	Support	3	Red	Support	3	9.45	Unknown	1
5.2+5.2	Red	Leaf	1	Red	Leaf	1	9.7	Unknown	4
5.2+5.3	Red	Leaf	1	Red	Leaf	1	9.16	Unknown	8
5.2+5.3	Red	Leaf	1	Red	Support	4	9.25	Unknown	1
5.2+5.3	Red	Support	2	Red	Leaf	1	9.28	Unknown	4
5.2+5.3	Red	Leaf	1	Red	Support	4	9.36	Unknown	2
5.2+5.3	Red	Support	3	Red	Leaf	1	9.37	Unknown	4
5.2+5.3	Red	Support	2	Red	Support	4	9.41	Unknown	1
5.2+5.3	Red	Support	3	Red	Support	4	9.46	Unknown	1
5.2+5.3	Red	Leaf	1	Red	Leaf	1	9.9	Unknown	4
5.3+5.3	Red	Leaf	1	Red	Leaf	1	9.23	Unknown	16
5.3+5.3	Red	Support	4	Red	Leaf	1		Unknown	8
5.3+5.3	Red	Support	4	Red	Support	4	9.47	Unknown	1

VITA

DENISE KOESSLER

Education:	B.A. Mathematics, The College of Wooster, Wooster, Ohio 2008 M.S. Mathematics (Graph Theory), East Tennessee State University Johnson City, Tennessee 2010
Professional Experience:	National Science Foundation Graduate GK-12 Fellow, East Tennessee State University Johnson City, Tennessee, 2008-2010
Publications:	D. Koessler, D. Knisley, J. Knisley, and T. W. Haynes "A Predictive Model for Secondary RNA Structure: Using Graph Theory and a Neural Network" Submitted March 2010.
	T. W. Haynes, K. R. S. Holmes, L. Sewell, and D. Koessler, "'Locating-domination in Complementary Prisms of Paths and Cycles," *CONGRESSUS NUMERANTIUM*, **199** (2009), pp. 45 - 55
	T. W. Haynes, K. R. S. Holmes, and D. Koessler, "Locating-domination in Complementary Prisms," Submitted May 2009.