

Genetic Principles and Applications to the Genus Hemerocallis¹

Terrence P. McGarty

Abstract

This paper is a review paper for the subsequent document on the Genus Hemerocallis. This paper establishes the baseline facts as are currently experimentally known and which are at the heart of understanding the genetics of the Genus Hemerocallis. This paper develops models which will be used elsewhere in the analysis and synthesis of color and patterning of the various hybrids as well as establishing an understanding of the underlying sets of species and their resulting hybrids.

¹ Copyright © Terrence P McGarty 2008 all rights reserved. tmcgarty@telmarc.com , www.telmarc.com Charts and Tables have been used and their source has been referenced. This is covered under the fair use doctrine. The material contained herein is the opinion of the author and no representation is made as to its accuracy. Reliance upon material contained herein is at the sole risk of the user.

Contents

1	INTRODUCTION	3
2	PRELIMINARY CONCEPTS AND DEFINITIONS.....	3
2.1	<i>Chromosomes and Genes</i>	<i>4</i>
2.1.1	Chromosome.....	5
2.1.2	DNA	6
2.1.3	Gene.....	11
2.2	<i>Genotype and Phenotype</i>	<i>13</i>
3	GENETICS.....	20
3.1	<i>Plant Cells</i>	<i>20</i>
3.2	<i>Plant DNA.....</i>	<i>22</i>
3.3	<i>Plant Gene Processes.....</i>	<i>23</i>
3.4	<i>Activators.....</i>	<i>26</i>
3.5	<i>Repressors</i>	<i>29</i>
3.6	<i>Summary of Actions.....</i>	<i>30</i>
4	EXPRESSION ANALYSIS AND IMPLICATIONS	30
4.1	<i>Approach: Engineering versus Science</i>	<i>31</i>
4.2	<i>A Control Paradigm.....</i>	<i>31</i>
4.3	<i>Cell Signaling: Intra and Inter Cell.....</i>	<i>32</i>
5	FLOWER COLOR EXPRESSION	34
5.1	<i>Pathways and Enzymes.....</i>	<i>35</i>
5.2	<i>Anthocyanins</i>	<i>36</i>
5.3	<i>Other Color Elements</i>	<i>38</i>
5.4	<i>Carotenoids</i>	<i>39</i>
5.5	<i>Flavones</i>	<i>40</i>
5.6	<i>Pathways</i>	<i>41</i>
5.6.1	Anthocyanin Pathway.....	42
5.6.2	Carotenoid Pathway.....	46
5.6.3	Flavonol Pathway	47
6	CONCLUSIONS.....	47
7	REFERENCES	47

1 INTRODUCTION

The genetic structure of the genus *Hemerocallis* and its impact on the color and patterning requires an understanding of a few essential facts from the now well understood operations of the gene and the secondary pathways associated with them. This paper is a review of these principles.

Specifically we review the following:

1. Gene structure and operation. This includes the basic Watson and Crick model as is currently understood. The development that we use is a functional model and note one that would be more familiar to the biologist. In all our analyses we will build models of functions and leave the basic principles and their modifications to the bench scientist.
2. Secondary Pathways are introduced and the related gene controls are presented. The secondary pathways which create the chemicals which in turn create colors are discussed in some detail.

This discussion should provide the basic principles to address the other issue we seek to develop.

2 PRELIMINARY CONCEPTS AND DEFINITIONS

We want first to develop some concepts and definitions. To fully understand the genus *Hemerocallis* and to be able to employ the techniques of breeding, one must have a common framework of concepts, the building blocks of the ideas we will develop and employ. This chapter begins that process.

There are several concepts we can begin to define. They are:

Chromosomes and Genes: The essence of understanding and growing new types of *Hemerocallis* is the understanding of the chromosome and gene. The *Hemerocallis* gene is somewhat simple and akin to that of a human. We humans have 22 chromosomes plus a sex chromosome. For a total of 23 chromosomes. The *Hemerocallis* has 11 chromosomes and no sex chromosome. Both generally have chromosomes in pairs, the human has 23 pairs of 46 chromosomes and the *Hemerocallis* has 22 pairs of 11 chromosomes².

Genotype and Phenotype: We all know what a specific plant "looks" like if we see it. We know its color, its size, its shape, and other characters or characteristics which we could then communicate in a somewhat unambiguous fashion to others so that they could in turn say whether they have found the same "type" of plant. In contrast, we can now ascertain the genotype of a plant, at least on the small. We can look at certain gene loci and from them determine what the plant is. In today's world we can use this genetic information perform various analyses which in turn will allow us to "characterize" a specific plant. But we know that no two

² See Kang and Chung (1997) p. 210, *Journal of Plant Research*, Japan. The authors state that they have independently verified this number.

plants have exactly all the same genes, some genes may not be expressed, so they may "look" alike in all aspects, but hidden in sections of their DNA are segments which do not speak but are different.

Species: Just what is a species and what does it mean for us as we proceed through this study. This is the most critical question that we shall pose and we shall spend considerable time discussing its meaning.

Breeding versus Hybridizing: Daylily people consider themselves hybridizers. Agricultural botanists look at breeding, as do say people who raise dogs, horses and the like. What is the difference between breeding and hybridizing and which applies or should apply in this area.

Pollination, Self and Cross: Obtaining variability in a plant means we must work with what is in nature or has already been developed by others. The plant in the wild will pollinate itself or will cross pollinate with others. What do we really mean by these terms and how does that influence the concepts we are trying to develop.

2.1 *Chromosomes and Genes*

Let us start with the chromosome. We will return in some detail to this latter but at this point we want to establish a few basic definitions. The plant has 11 pairs of chromosomes, for a total of 22 chromosomes³.

The Figure below is a graphic of a typical plant cell showing the nucleus and one of the chromosome pairs. This graphic is not at all what one would see in reality but it is typical of the generic elements.

³ See Kang and Chung, 1997, Journal of Plant Research.

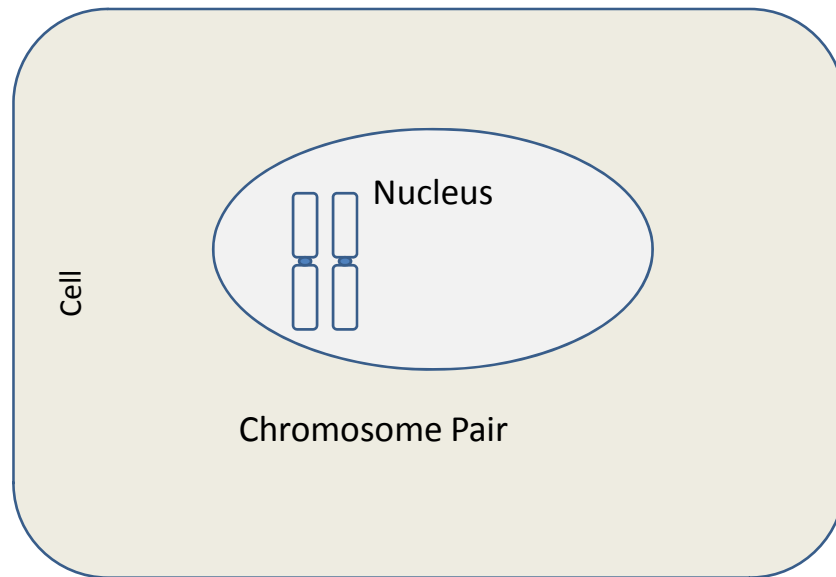


Figure 1 Basic construct of a plant cell.

2.1.1 *Chromosome*

The chromosomes are the collection of DNA which agglomerates together into separate units. They bind together as pairs and it is these pairs which make up the chromosomes we see in the nucleus of a mature cell.

The Figure below depicts the types of possible chromosome combinations we would see in a typical *Hemerocallis*. This is called ploidy, haploid being one chromosome and diploids being pairs of chromosomes.

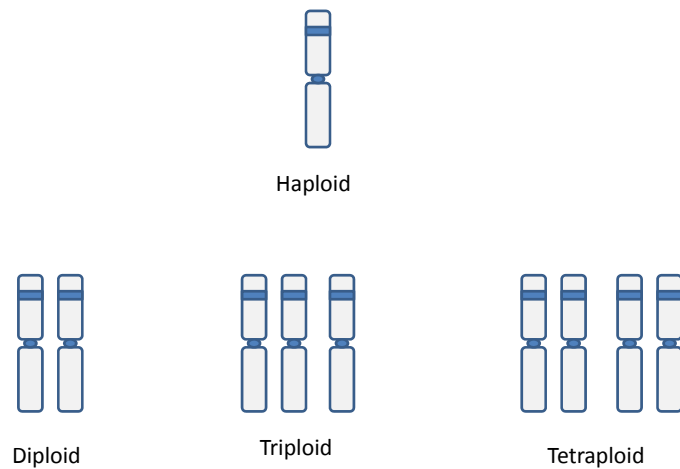


Figure 2 Plant ploidy

The types of ploidy are:

Haploid: The haploid is the single chromosome strand that one may be able to see in the sex cells of a plant. Namely in the pollen or in the ovary cells. The haploid is a single stranded non-banded collection of DNA.

Diploid: The diploid is the prototypical collection of DNA in the mature *Heimerocallis* as is normally found in species and in many hybrids. The diploid is merely two, one from the male and one from the female.

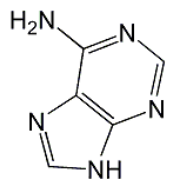
Triploid: This type of three way bonding is found in many *Heimerocallis* which do not produce sexually such as the *H. fulva* Europa, the common garden variety orange daylily and the doubles we see frequently the *H. fulva* Kwanso and *H. fulva* Flore Pleno. These triploids are not at all readily used in crossing but it has been recorded that from time to time they do manage a cross. The details of the crossing mechanism are not fully understood.

Tetraploid: Since the mid 20th century, with the use of colchicine an alkaloid from the *Colchicum* genus, also used for gout, the creation of tetraploids was possible. Tetraploids have four chromosomes per grouping and thus the nucleus has a total of 44 chromosomes. This is twice the DNA of the normal diploid and this doubling introduces many additional variations which we shall show later.

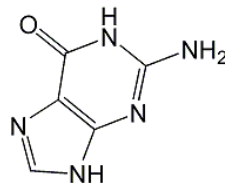
2.1.2 DNA

DNA, deoxyribonucleic acid is the heart of the gene. It is the basis of the code we can understand to determine the relationship between genes and their phenotypic responses.

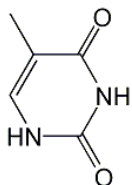
We briefly layout the ideas concerning DNA in this section. DNA is constructed in the following manner. There are four base elements; Adenine (A), Glutamine (G), Tyrosine (T) and Cytosine (C). They are shown below.



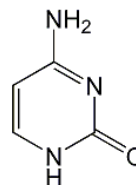
Adenine



Glutamine



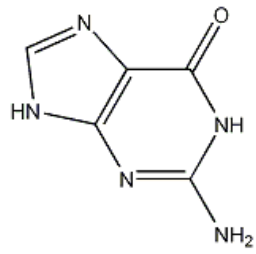
Thymine



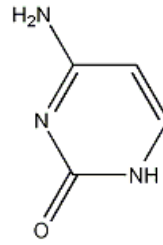
Cytosine

Figure 3 Bases of DNA elements

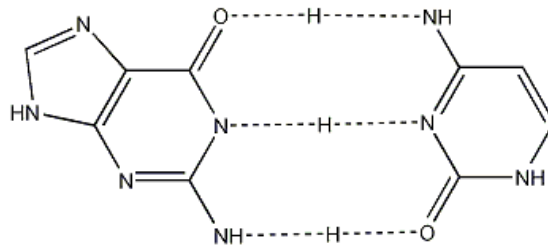
These Base elements can combine in only a specific manner, namely A with T and G with C. These bonds are shown below. This was one of the seminal observations which drove Watson and Crick towards their great discovery. The bonding also is the basis for how these Bases combine in pairs, the Base Pairs, and then how these Base Pairs link up to form the now famous DNA chain.



Guanine



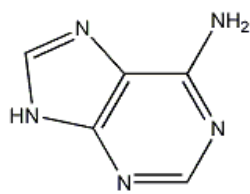
Cytosine



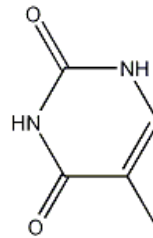
G

C

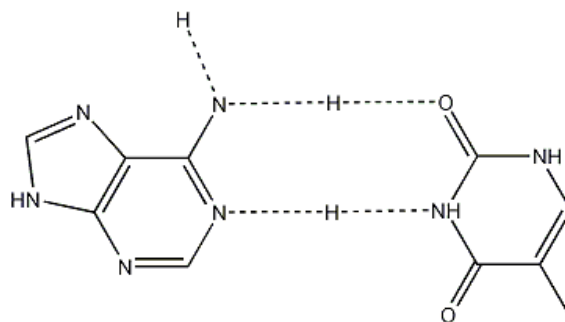
Figure 4 CT Base elements and their bonding.



Adenine



Thymine



A

T

Figure 5 A-T Base Pairing

Now these Base Pairs are connected to sugar molecules, a cyclic ribose, to create a Nucleoside, such as deoxyadenosine. Then the nucleosides are enhanced with a phosphate constellation, a phosphorous molecules surrounded by oxygen and hydrogen. This combination of the nucleoside and the phosphate is called a Nucleotide. It is these nucleotides which connect on a backbone on the outside and in another backbone on the inside to form the DNA molecule. The following Figure shows a Nucleotide connection, we do not show the base pair connections. The Nucleotide has two defined ends; a 3' end which of the OH molecule and the 5' end which is the phosphate. We show these in the following Figure. These ends will play an important part in the generation of the products of DNA.

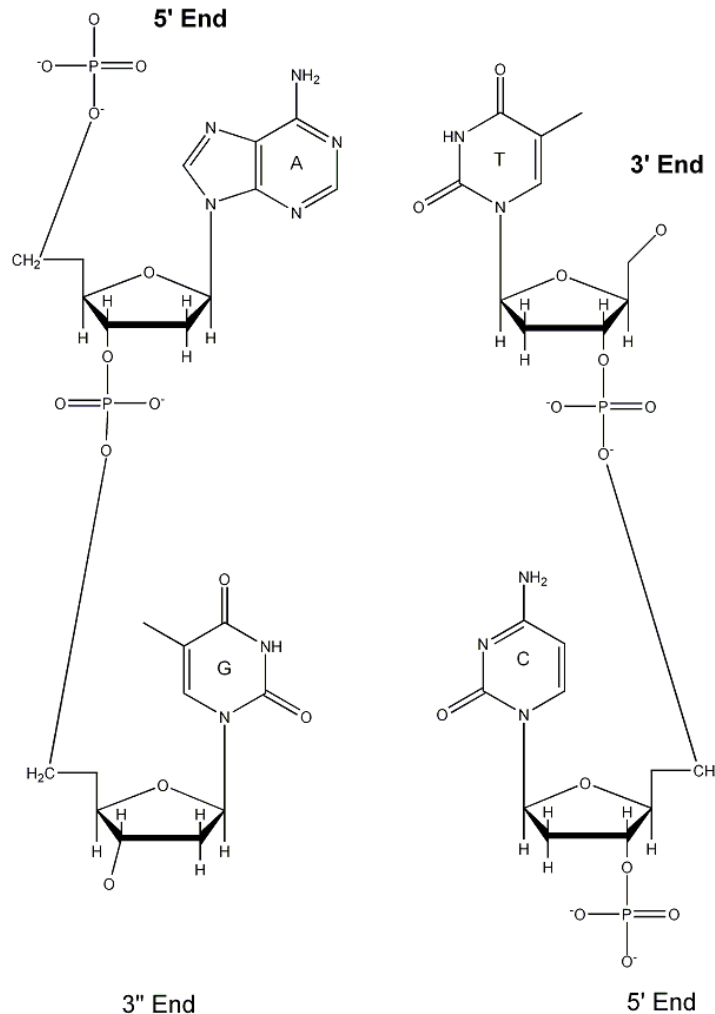


Figure 6 DNA Nucleotides and 3' and 5' Ends.

The nucleotides are then connected into the long DNA wrapped double helix which is generally well known. This is shown below. Our interest will be in the genes themselves and we will look at them in some detail. One of the key questions will be just what is a gene? That will be a challenging question. It will go to the heart of hybridizing. It can be answered in many ways but clearly the simple ideas of Mendel must be revisited.

In the Figure below we set forth a paradigm of the opposite bases and they are lined up in a stretched out set of nucleotides where we are looking solely at the base elements, the A, T, G and C.

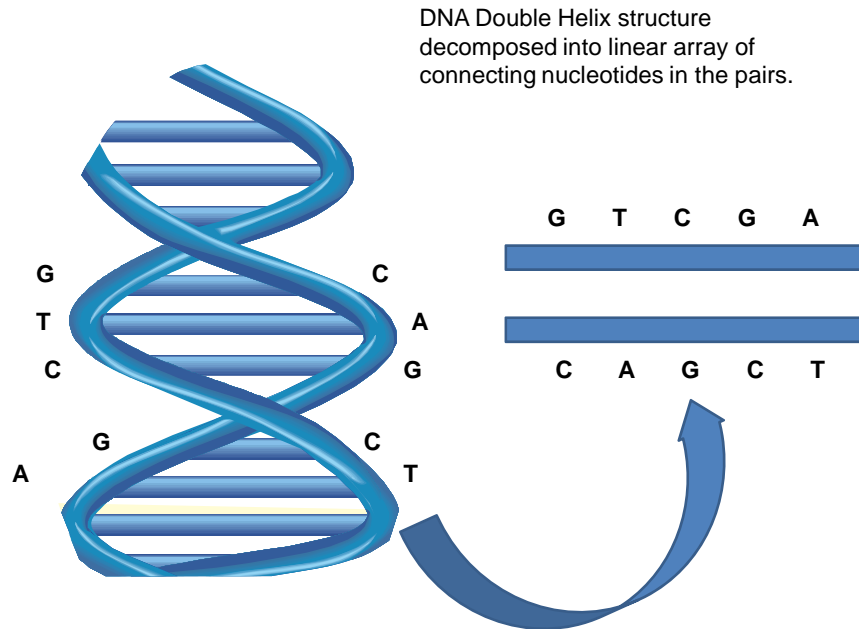
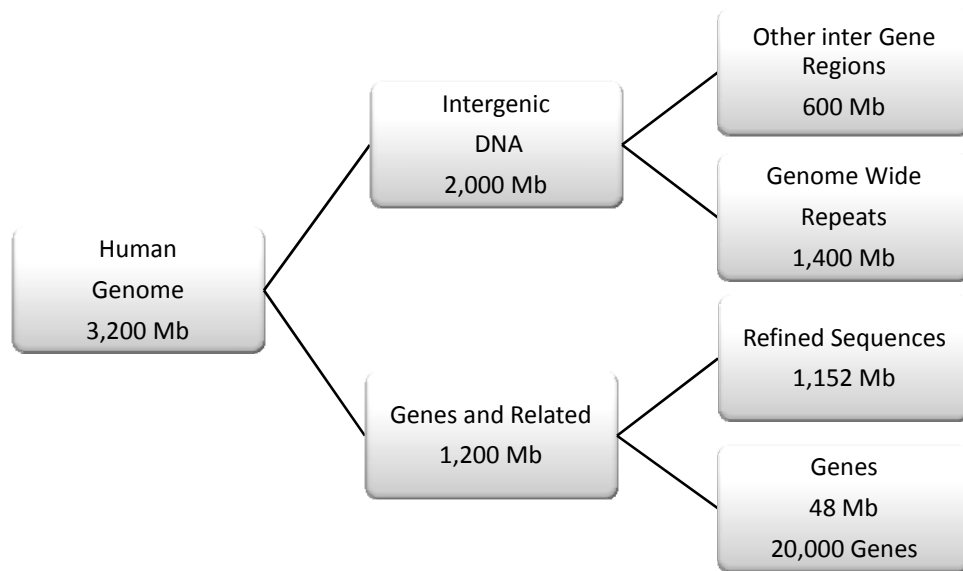


Figure 7 DNA double helix and Base Pair paradigm.

In the human the DNA is of moderate size, about 3,200 Mb, that is 3.3 billion G, T, A, or C. However as shown below the DNA is broken down into many small elements. The actual operating genes constitute a mere 48 million bases and this constitute about 20,000 genes. That is an average of 2,400 bases per gene. As we shall see it takes three bases to create one chemical compound on a protein, this there are a total of 800 per protein on average.

The main conclusion is that there is a great deal of what has been called junk DNA. That DNA is useful for identifying people, namely that is used in DNA identification, and it may or may not play great roles in protein generation and gene modulation.



Ref: Watson et al, Molec. Bio Gene 5th Ed, p. 137

2.1.3 Gene

The gene is the fundamental building block of any living creature. It is not the single expressive element to control a phenotype, it may contribute to that control but it is not the one to one element in the process. Thus a red flower may be controlled by several genes and in addition those genes may be affected by several epi-genetic factors ranging from the environment to other genes.

The human is now thought to have about 20,488 genes⁴. Not a large number and greatly lower than what literally all the experts thought before the human gene was fully analyzed. Many experts had guessed that there were well above 300,000 genes in the human. The Human genome is composed of slightly more than 3 Billion base pairs, combinations of G, T, C or A. The Hemerocallis genome is approximately 4 Billion base pairs. The number of active genes in Hemerocallis is at this time unknown. But it is close in size to the human genome.

The simple construct of a gene is shown below. It is a collection of DNA bases which combine together in terms of the effect. We show in the Figure the Introns, namely the unused DNA bases, and the exons, the used DNA bases. The exons are "combined" to effect what a gene does.

⁴ See Pennisi, Working the (Gene Count) Numbers: Finally, a Firm Answer? SCIENCE Vol 316 25 May 2007
<http://www.sciencemag.org/cgi/content/full/316/5828/1113a?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&fulltext=gene+count&searchid=1&FIRSTINDEX=0&resourcetype=HWCIT>

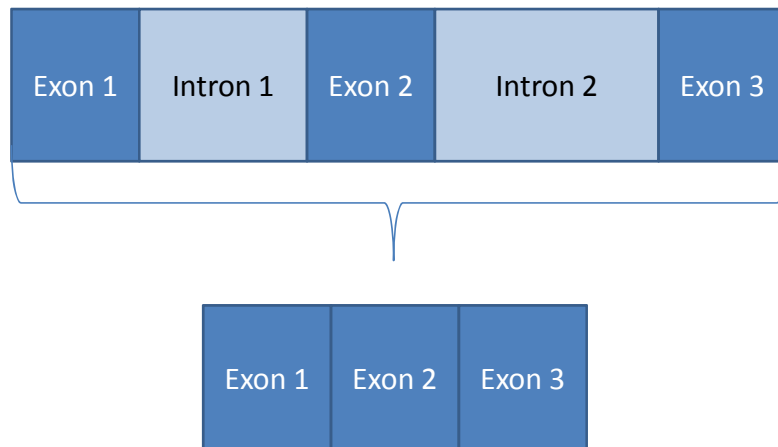


Figure 8 Genes and Introns and Exons

What then is a gene? For our purposes and to be consistent with contemporary understanding we define a gene as:

"A gene is a collection of DNA bases which when combined in a determinable manner can express the combination of bases via the production of some effect upon the cell and potentially its surrounding environment. A gene is an expressible collection of base pairs, when acting in concert, in the internal environment of a cell."

Thus we understand a gene by its effects, not just by its structure. Its effects may be complex. It may produce some RNA, and in turn a protein, it may activate or suppress another gene, or it may be the basis for creating a new gene in this construct. Based upon what we know and understand today, a gene is not some well defined coherent set of contiguous DNA. Genes can even be created on the fly within the cell based upon the environment that is if we define a gene by what it creates and affects.

The classic paradigm for DNA influence is shown below. Namely that DNA generates RNA via transcription and RNA generates proteins via translation. We will not get into further details other than saying that this process has many sub elements which will be regarded in further detail latter.

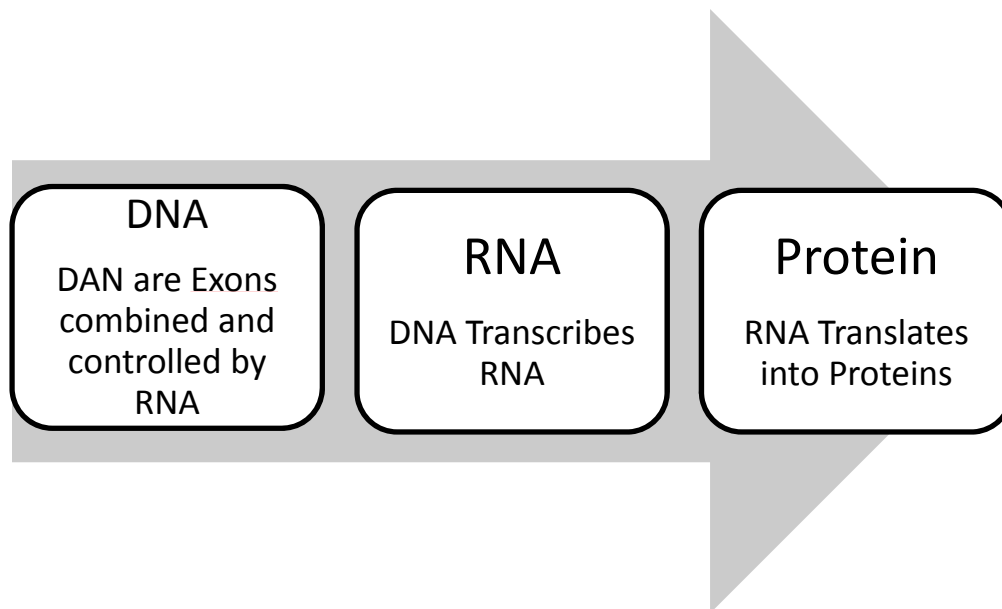


Figure 9 Classic DNA Paradigm

The above understanding of the gene and its relationship to its environment states that there exists a gene, a construct, which uniquely generate an RNA strand, which in turn uniquely generates a protein. We now know that these are all subject to further analysis. For example, the gene is not just a connected set of DNA bases, it is a set of exons, which may be combined in a sequence, or may even be broken or reassembles. Thus the gene is determined by what it does, not by any unique set of base pairs.

The protein that results from the above model is then related to some phenotypic response.

2.2 *Genotype and Phenotype*

Phenotypes are what we see, smell, hear, touch, taste; they are the interactions between some creatures, in our case a plant, which we may use to identify the plant. In the genus *Hemerocallis*, the phenotype may be the color, color patterns, size, time of bloom, odor, texture of the flower, and other definable characteristics that we see when we observe the plant.

Genotype is what the gene has as specific content, its specific DNA. The production of a phenotype is frequently driven by the expression of a gene. The gene "expresses" itself in a very special manner. The DNA is wrapped in tight coils.

The model we will build upon appears as in the Figure below. This is the canonical model for gene expression. We assume that there is some collection of secondary pathways, and that these pathways result in chemical products that are directly related to a phenotype; a darker red flower, a longer leaf, a taller scape. That these pathways are modulated in some manner by proteins generated from within a cell. That the proteins are the result of some entity called a gene. That the gene can be an assembly of bases and the gene may itself be modulated up or down by activator or repressor proteins respectively generated by other genes or even the same gene. Thus

we model the cell as a dynamic system and further we argue that this system has certain random elements which we shall include latter.

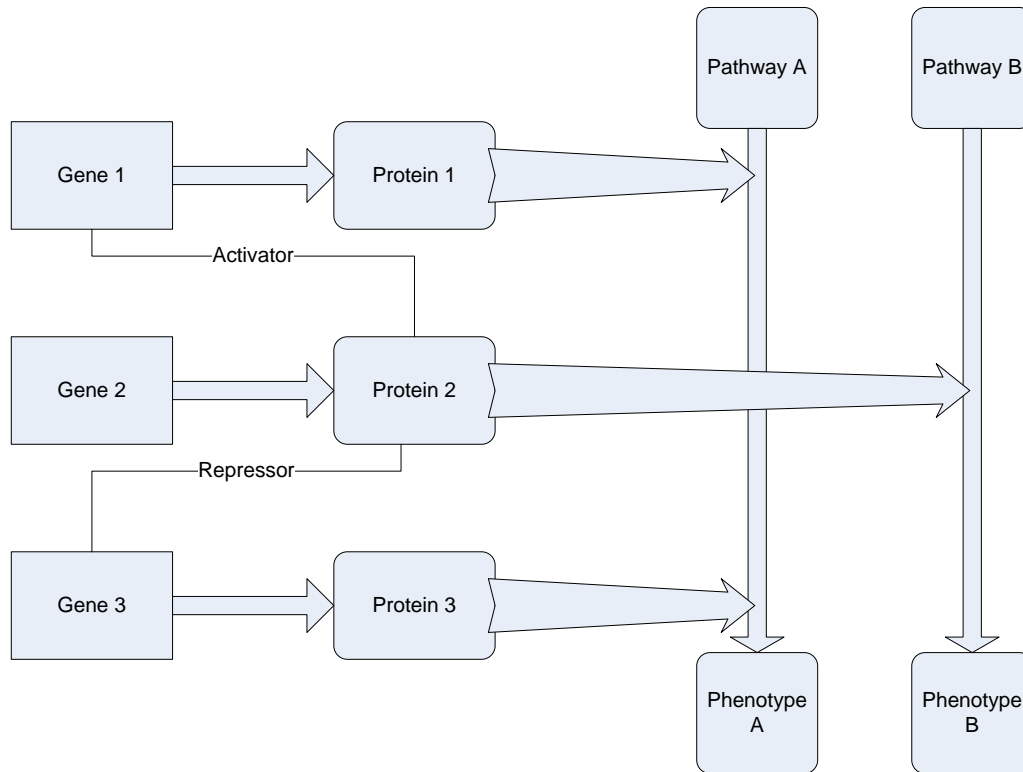


Figure 10 Canonical Model for Gene Expression

It is the output of this genetic process that we get the plant in its full temporal and spatial existence.

The above model of the gene is one in which we see the beginnings of some form of feedback. We see the activator and repressor genes as the basis for this element. However this may be expanded even further, We show this below. Note we show that the Gene K can be influenced by other Genes, as well as the products of the pathways as well as by the environment. The Environment can modulate the pathway which by being fed back to another controlling gene can then modulate the activating gene. This process is a complex process and exceeds what we would have imagined from the simple Mendellian gene theory.

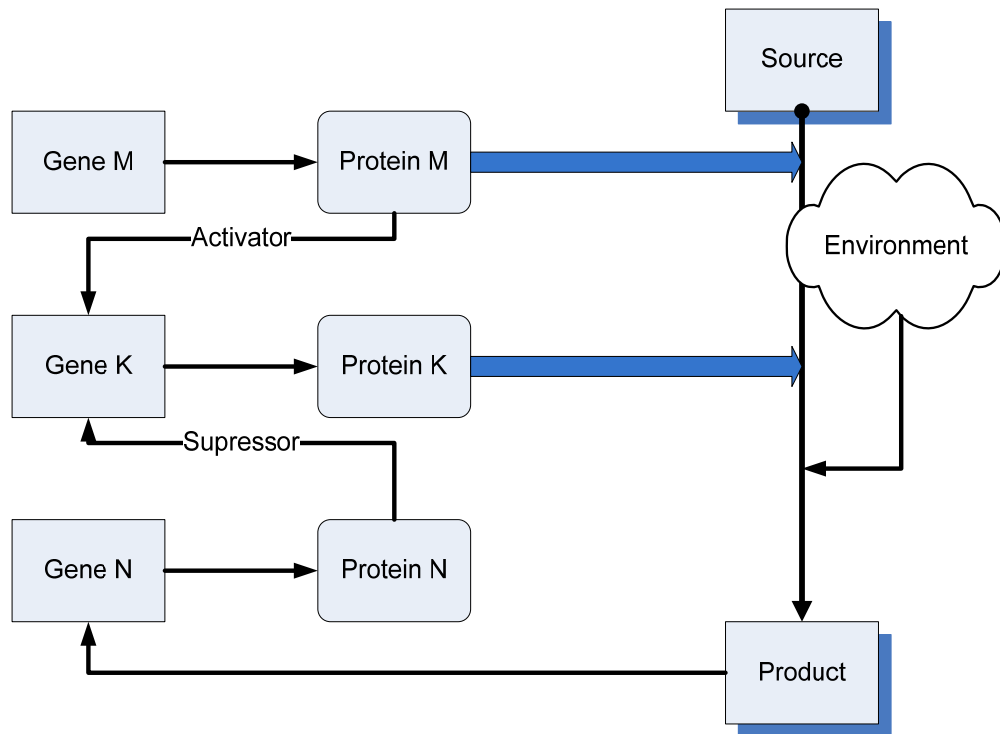


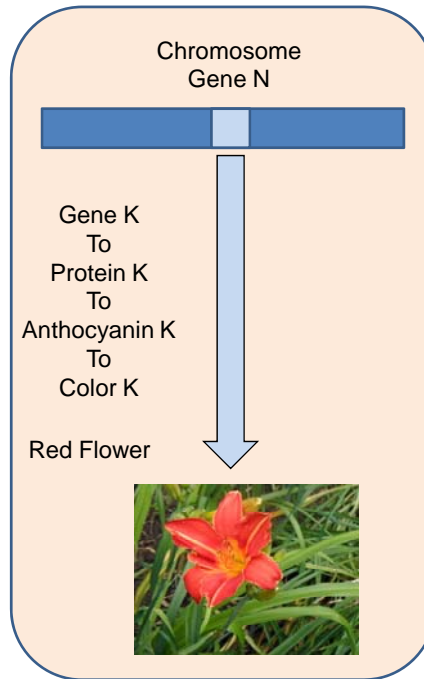
Figure 11 Dynamic Gene Model

Now back one again to the Mendellian Gene model. Although Mendel and his model was not so rigidly simple, for he did admit some other influences as well as variation, we will call the simple Gene and Phenotype combination the Mendel Model. Namely in this model we assume the existence of a Gene and then we further assume that there is some phenotypic characteristic such as flower color which maps one to one onto this gene. One gene and one phenotypic character. The phenotypic characters further have countable and discrete values. The flower is red, yellow, and green. There are no blends and there are a limited numbers. Then there is a gene for red, a gene for yellow and a gene for green. The gene is at the same place on the chromosome and the gene just somehow changes to produce a different color. In addition the genes are dominant in some order. That is if there is one red gene, of the two on the chromosome, then we get red, if not a red but a yellow we get yellow, and we get green if and only if there are all green genes, namely two.

**Mendel's
World View:**

One gene to
one protein to
one phenotype.

Qualitative
Genetics
Namely
One to One



$$G_K \rightarrow P_K$$

Now there is a second model, based upon our understanding of DNA and the Watson Crick world. However this model goes well beyond the simple Watson Crick model. Here we assume we have long segments of DNA with many exons and many more introns. The gene as we know it is the result of the cellular processes which assemble the exons into a block of DNA which RNA will use to in turn generate a protein. In reality what happens is that the exons may be recombined to generate RNA in a variety of fashions. The result of that process, as well as the dynamic model we depicted above is that the phenotypic characteristic, say leaf length or width, or date of first bloom, takes on the character of a random variable. It has a set of values whose probability distribution may be of some form. We use as an example a standard Gaussian curve. This is shown below.

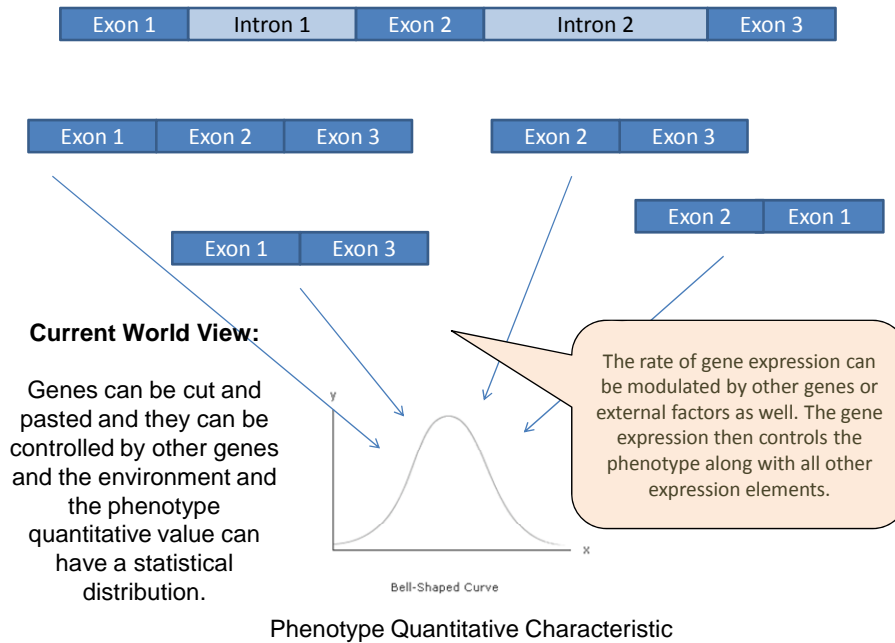


Figure 12 Current View of Genetic Control

For example we show in the following Table results from Hasegawa et al as modified:

Table 1. Morphological traits (mean and standard deviation) of *Hemerocallis citrina* and *Hemerocallis fulva*, their F1 hybrids, and individuals in the hybrid population The standard deviation is given in parentheses

	H fulva	H citrina	F1
No. of scapes	72	74	55
Flower tube length (mm)	32.40	45.28	32.27
Petal length (mm)	92.51	81.13	78.60
Petal width (mm)	15.42	14.50	16.01
Stamen length (mm)	77.02	68.04	64.39
Pistil length (mm)	97.34	76.19	77.73

We now use the data from Hasegawa and present their curves shown the statistical distribution of scape and dimension.

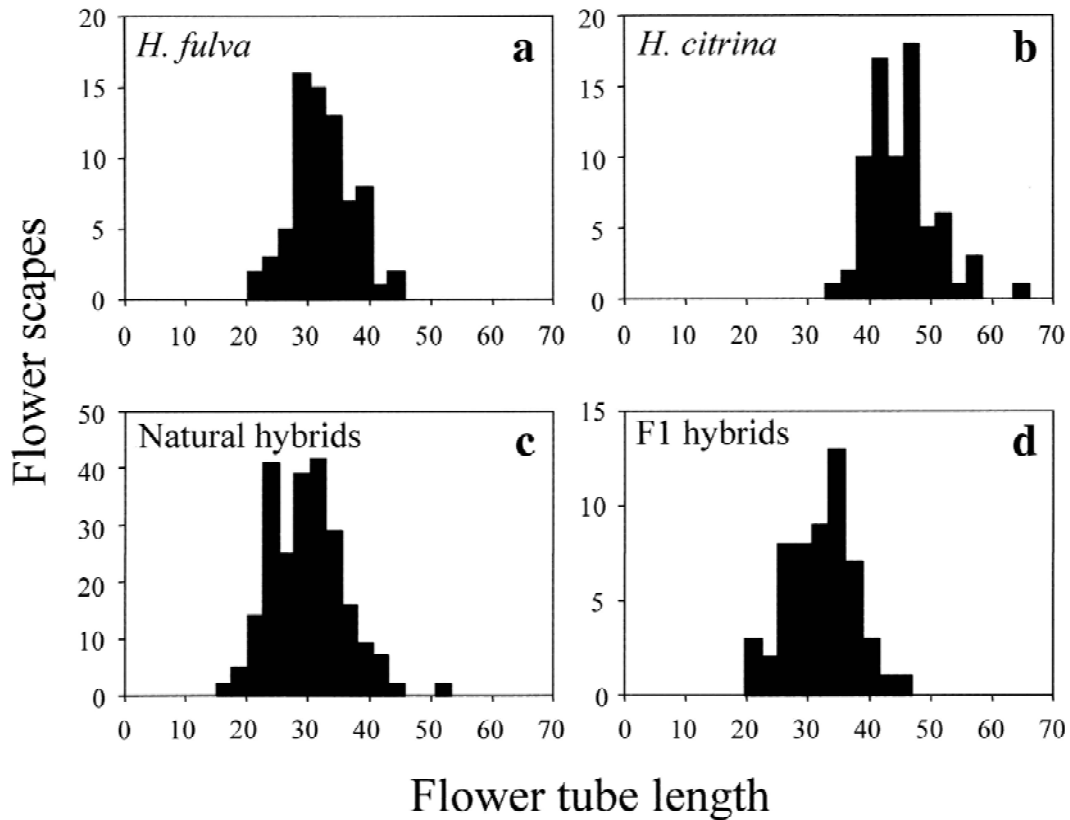


Figure 13 From Hasegawa et al, Distribution of *H. fulva* and *H. citrina*

In a further step using Hasegawa and their study of *H. fulva* and *H. citrina* we can see the distribution in flowering time of the two species as recorded by the authors. Clearly several things can be observed. First, the time of bloom is clear bimodal and this form of separation will enhance the separateness of the species. The pollinators further reinforce the separation. Second, and in line with our above discussion, the blooming is spread out over a large period of time. This gives a probability distribution, albeit not a Gaussian one,

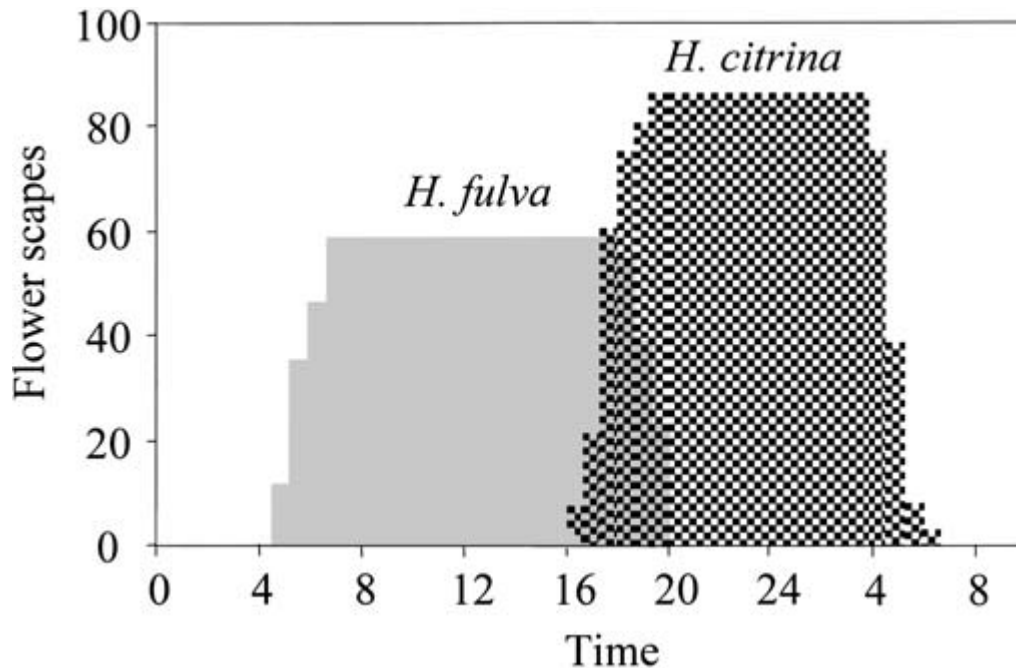
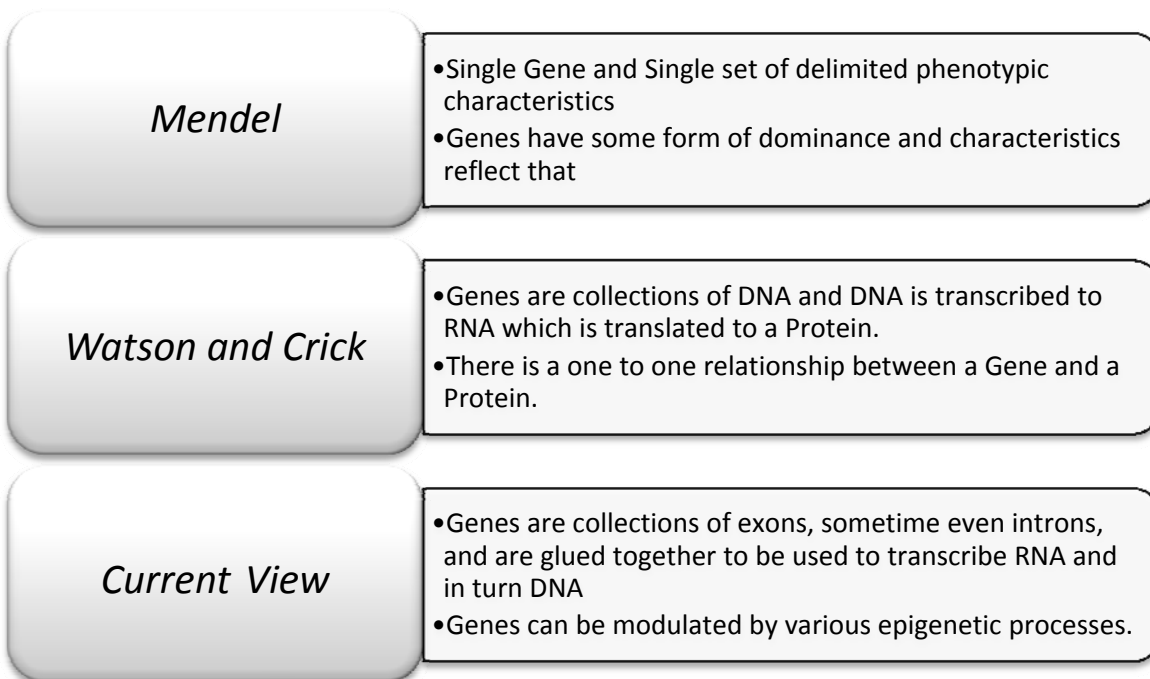


Figure 14 From Hasegawa et al, Flowering time of H fulva and H citrina

The views of gene impact are summarized in the following Figure. We shall use this model as a way to better understand how one can better seek hybridizing opportunities.



3 GENETICS

In this section we present an overview of the classic Mendellian analysis.⁵ The Mendellian analysis makes classic assumptions which prevailed until the advent of the Watson and Crick model, and even slightly beyond. In fact many breeding programs build upon a Mendellian approach. We argue that such an approach is partially correct but lacks most of the key elements which must be considered.

In this section we briefly review the molecular genetics of a plant cell. We do not get into any significant details but merely review the elements which we can use later in developing the mathematical models for plant regulation. As we have shown in the previous section, plant colors are the result of the expression of three types of secondary plant cell products; anthocyanins, flavones and carotenoids. We have focused mainly on the anthocyanins but have shown the details on all three. What we focused on is that the production of any one of these is a result of a specific pathway and that the production in that pathway is controlled by a set of enzymes. The enzymes are proteins produced within the cell. The proteins are the result of the expression of a set of genes.

In this section we now by reviewing the current understanding of plant cell micro genetics show that the proteins are expressed by the normal process understood since Watson and Crick's seminal work and that there are factors which activate their production, indeed enhance their production, or repress their production. These are the activators or repressor proteins. The activator and repressor proteins are in effect other genes expressing themselves. We will combine the last section with the results in this section to affect a dynamic system model for plant color generation in the next section.

What will be critical to understand here is that we just want to place the process of activators and repressors in context. We discuss in the next section what our overall design approach will be; that of an engineering model development and not a detailed understanding at the cell level. Frankly, we are not interested in the lower level detail, only gross modeling of cells, genes, and their proteins. They will become the inputs, outputs and control mechanisms of our design approach.

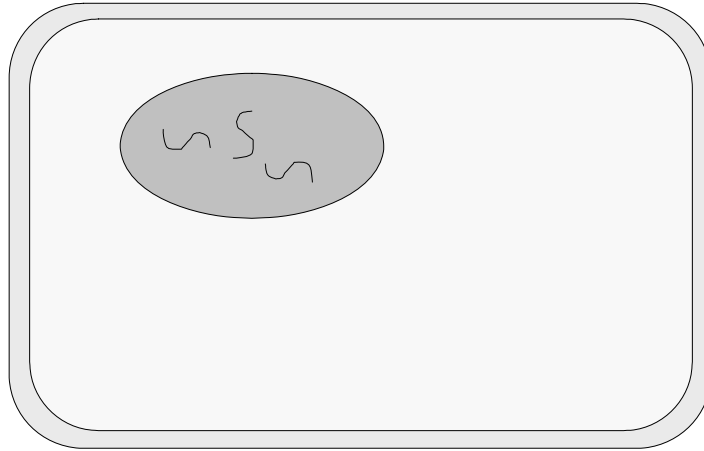
3.1 *Plant Cells*

Plant cells are a class of eukaryotic cells which are characterized primarily by have a rigid cell wall. In almost all other ways they are similar to animal cells. Plants generate all of the amino acids they need for protein generation unlike animal cells but other than that, for our purposes, they function very much the same. Thus as we develop a model for plants the model has no restrictions in its applications to animals as well.

The typical plant cell is shown below. The cell wall and the nucleus are depicted.

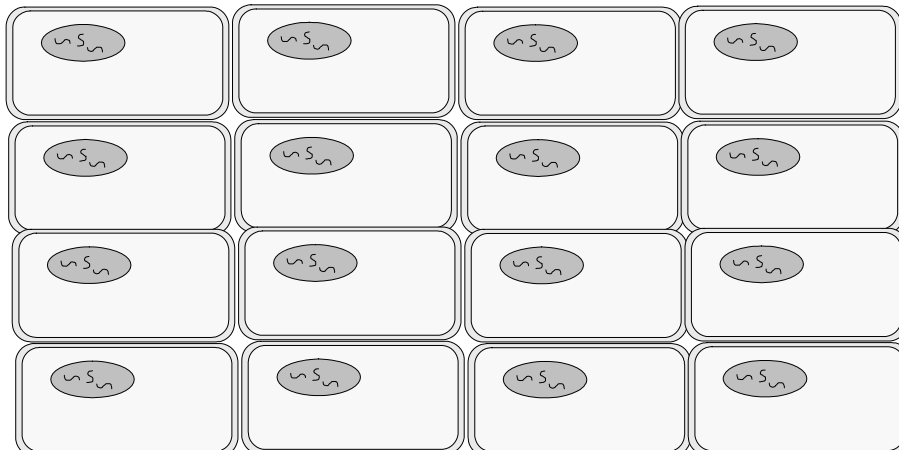
⁵ See Griffiths. This is an excellent overview of genetic analysis.

Plant Cell



When we look at a collection of plant cells they appear as below. They are aligned and interconnect via various channels. Unlike animal cells plant cells have a much more rigid structure due to the cell wall however the general intercell signaling is identical.

Plant Cell Matrix

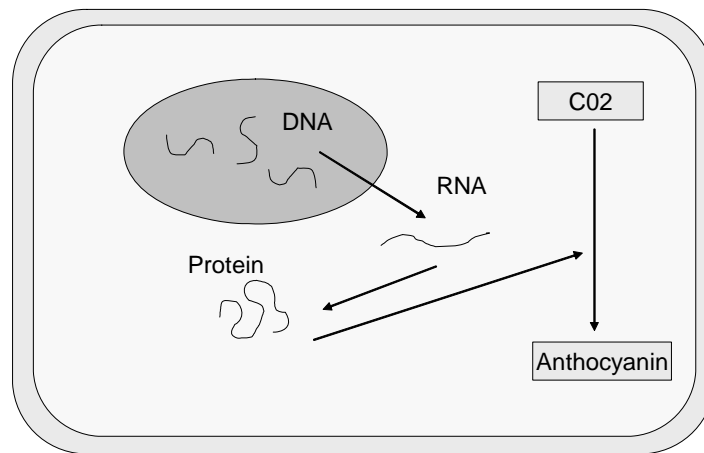


Our interest will be to focus on both the intracell and intercell signaling and control of the pathways.

3.2 Plant DNA

Plant DNA processes are almost identical to those of animals. The graphic below summarizes the view we shall take. Each cell has DNA and the DNA uses a mRNA to create proteins. The proteins are then used in the management of the pathways to create the secondary products of the cell, in our case the anthocyanins.

Plant Cell DNA Process



For a single cell the model is quite straight forward. Gene expression causes RNA which causes Protein, which is enzyme in anthocyanin pathway generating the anthocyanin.

We do however want to stress certain issues. There are two extreme views of cells:

Micro/Time View: The micro view looks at a cell at each instant of time and considers what is happening. Is the cell generating a protein and a secondary and if so how and what is the sequence in which this process occurs. It is a focus on a single cell over some time period and we see many things happening.

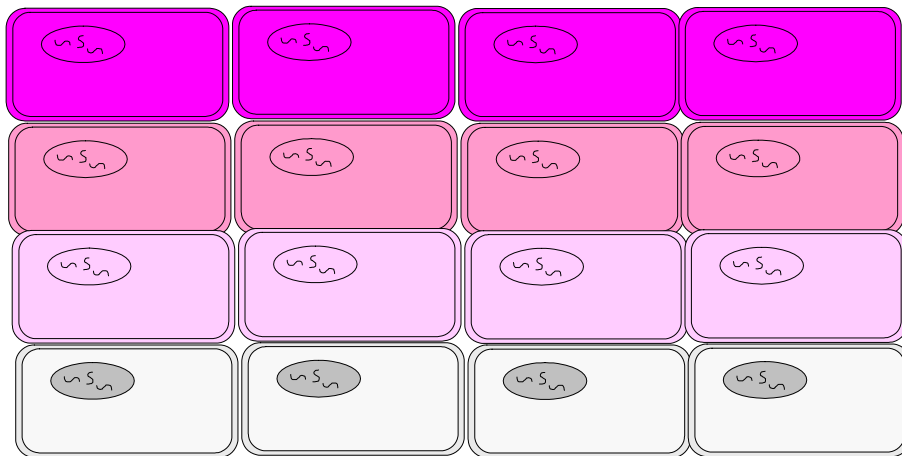
Ensemble View: In this case we look at the cell on average. Namely we say a cell can “on average” produce a protein and can then in turn produce a secondary.

These two views have analogs in mathematical analysis; they are the time averages versus the ensemble average. In mathematical statistics we have the concept of looking at a single cell and time averaging say the concentration of a certain secondary. We know how it is produced and thus over some time window we can look at the average of say pelargonidin and we than measure its average value. In contrast we can take a collection of similar cells and measure the pelargonidin in each cell and take that average. The latter is called the ensemble average. The

equivalence of the two is called the Ergodic Theorem and was developed by Norbert Wiener⁶. The microbiologist typically focuses on the time view. We in this paper will focus on the ensemble view. The latter view will allow us to model, predict and control large collections of cells.

Now the figure below depicts a typical problem we want to understand. Consider an array of cells. Consider that they are arranged in ascending order up the petal of the flower, from base to outer edge. Consider now that at each vertical increase that the cells at the same level all have the same color yet at each level they have a differing shade of color. This implies that the anthocyanin concentrations are different at each level but identical at each cell within a level. We will assume we can understand a single cell from our discussions in the last section, if we understand the pathways and their enzyme controls. Now we ask how does one create a mathematical system model which can “explain” the color patterns we see below. This will be a critical question to answer.

Plant Cell Matrix Colors



How do the cells communicate? Why does one cell generate more anthocyanin than other cells. Why is this not just random? What is the control mechanism?

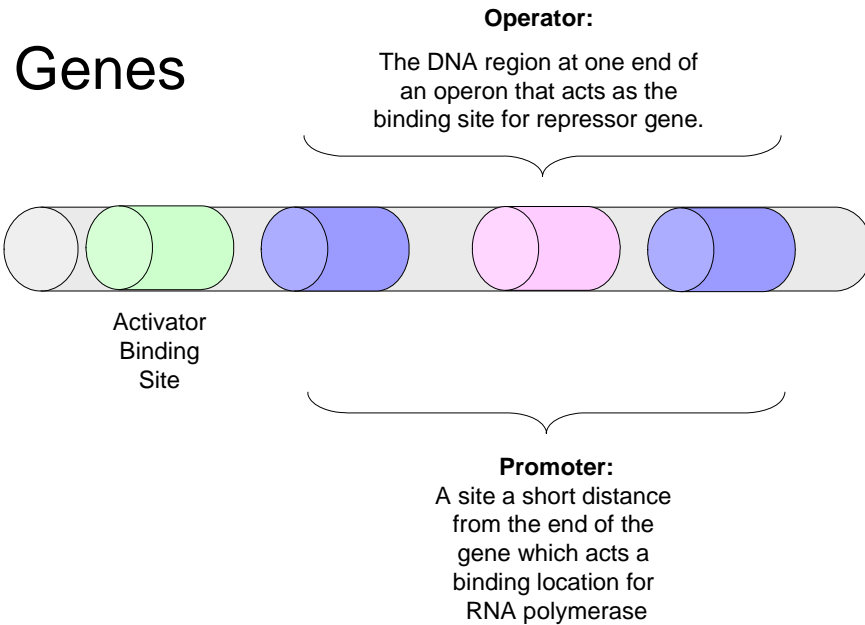
Before we can answer this question we need to delve a bit deeper into the genetics of gene expression.

3.3 Plant Gene Processes

The processes in plant genes are generally identical to those in animal and thus human genes. The figure below shows a typical gene structure along with key sites. This structure shows the

⁶ See McGarty, Stochastic Systems and State Estimation.

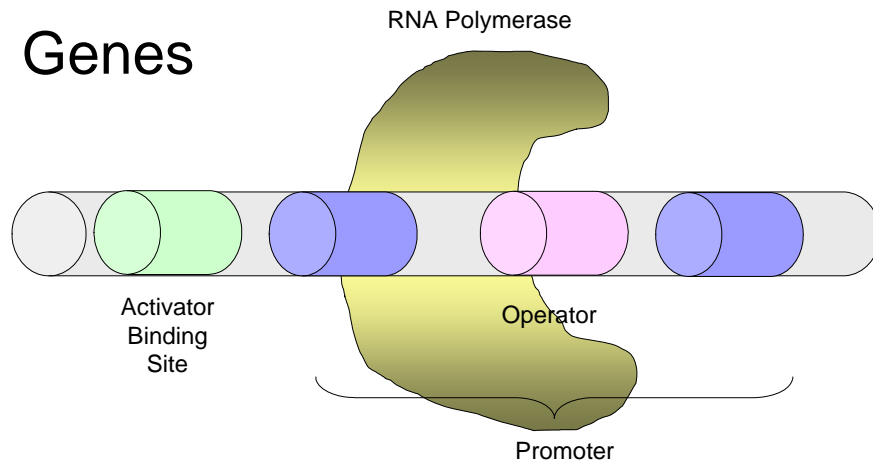
gene activator site which is where activator proteins can bind to start or enhance the expression of the gene. The operator sits and the overall promoter sequence are shown down from the activator site.⁷



Genes express themselves with the assistance of RNA polymerase. The RNA polymerase is key in that it binds to the DNA and then opens it up to allow for the transcription creating the mRNA required for the translation process. In the figure below we show this process.

⁷ This is detailed in Watson et al. Also see Griffiths et al.

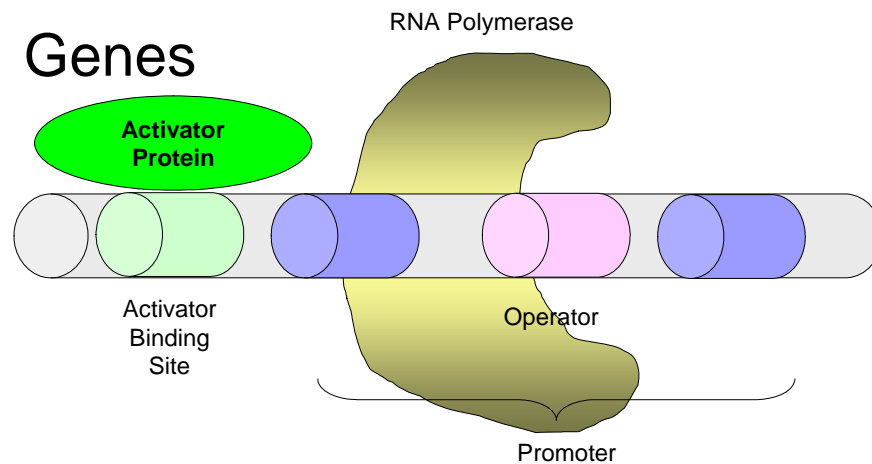
Genes



We will now focus on two actions which control the gene expression; activators and suppressors.

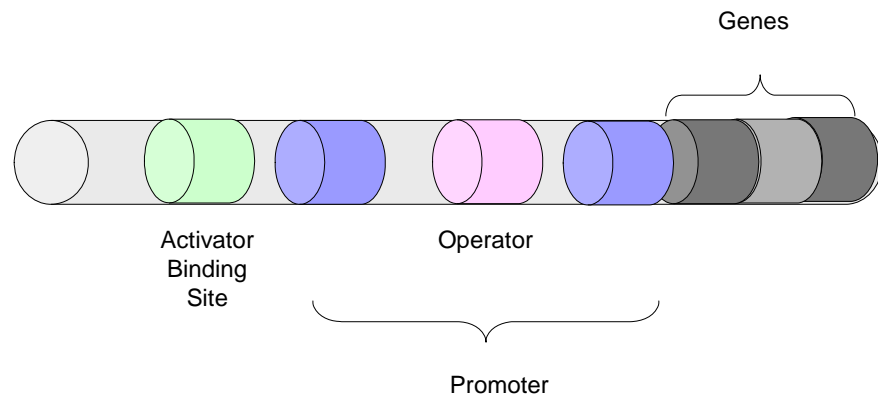
3.4 Activators

Activators are proteins which when attached to the gene assist in the expression of the gene. An activator is a protein resulting from another gene which can assist and facilitate the expression of a gene. Remember we want to look at the ensemble view, not the time view. Thus we assume that the RNA polymerase is continuously acting to produce proteins and that there is a continuous flow at some level of the activators. The cell process from the time view is shown below. An activator binds facilitates the RNA polymerase binding which in turn produces the mRNA and then in turn the proteins via the translation process.



If there is an activator then the gene can be readily expressed. The RNA polymerase then binds, creates the mRNA and this in turn produces the related protein. Activators stimulate this process. The Figure below depicts the location of the gene downstream from the activator and the promoter.

Genes



Now it is important to understand the activator from a time perspective and then from the ensemble perspective.

1. Activators are proteins generated by other genes in the cell.
2. Activators bind to the DNA and facilitate the production of the gene, which in turn produces another protein.
3. Activators can bind, release and then rebind. Each time they do that they produce another mRNA and that in turns produces another protein molecule.
4. From a time perspective, it is activator, produces gene reading, produces mRNA, and produces protein.
5. From an ensemble perspective we have a concentration of activator proteins and then we get a concentration of result proteins.

This then leads to a simple model:

$$P_o = \text{Output Protein Concentration}$$

$$P_i = \text{Input Protein Concentration}$$

$$P_o = A_{o,i} P_i$$

But there is also a dynamic model which we can state; to some degree this model is a hybrid of the time and ensemble approach. The model states:

$$\frac{dP_o}{dt} = f(P_o(t), P_i(t), t)$$

$$P_o(0) = P_o^0$$

$$P_i(0) = P_i^0$$

Now we must remember that this simple two protein, two gene model is just a simplification. In reality we may have dozens or hundreds of genes in this process. Now consider a simple linear model for this two gene system:

$$P_i(t) = P_i^0 \exp(-\lambda_i t)$$

$$\frac{dP_o(t)}{dt} = A_{o,i} P_i(t) + A_{o,o} P_o(t)$$

We can solve this differential equation. It is:

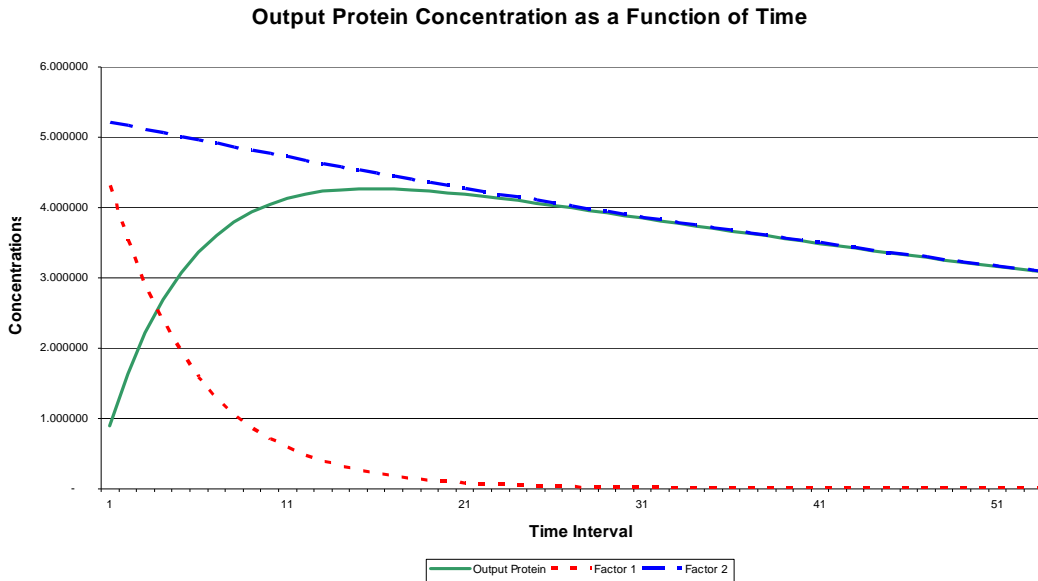
$$P_o(t) = k_{o,i} P_i(0) \left[\frac{\exp(-\lambda_i t) - \exp(-k_{o,o} t)}{\lambda_i - k_{o,o}} \right]$$

where;

$$A_{o,o} = -k_{o,o}$$

$$A_{o,i} = +k_{o,i}$$

We have solved this for a simple example using constants of 0.01 and 0.2 respectively.

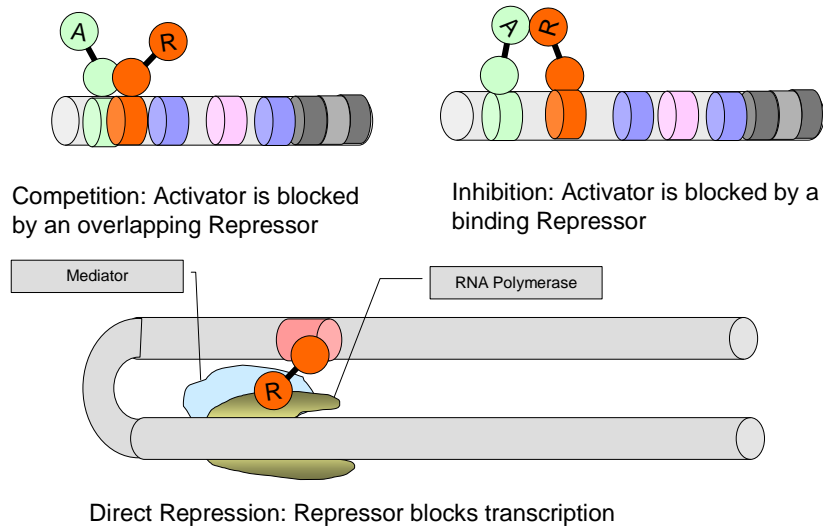


Note that the output protein concentration reaches a peak and then decays as per the driving protein. We will see this phenomenon again.

3.5 *Repressors*

In contrast to activators we also have genes which are suppressors. Three methods of suppressor action are shown below. A suppressor does the opposite of an activator. It suppresses the expression of a gene. The same logic will follow the repressor as was with activators. We again also want to view this from an ensemble perspective.

Repressors



As we did with the activator, we see a repressor stops the generation of the protein. This it is nothing more than a negative driver to protein generation.

3.6 Summary of Actions

We can now summarize what we have presented here:

1. Color is the result of anthocyanin production.
2. Anthocyanin production is a product of a specific pathway.
3. Pathways are mediated by enzymes, which are proteins generated by genes in the cell.
4. Proteins are generated by genes.
5. Gene activation is modulated by activator proteins and suppressor proteins.
6. Activator and suppressor proteins are generated by other genes.
7. One can model this overall process by a linked set of equations, both of a time varying nature and an ensemble, average steady state, nature.
8. An overall state model can be developed for the genetic control of color in plants.

We can now take this set of conclusions and use it to construct the state model.

4 EXPRESSION ANALYSIS AND IMPLICATIONS

In this section we develop a systems approach to the problem of color analysis and synthesis. This work is based upon the recent work of [Szallasi](#) and others. However this also builds upon the work in McGarty (1971) which focused a systems approach to the overall identification problem.

4.1 Approach: Engineering versus Science

The approach we take in this paper is an engineering approach rather than a biological approach.⁸ Our interest is in developing a model or sets of models which allow us by a verifiable means to show how the genes react and interact to produce the plant colors. We can compare this to the engineering approach to circuit design of transistor circuits versus the science of understanding the semiconductor from the point of view of detailed quantum mechanical models.

The biologist in our approach is akin to the physicists and engineers who approach the cell from the bottom up, trying to understand all of the intricate processes and steps that lead at the micro level to the developments we look at herein. In our approach it is akin to the engineer knowing that there is some function inside the semiconductors which may clearly be important but the engineer's interest is in designing and analyzing the transistor as a circuit element.

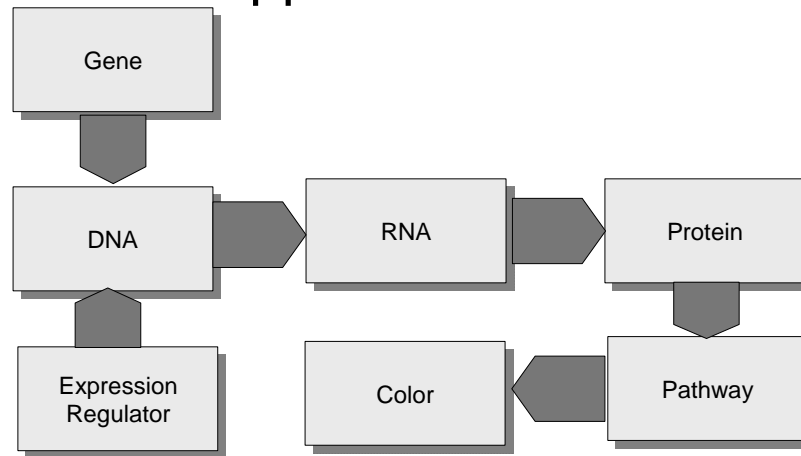
Thus for an engineer, if we increase a current here we get a decrease or an increase at some other point. The engineer creates a world view of a macro set of processes and models the details of the biologists in our case with a few set of equations which show the results of increases and decreases. This model must then be valid table and verifiable. One must be able to make measurements to show that the processes predicted indeed occur, to a reasonable degree of accuracy. Then one can analyze a genetic circuit and then in addition one can design a genetic circuit. We then can understand where the colors come from and possibly engineer the genes to develop and deliver on colors we desire.

4.2 A Control Paradigm

The basic control paradigm is contained in the following Figure. The expression regulator may be an activator or suppressor. It may be a result of a gene expression in the cell itself or quite possibly as we shall discuss fed through from another cell. There are many of these regulatory cycles and they are all interconnected. This basic paradigm is one of hundreds or thousands of such interconnected flows.

⁸ There has been a significant set of development recently in analyzing genetic data from a systems perspective. In this paper we have taken such an approach. The recent work by such authors as Perkins et al, Vohradsky, Hatzimanikatis et al, and the recent book by Szallasi are seminal. However, there is an issue here also or world view and what does one really want from the analysis. The bench scientist looks to understand all the details of the underlying processes. The engineer seeks to understand enough to model the process and to do so with a reasonable degree of accuracy but the ultimate goal for the engineer is control of the process and generation of new processes.

Current Approach

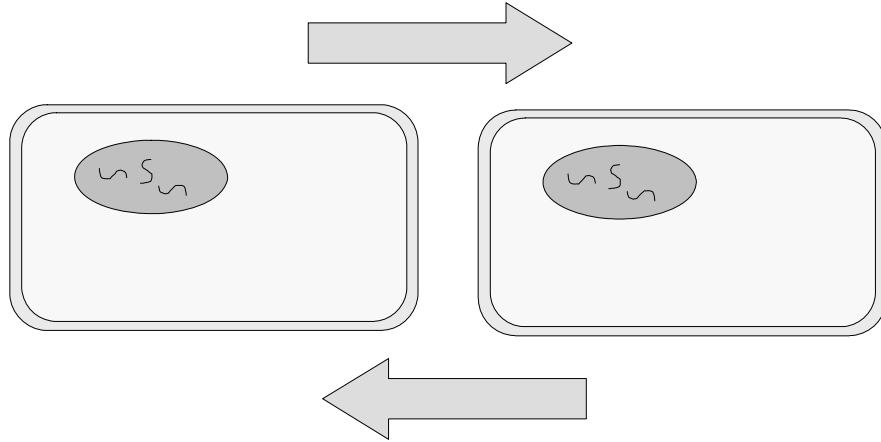


In developing our models we will use this construct. However, we can frequently focus on natural clusters of related genes. They may be a dozen or more such related genes in each cluster and possibly hundred of such clusters. Although cells and their proteins may affect all other cells, only a few of the genes regulated have a significant level of regulation. The low levels of “regulation” we shall consider just as noise.

4.3 Cell Signaling: Intra and Inter Cell

We must also better understand the inter cell signaling. Although we include it in this paper we have not as of yet produced a robust enough model for this set of processes. The Figure below presents the essence of the problem.

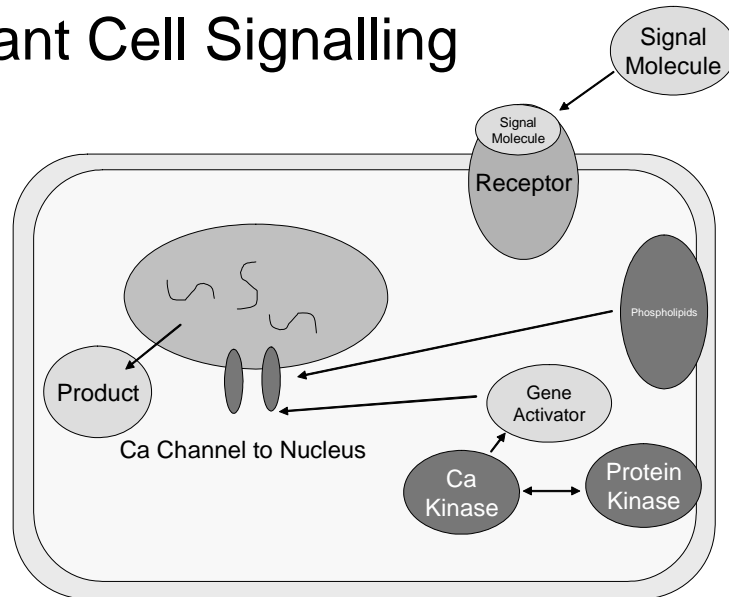
Plant Inter Cell Communications



What do the cells use to communicate and how. What are the elements?
Proteins?

Key to intercell signaling will be the receptor elements which control the flow of the controlling elements. This means that we must be able to introduce certain additional elements in the model which at this time are not yet fully developed. The Figure below highlights the issues of concern in this area.

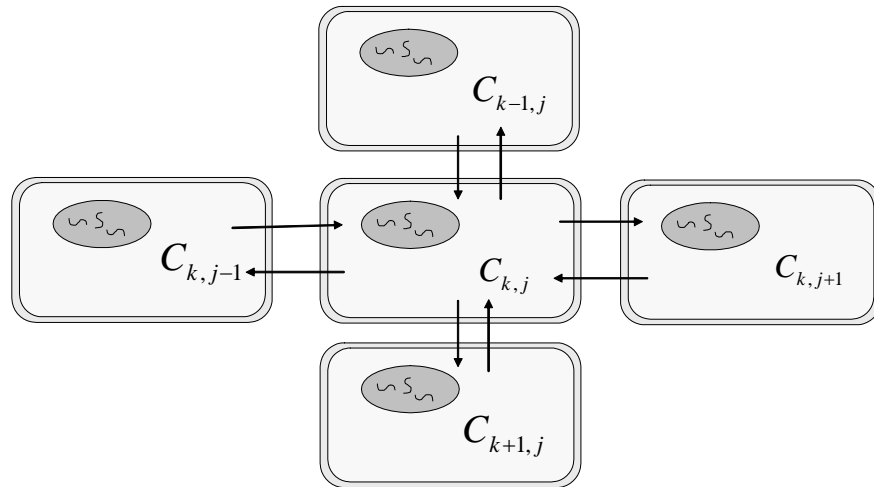
Plant Cell Signalling



Dey, Plant Biochemistry, p. 373, 1997 Academic

Then we must be able to establish a full network view of the signally processes. There has been considerable work looking at this from a meta perspective as some neural network. However the approach does not yet provide an adequate refection of a gene by gene analysis.

Plant Inter Cell Communications



What do the cells use to communicate and how. What are the elements?
Proteins?

5 FLOWER COLOR EXPRESSION

We have just shown that there are a wide variety of coloration in the daylily. In a little over a hundred years we have taken the dozen or so species and intermixed them and as a result have created a very complex set of flowers with characteristics which differ dramatically from the species.⁹ The species have managed to maintain their separate identities over thousands of years but in a small fraction of time we have been able to introduce multiple forms and colors. To understand this process we first have to understand where the colors come from. How do we get purple from a plant which is red, yellow, orange and possibly even brown? How are the colors made and how do we get from there to where we are today.

The first step in understanding that process is to understand the pathways that lead to color production in a single cell. Then we can address the issue of multiple cells and finally how the cells communicate. How do we get an eyezone for example. Why if a cell is whit do we go so abruptly to a purple eyezone. What is the mechanism for this process? We begin the exploration of this issue with a analysis of the underlying pathways.

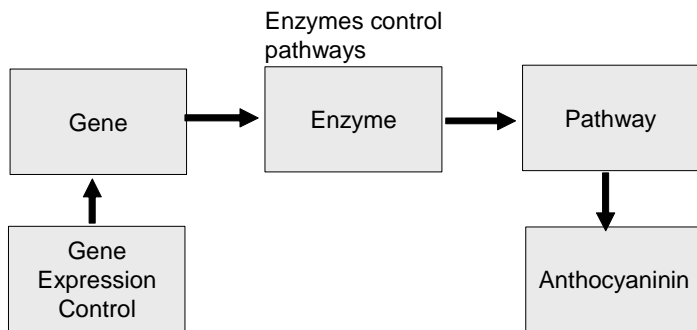
⁹ See Lensaw and Ghabrial for an excellent discussion of the tulip. In contrast to the daylily, the tulip craze of the seventeenth century was a dramatic bubble, and the irony was that most of the color variations were induced by viruses.

5.1 Pathways and Enzymes

Pathways are nothing more than a set of chemical reactions which get us from some primitive chemical to a more complex but useful chemical structure.¹⁰ In fact the pathways may be just a set of processes going from any one chemical structure to another independent of the nature of the starting and starting chemical. Some pathways are linear going from a beginning to an end and some are circular taking us from the beginning and back again; the Krebs cycle is an example. What makes the pathway work? Just three elements are required: (i) the underlying chemical constituents, (ii) some form of energy, (iii) generally some form of facilitation such a catalyst and in our analyses this is an enzyme.

The general flow structure we look at is shown below. In our view, not the only such view but one convenient for the development of our argument, we have the pathway but it facilitated by an enzyme, a protein. The protein is generated by a gene. And the gene is activated by some other element, generally another protein. In our case shown below the output is some anthocyanin. The more of the enzyme, namely the more the gene expresses itself the more anthocyanin we get. Thus if we can get the gene to express then we get more of that specific anthocyanin, more pelargonidin for example. We defer to the next section how we get this gene to express so strongly.

Pathways, Enzymes and Expression



Many factors control the expression of the gene. Even the cell which is next to the one producing the enzyme.

Each anthocyanin creates a color element. The more of that one type the richer that element. Combining them together creates a totally new color.

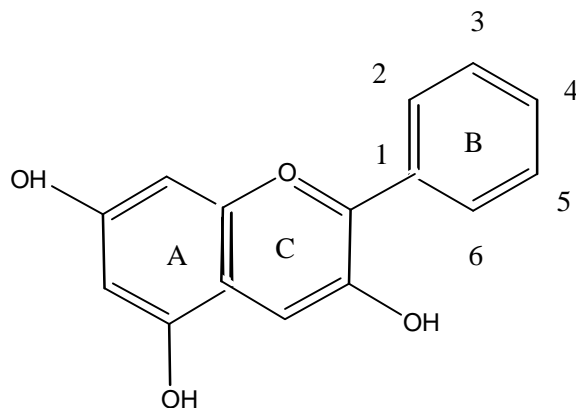
¹⁰ See Taiz for an excellent overview. Dey is also a superb and current reference. The older references by Goodwin are useful but they fail to account for the genetic effects.

The opposite is also true. Namely if we can suppress the gene then we can get less and even possibly no anthocyanin from the pathway. This is the first step in the development of an overall system model.

5.2 Anthocyanins

Let us consider our first pathway. This is the pathway which creates anthocyanins.¹¹ The anthocyanin molecules is shown below. Note on the B ring we have six sites to which we can attach differing molecular chains. This will be an important element when we see the different configurations and their implications.

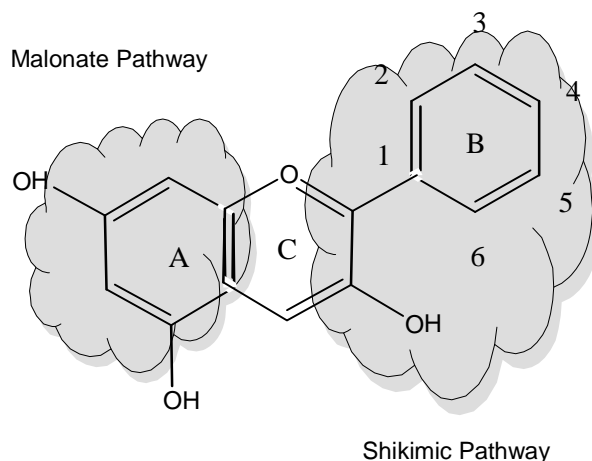
Anthocyanidin



The anthocyanin or anthocyanidin molecules comes from two different pathways. In the figure below we have taken the basic resulting molecule and have shown that there are two elements; one is from the shikimic pathway and the other from the malonate pathway. This means that we have to understand both pathways to understand the ultimate abundance of the product.

¹¹ See the papers by Mol and also by Winkel-Shirley. They are excellent in the characterization of the pathways. Also the papers by Holton and the one by Jaakola are quite useful here as well.

Anthocyanidin



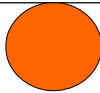
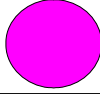



Before continuing we want to look at what the results would look like if we have different substituents on the B ring. In the Table below we show that the terminations on the 3, 4 or 5 elements yield different results. The results give pelargonadin, cyanidin, delphinidin, peonidin, and petunidin. Each obviously named after their related flower and each resulting an anthocyanin of a different color.

Colors

<i>Anthocyanidin</i>	<i>Substituents</i>	<i>Color</i>
Pelargonadin	4'-OH	orange-red
Cyanidin	3'-OH, 4'-OH	purplish red
Delphinidin	3'-OH, 4'-OH, 5'-OH	bluish purple
Peonidin	3'-OCH ₃ , 4'-OH	rosy red
Petunidin	3'-OCH ₃ , 4'-OH, 5'-OCH ₃	purple

In the Table below we have shown the colors of each of these as well as the weighting of a red, green and blue combination which best matches the color. Thus one can in an 8 bit color schemes, as one would find in any PC color scheme, get the resulting anthocyanin colors by blending the R, B, G elements to yield what we are seeking. This relating the colors back to RGB is critical since it get reflected in the ultimate flower color.

Colors (R, G, B)

Pelargonadin (255, 102, 0)	
Cyanidin (255, 0, 255)	
Delphinidin (153, 102, 255)	
Peonidin (255, 153, 204)	
Petunidin (153, 0, 153)	

Now if we assume we have only anthocyanins for color, and that we have the above combinations available, we ask how do we combine these colors in a weighted manner to obtain the desired color. This approach is critical to the overall understanding. First we show by a weighted RGB we get the color we seek or the color which is presented. Then we assume that if we can then do the same for each anthocyanin, then we can create any desired color from a weighted collection of anthocyanins. This means that we can then determine what the relative percents of expression of any anthocyanin is and this lets us then go back to how strongly the gene for that anthocyanin is expressed. The model we presented earlier will be a key element in this overall process.

5.3 *Other Color Elements*

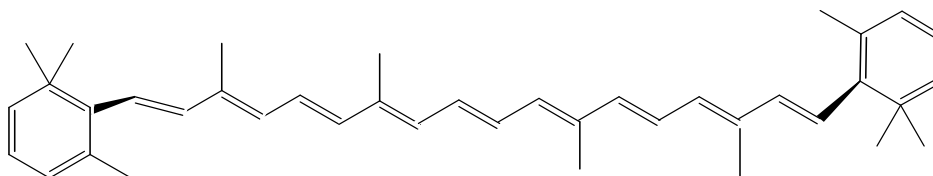
Anthocyanins are not the only elements which are secondary products which produce color. There are three classes of chemicals which give rise to color; anthocyanins, flavones or flavonols, and carotenoids. The Table below depicts the different elements and their colors. The approach we took above for the anthocyanins can be take for the flavones and carotenoids as well. It should be noted that there may not be a unique solution here but there are several possible but they can be narrowed down by actual determination of one to three elements as baseline.

<i>Class</i>	<i>Agent</i>	<i>Color</i> ¹²
Anthocyanidin		
	Pelargonidin	orange-red
	Cyanidin	purplish-red
	Delphinidin	bluish-purple
	Peonidin	rosy red
	Petunidin	purple
	Malvinidin	
Flavonol		
	Kaempferol	ivory cream
	Quercetin	cream
	Myricetin	cream
	Isorhamnetin	
	Larycitrin	
	Syringetin	
	Luteolin	yellowish
Agipenin	Cream	
Carotenoids		
	Carotene	orange
	Lycopene	Orange-red

We now summarize the other element classes.

5.4 Carotenoids

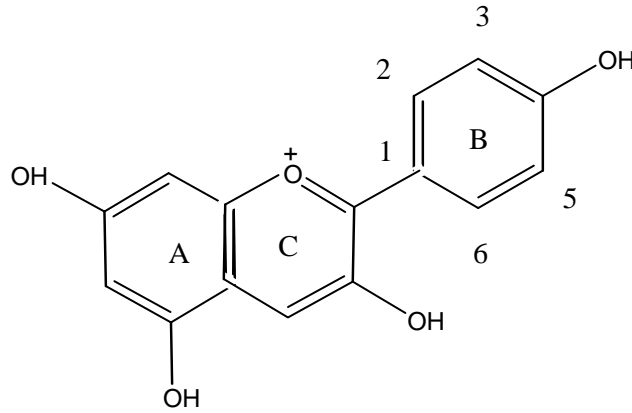
Carotenoids are what is quite common in the carrot, the orange hew we see in that root. Its molecular structure is shown below, this is beta carotene.



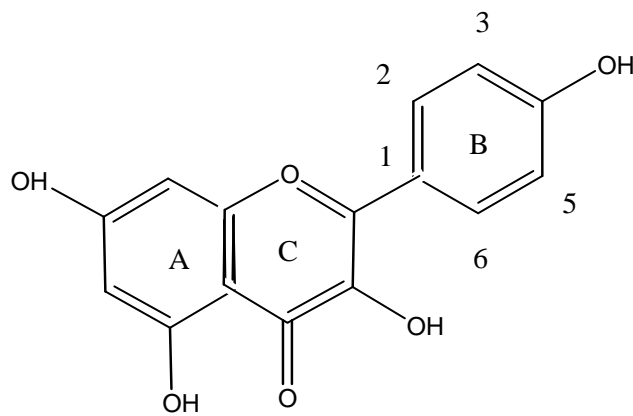
¹² See Taiz p. 334 for the anthocyanidin color and Bernhardt for the flavonol and carotene.

5.5 Flavones

The flavonols, or flavones are quite similar to anthocyanin. Their structure is shown below. Note that we have compared it to that of anthocyanin.



Anthocyanidin



Flavonol

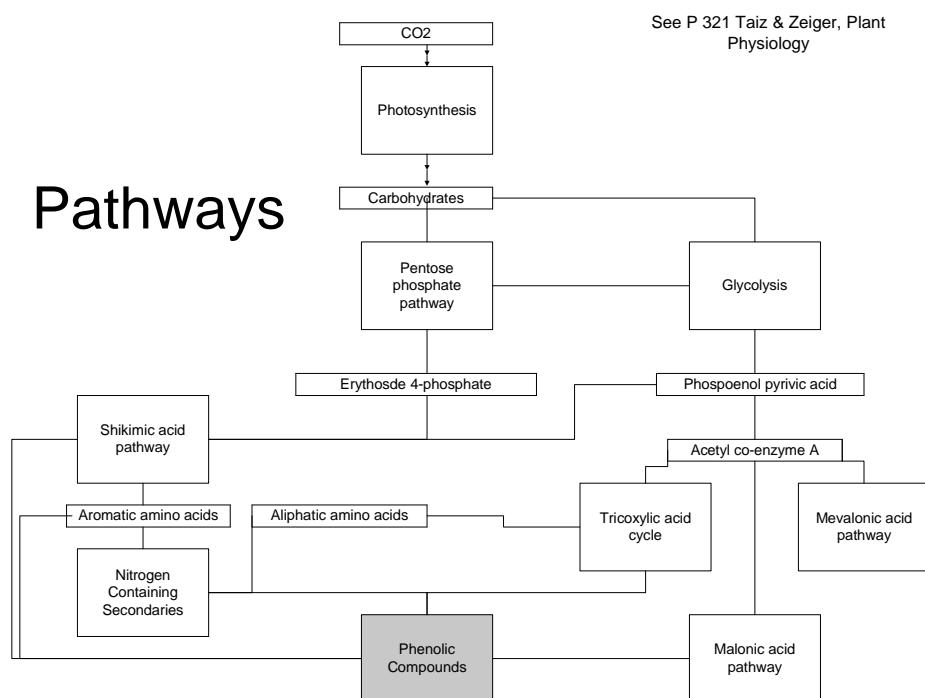
We can also show how closely they relate in substitutions and colors. This is shown in the Table below.

Flavonol	Anthocyanidin	Substitution	
		3'	5'
Kaempferol	Pelargonidin	H	H
Quercetin	Cyanidin	OH	H
Myricetin	Delphinidin	OH	OH
Isorhamnetin	Peonidin	OCH ₃	H
Larycitrin	Petunidin	OCH ₃	OH
Syringetin	Malvinidin	OCH ₃	OCH ₃

5.6 Pathways

In this section we present the pathways for the three classes we have described above. We first present an overview of the pathway and then we present the details of the pathway and the enzymes used in each step. The key observation is that we must have enzymes to go from step to step in the pathways and that if any one enzyme is missing we cannot proceed on that path, and further the path with the small amount of enzyme becomes the limiting path. Thus, we do not have a one to one map here. The production of any one anthocyanin, for example, is limited by the lowest produced enzyme, and the other enzymes may be present in abundance.

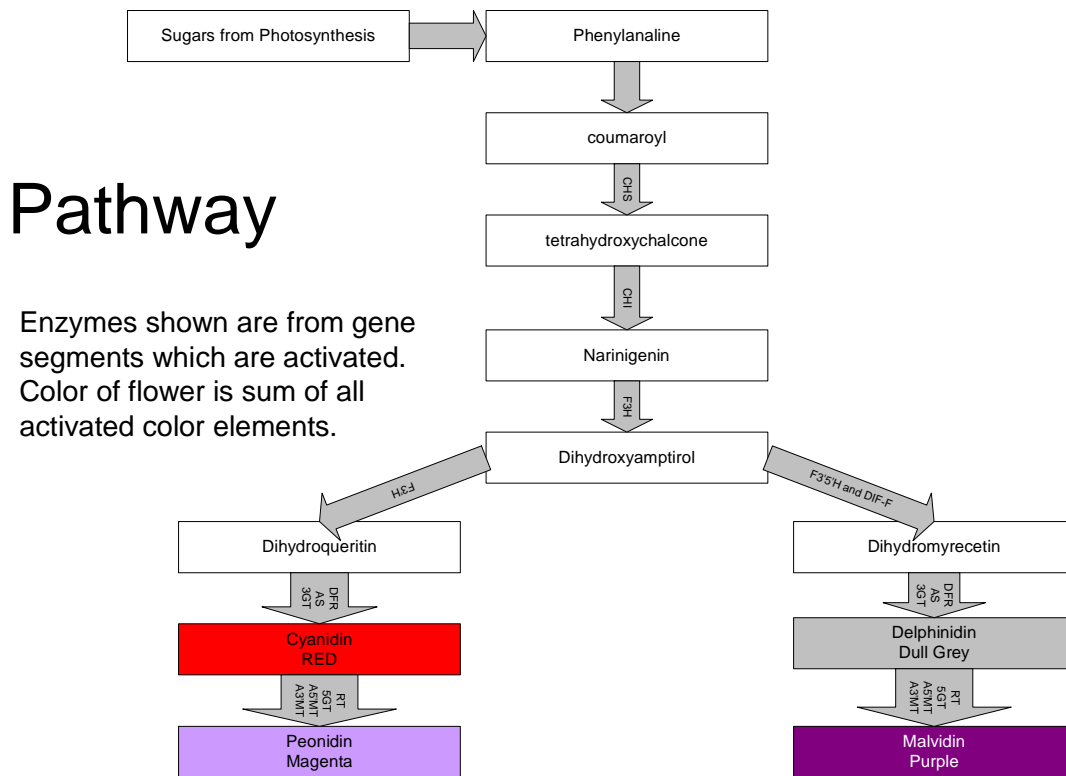
The following is the overall pathway for all elements.



The above shows how we start from CO₂ and then go through a variety of other pathways. We will review those pathways in some detail since it is the enzyme control in them which is key.

5.6.1 Anthocyanin Pathway

The anthocyanin pathway with the controlling enzymes is shown below. The enzymes are presented in the arrows linking each step in this pathway. This pathway shows the start as a sugar element and then goes to phenylalanine and then down through the chain to one of the four indicated anthocyanins.

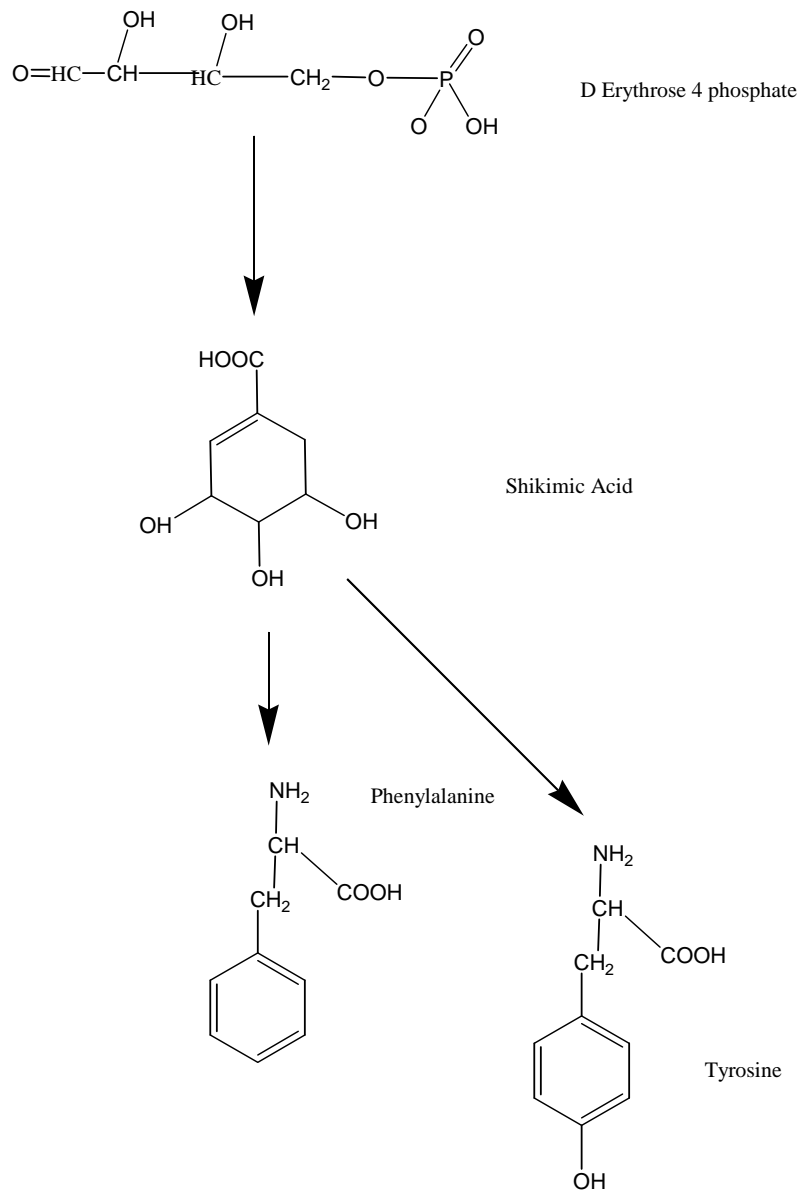


Note that at each step there is an enzyme element. The genetic loci for cloned flavonoid enzymes in Arabidopsis are shown in the following Table.¹³

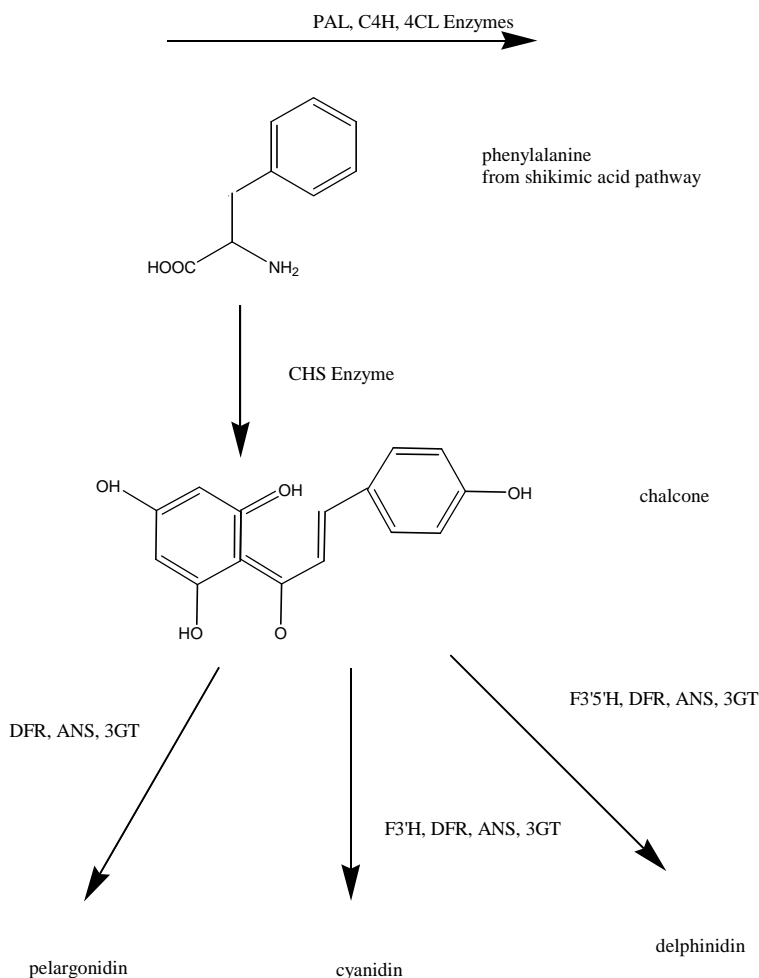
¹³ See Similar information for maize, petunia, and snapdragon is described by Holton and Cornish (1995). Based on the AGI map, 11/12/00; numbers in parentheses refer to P1 or bacterial artificial chromosome clones on which these sequences reside. Transposon- tagged mutant for FLS1 (Wisman et al., 1998).

Enzyme	Locus	Chromosome	Map Position
CHS	tt4	5	7,050 kb (MAC12)
CHI	tt5	3	21,000 kb (T15C9)
F3H	tt6	3	19,600 kb (F24M12)
F39H	tt7	5	4,400 kb (F13G24)
FLS	fls1<Enc	5	FLS1: 4,700 kb (MAH20) FLS2-5;: 32,150 kb (MBK5) FLS6: 24,350 kb (MRH10)
DFR	tt3	5	23,800 kb (MJB21)
LDOX	tt19	4	16,900 kb (F7H19)
LCR	ban,ast d	1	26,800 kb (T13M11)

The pathway for the conversion of the sugar erythrose to penylanaline is shown in the reaction below. This accounts for the upper part of the pathway which we have shown. It uses the Shikimic pathways which we have shown in the initial discussion on the pathways.



The conversion details from phenylalanine through chalcone to the anthocyanins is shown in the reaction below. We have reiterated by transition the enzymes which facilitate each step in this process.



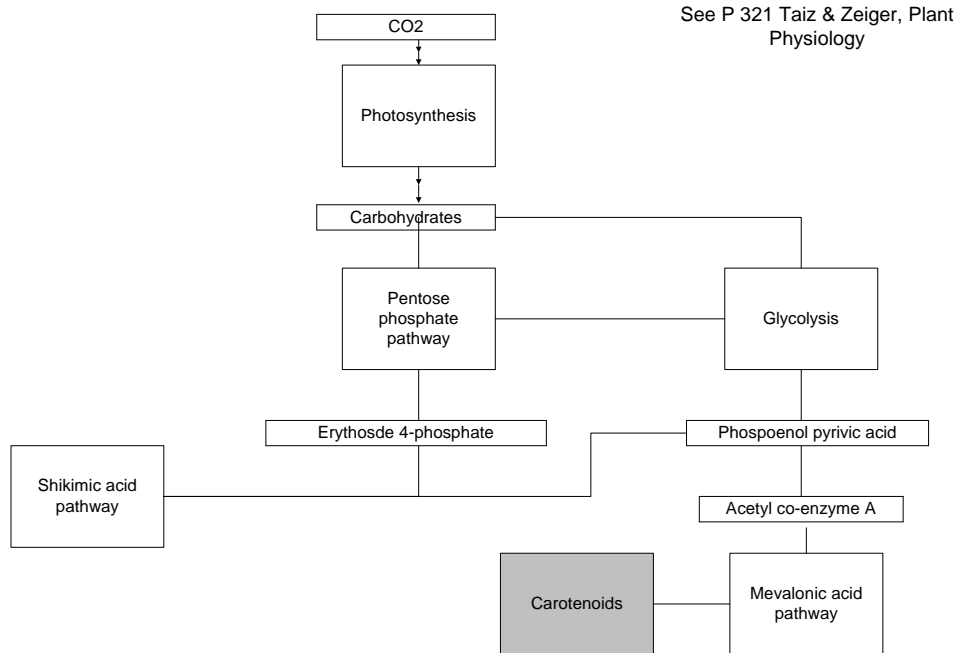
What these process point out can be summarized as follows:

1. There are common pathways which are operational in all plants for the generation of the pigments.
2. Enzymes used as activators modulate the amount of production of the enzymes.
3. The products of these pathways, the anthocyanins, are driven by the concentration of the facilitating enzymes.

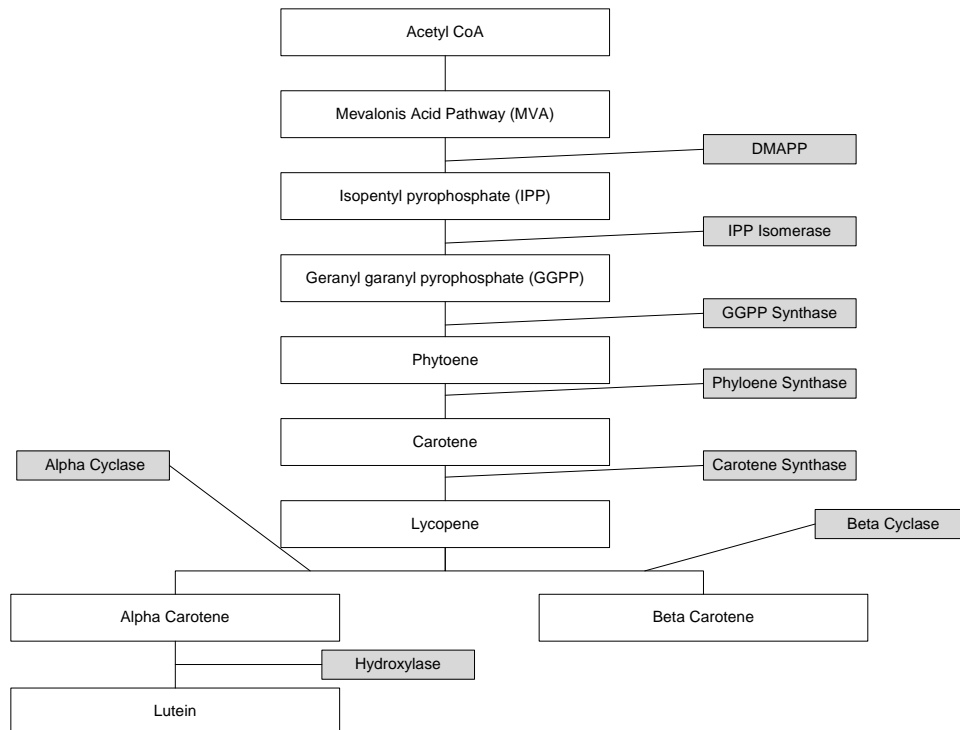
Secondary products always have this type of production process. As we look at a cell, from a system point of view we see facilitating proteins and secondary products. The concentration of the secondaries are proportional, in some general way, to the concentration of the facilitating proteins. However we see there are many facilitating proteins which may make this a more complex analysis, however doable.

5.6.2 Carotenoid Pathway

We have shown the carotenoids as above. The carotenoid pathway is shown below. We have demonstrated this in general terms earlier but in this case below we see the specific details.



We show below the pathways and the facilitating enzymes. In many ways it appears identical to the anthocyanin pathway and the facilitating enzymes.



5.6.3 Flavonol Pathway

The flavonol pathway is identical to that of the anthocyanin. See Winkel-Shirley.

6 CONCLUSIONS

The basic principles of genetics is essential to understanding the overall issues of species and their related colors, patterns forms and growth. The principles that we believe are critical are those related to activators and repressors and the interlinking of them in a dynamic system. As we have stated herein, we look at this as an engineering and design problem and not as a scientific assault on the unknown. As engineers we look and say; what can we achieve with what we know, and if that does not work then go back and question the assumptions. The alternative is to make what we don't know acts as "noise". That is the engineering approach.

The issues of the secondary pathways will become essential when we look at the issues of how genes control, the pathways and then how the pathways generate the color we see in flowers. We have expanded upon this extensively in two areas; color and its control and patterns and their control.

7 REFERENCES

1. Allard, R., **Plant Breeding**, Wiley (New York) 1960.
2. Atkins, P. **Physical Chemistry**, Freeman (New York) 1990.
3. Baici, A., *Enzyme Kinetics, The Velocity of Reactions*, Biochem Journal, 2006, pp. 1-3.
4. Bartel, B., S. Matsuda, *Seeing Red*, Science, Vol 299, 17 Jan 2003, pp 352-353.
5. Benson D. L., *Unraveling the Turing Bifurcation Using Spatially Varying Diffusion Coefficients*, Jour Math Bio, Vol 37 1998, pp. 381-417.
6. Benson, D. L., et al, *Diffusion Driven Instability in an Inhomogeneous Medium*, Bull Math Bio, Vol 55 1993, PP. 365-384.
7. Berns, R. S., **Principles of Color Technology**, Wiley (New York) 2000.
8. Born, M., E. Wolf, **Principles of Optics**, 4th Ed, Pergamon (New York) 1970.
9. Brown, A., et al, **Plant Population Genetics, Breeding, and Genetic Resources**, Sinauer (Sunderland, MA) 1990.
10. Campbell, A., L. Heyer, **Genomics, Proteomics, and Bioinformatics**, Benjamin Cummings (New York) 2003.
11. Carey, F. A., **Organic Chemistry**, McGraw Hill (New York) 1996.
12. Causton, H. et al, **Microarray Gene Expression and Analysis**, Blackwell (Malden, MA) 2003.
13. Cavalli-Sforza, L. L., W. F. Bodmer, **The Genetics of Human Populations**, Dover (Mineola, NY) 1999.
14. Chase, M., *Monocot Relationships, an Overview*, Journal of Botany, Vol 91 2004 pp 1645-1655.
15. Chen, T., et al, *Modeling Gene Expression with Differential Equations*, Pacific Symposium on Biocomputing, 1999 pp. 29-40.
16. Chung, M., J. Noguchi, *Geographic spatial correlation of morphological characters of *Hemerocallis middendorfii* complex*, Ann Bot Fennici Vol 35, 1998, pp. 183-189.
17. Chung, M., *Spatial Structure of three Populations of *Hemerocallis hakuunensis**, Bot. Bull. Acad. Sci., 2000, Vol 41, pp. 231-236.
18. Cilla, M. L., D. Jackson, *Plasmodesmata Form and Function*, Current Opinion in Cell Bio, Vol 16 2004 pp. 500-506
19. Crawford, D., **Plant Molecular Systematics**, Wiley (New York) 1990.
20. Cronquist, A., **The Evolution and Classification of Flowering Plants**, New York Botanical Garden Pres (Bronx, NY) 1986.
21. Dahlgren, R.M.T., **The Families of Monocotyledons**, Springer (New York) 1985.

22. Daly, D., et al, Plant Systematics in the Age of Genomics, *Plant Physiology*, Dec 2001, pp 1328-1333.
23. Dey, P. M., J. B. Harborne, **Plant Biochemistry**, Academic Press (New York) 1997.
24. Dunn, G., B. Everitt, **Mathematical Taxonomy**, Dover (Mineola, NY) 2004.
25. Durbin M. L. et al, *Genes That Determine Flower Color, Molecular Phylogenetics and Evolution*, 2003 pp. 507-518.
26. Durbin, R. et al, **Biological Sequence Analysis**, Cambridge (Cambridge) 1998.
27. Durrett, H., **Color**, Academic Press (New York) 1987.
28. Eisen, M., et al, *Cluster Analysis and Display of Genome Wide Expression Patterns*, *Proc Nat Acad Sci* 1998 Vol 98 pp 14863-14868.
29. Erhardt, W., **Hemerocallis**, Timber Press (Portland, OR) 1992.
30. Esau, K., **Anatomy of Seed Plants**, Wiley (New York) 1977.
31. Felsenstein, J., **Inferring Phylogenies**, Sinauer (Sunderland, MA) 2004.
32. Fox, M. A., J. K. Whitsell, **Organic Chemistry**, Jones and Bartlett (Boston) 1997.
33. Gitzendanner, M., Soltis, P, *Patterns of Genetic Variation in Rare and Widespread Plant Congeners*, *Journal of Botany*, Vol 87 2000 pp 783-393.
34. Goodwin, T.W., **Chemistry and Biochemistry of Plant Pigments**, Vols 1 and 2, Academic Press (New York) 1976.
35. Griffiths, A., et al, **Genetic Analysis 5th Ed**, Freeman (New York) 1993.
36. Guerro, C., et al, Analysis of the Expression of Two Thioprotease Genes from Daylily (*Hemerocallis* spp.) during Flower Senescence, *Plant Molecular Biology*, Vol 36 1998 pp 565-571.
37. Gusfield, D., **Algorithms on Strings, Trees, and Sequences**, Cambridge (New York) 1997.
38. Harborne, J. B., C. A. Williams, *Anthocyanins and Flavonoids*, *Nat Prod Rep* 2001 Vol 18 pp. 310-333.
<http://www.rsc.org/ej/NP/1998/a815631y.pdf>
39. Harborne, *Spectral Methods of Characterizing Anthocyanins*, *Biochemical Journal*, pp 22-28, 1958.
<http://www.biochemj.org/bj/070/0022/0700022.pdf>
40. Harvey, P., M. Pagel, *The Comparative Method in Evolutionary Biology*, Oxford (New York) 1991.
41. Hasegawa, M., T. Yahara, Bimodal Distribution of Flowering Time in a Natural Hybrid Population of Daylily (*Hemerocallis fulva*) and nightly (*Hemerocallis citrina*), *Journal of Plant Research*, 2006 pp 63-68.
42. Hatzimanikatis, V., *Dynamical Analysis of Gene Networks Requires Both mRNA and Protein Expression Information*, *Metabolic Engr*, Vol 1, 1999, pp. 275-281.
43. Haywood, V. et al, *Plasmodesmata: Pathways for Protein and Ribonucleoprotein Signaling*, *The Plant Cell*, 2002 PP 303-325.
44. Hildebrand, F. B., **Numerical Analysis**, 2nd Edition, Dover (New York) 1987.
45. Holton, T., E. Cornish, *Genetics and Biochemistry of Anthocyanin Biosynthesis*, *The Plant Cell*, Vol 7, 1995, pp 1071-1083.
46. Innan, H., et al, *A Method for Estimating Nucleotide Diversity from AFLP Data*, *Genetics*, Vol 151 March 1999, pp. 1157-1164.
47. Jaakola, L. et al, *Expression of Genes Involved in Anthocyanin Biosynthesis*, *Plant Physiology*, Vol 130 Oct 2002, pp 729-739.
48. Jenkins, F. A., H. E. White, **Fundamentals of Optics**, McGraw Hill (New York) 1957.
49. Judd, D. B. et al, **Color 2ND Edition**, Wiley (New York) 1963.
50. Judd, W., et al, **Plant Systematics**, 3rd Ed, Sinauer (Sunderland, MA) 2008.
51. Judd, W., R. Olmstead, A Survey of Tricolpate (Eudicot) Phylogenetic Relationships, *Journal of Botany* Oct 2004, pp 1627-1644.
52. Kadar, S., et al, *Modeling of Transient Turing Type Patterns in the Closed Chlorine Dioxide-Iodine-Malonic Acid-Starch Reaction System*, *J Phys Chem*, Vol 99 1995 pp. 4054-4058.
53. Kang, S., M. Chung, *Hemerocallis taeanensis* (Liliaceae) a New Species from Korea, *Systematic Botany* 1997 Vol 22 pp 427-431.
54. Kang, S., M. Chung, Genetic Variation and Population Structure in Korean Endemic Species, IV *Hemerocallis hakuunensis*, *Journal of Plant Research*, Vol 110 1997 pp 209-217.
55. Kang, S., M. Chung, High Levels of Allozyme Variations Within Populations and Low Allozyme Divergence within and Among Species, *American Journal of Botany*, Vol 87 2000 pp 1634-1646.
56. Kohane, I., et al, **Microarrays for an Integrative Genomics**, MIT Press (Cambridge) 2003.
57. Koopman, W., et al, *AFLP Markers as a Tool to Reconstruct Complex Relationships: A Case Study in Rosa*, *Journal of Botany*, Vol 95 2008 pp 353-366.
58. Lee, D., **Nature's Palette**, University of Chicago Press (Chicago) 2007.

59. Lesk, A., **Bioinformatics**, Oxford (New York) 2002.
60. Levi, L., **Applied Optics**, Wiley (New York) 1968.
61. Mauseth, J. D., **Plant Anatomy**, Benjamin (Menlo Park, CA) 1988.
62. Mayo, O., **The Theory of Plant Breeding**, Oxford (New York) 1987.
63. Mayr, E., **Evolution and the Diversity of Life**, Harvard University Press (Cambridge) 1976.
64. Mayr, E., **Population, Species and Evolution**, Harvard University Press (Cambridge) 1970.
65. Mayr, E., **The Growth of Biological Thought**, Harvard University Press (Cambridge) 1982.
66. Mayr, E., **Toward a New Philosophy of Biology**, Harvard University Press (Cambridge) 1988.
67. McGarty, T. P., *Flower Color and Means to Determine Causal Anthocyanins And Their Concentrations*, MIT 2008,
<http://www.telmarcgardens.com/Documents%20Papers/Flower%20Color%20and%20Means%20to%20Determine%2002.pdf>
68. McGarty, T. P., *On the Structure of Random Fields Generated by a Multiple Scatter Medium*, PhD Thesis, MIT 1971. <http://mit.edu/mcgarty/www/MIT/Paper%20Hypertext/1971%20PhD%20MIT.pdf>
69. McGarty, T., *Gene Expression in Plants: Use of System Identification for Control of Color*, MIT, 2007.
<http://mit.edu/mcgarty/www/MIT/Paper%20Hypertext/2007%20Gene%20Expression%20IEEE%2007%2002.pdf>
70. McGarty, T., **Stochastic Systems and State Estimation**, Wiley (New York) 1974.
71. McMurry, J., Begley, T., **The Organic Chemistry of Biological Pathways**, Roberts & Company Publishers, 2005.
72. Milgrom, L. R., **The Colours of Life**, Oxford (New York) 1997.
73. Mohr, H., P. Schopfer, **Plant Physiology**, Springer (New York) 1995.
74. Mol, J, et al, *How Genes Paint Flowers and Seeds*, *Trends in Plant Science*, Vol 3 June 1998, pp 212-217.
75. Mol, J., et al, *Novel Colored Plants*, *Current Opinion in Biotechnology*, Vol 10, 1999, pp 198-201.
76. Mueller, U., L. Wolfenbarger, AFLP Genotyping and Fingerprinting, *Trends Ecol. Evol.* **14: 1999**, pp 389–394.
77. Munson, R., **Hemerocalis, The Daylily**, Timber Press (Portland, OR) 1989.
78. Murray, J., **Mathematical Biology**, Springer (New York) 1989.
79. Murrell, J., *Understanding Rate of Chemical Reactions*, University of Sussex.
80. Naik, P. S., et al, *Genetic manipulation of carotenoid pathway in higher plants*, *Current Science*, Vol 85, No 10, Nov 2003, pp 1423-1430.
81. Nei, M., S. Kumar, **Molecular Evolution and Phylogenetics**, Oxford (New York) 2000.
82. Nicklas, K. J., *The Bio Logic and Machinery of Plant Morphogenesis*, *Am Jour Bot* 90(4) 2003 pp. 515-525.
83. Noguchi, J., H. De-yuan, *Multiple origins of the Japanese nocturnal Hemerocalis citrina*, *Int Jrl Plant Science*, 2004, Vol 16, pp. 219-230.
84. Norton, J. 1972. *Hemerocallis* Journal 26 (3) in Bisset, K. 1976. *Spectrophotometry, Chromatography and Genetics of Hemerocallis Pigments*. Dissertation, Florida State Univ.
85. Norton, J., *Some Basic Hemerocallis Genetics*, American Hemerocallis Society, 1982.
86. Oparka, K. J., A. G. Roberts. Plasmodesmata, *A Not So Open and Shut Case*, *Plant Phys*, Jan 2001, Vol 125 pp. 123-126.
87. Percus, J., **Mathematics of Genome Analysis**, Cambridge (New York) 2002.
88. Perkins, T., et al, *Inferring Models of Gene Expression Dynamics*, *Journal of Theoretical Biology*, Vol 230, 2004, pp. 289-299.
89. Petit, T. *The Patterned Daylily*, *The Daylily Journal*, Vol 62 No 2 2007 pp. 125-141.
90. Planet, P. et al, *Systematic Analysis of DNA Microarray Data: Ordering and Interpreting Patterns of Gene Expression*, Cold Spring Harbor, Genome Research, 2001 pp 1149-1155.
91. Rossi, B., **Optics**, Addison Wesley (Reading, MA) 1957.
92. Saitou, N., M. Nei, *The Neighbor Joining Method: A New Method for Reconstructing Phylogenetic Trees*, *Molecular Biological Evolution*, Vol 4 1987 pp 406-425.
93. Sansone, G et al, **Orthogonal Functions**, Academic Press (Mew York) 1958.
94. Sattath, S., A. Tversky, *Additive Similarity Trees*, *Psychometrica*, Vol 42, 1977 pp. 319-345.
95. Schnell, S, T. Turner, *Reaction Kinetics in Intracellular Environments with Macromolecular Crowding*, *Biophys and Molec Bio* vol 85 2004 pp. 235-260.
96. Sears, F. W., **Optics**, Addison Wesley (Reading, MA) 1949.
97. Sokal, R., P. Sneath, **Principle of Numerical Taxonomy**, Freeman (San Francisco) 1963.
98. Soltis, D. et al, *Evolution of Genome Size in Angiosperms*, *Journal of Botany*, Nov 2003, pp 1596-1603

99. Somasundaram, S., M. Kalaiselvam, *Molecular Tools for Assessing Genetic Diversity*, UN University Course, http://ocw.unu.edu/international-network-on-water-environment-and-health/unu-inweh-course-1-mangroves/Molecular_Tools_for_Assessing_Genetic_Diversity.pdf .
100. Stace, C., **Plant Taxonomy and Biosystematics**, Arnold (London) 1989.
101. Stout, A.B., **Daylilies**, Saga Press (Millwood, NY) 1986.
102. Stout, A., The Inflorescence in *Hemerocallis*, *Bulletin of the Torrey Club*, Vol 68 1941 pp 305-316.
103. Stout, A., C. Chandler, Pollen Tube Behavior in *Hemerocallis* with Special Reference to Incompatibilities, of the *Torrey Club*, Vol 60 1933 pp 397-417.
104. Strickberger, M., **Genetics**, 2nd Ed McMillan (New York) 1976.
105. Strong, J., **Concepts of Classical Optics**, Freeman (San Francisco) 1958.
106. Studier, J., K. Kappler, *A Note on the Neighbor Joining Algorithm*, *Molecular Biological Evolution*, Vol 5 1988 pp 729-731.
107. Stuessy, T., **Case Studies in Plant Taxonomy**, Columbia University Press (New York) 1994.
108. Szallasi, Z. **System Modeling in Cellular Biology: From Concepts to Nuts and Bolts**. MIT Press (Cambridge) 2006.
109. Taiz, L., E. Zeiger, **Plant Physiology**, Benjamin Cummings (Redwood City, CA) 1991.
110. Taubes, C. H., *Modeling Differential Equations in Biology*, Cambridge (New York) 2001.
111. Tetter, A., et al, *Primer on Medical Genomics Part III*, *Mayo Clinic Proc*, Vol 77 2002 pp 927-940.
112. Tinoco, I. et al, **Physical Chemistry**, Prentice Hall (Englewood Cliffs, NJ) 1995.
113. Tobias, A., *Directed Evolution of Biosynthetic Pathways to Carotenoids with Unnatural Carbon Bonds*, PhD Thesis, Cal Tech, 2006.
114. Tomkins, J. R., *DNA Fingerprinting in Daylilies*, Parts I and II, *Daylily Journal*, Vol 56 No 2 and 3 2001, pp. 195-200 and pp. 343-347.
115. Tomkins, J. R., *How much DNA is in a Daylily*, *Daylily Journal*, Vol 58 No 2 2003, pp. 205-209.
116. Tomkins, J., et al, *Evaluation of genetic variation in the daylily (*Hemerocallis*) using AFLP markers*, *Theor Appl Genet* Vol 102, 2001, pp. 489-496.
117. Turing, A., *The Chemical Basis of Morphogenesis*, *Phil Trans Royal Soc London B337* pp 37-72, 1945.
118. Van Trees, H. L., **Detection, Estimation and Modulation Theory**, Wiley (New York) 1968.
119. Vohradsky, J., *Neural Network Model of Gene Expression*, *FASEB Journal*, Vol 15, March 2001, pp. 846-854.
120. Vos, P., et al, *AFLP: A New Technique for DNA Fingerprinting*, *Nuclear Acids Research*, Vol 23 1995 pp 4407-4414.
121. Wade, L. G., **Organic Chemistry**, Prentice Hall (Saddle River, NJ) 2003.
122. Watson, J. et al, **Recombinant DNA**, 3rd Ed, Freeman (New York) 2007.
123. Watson, J., et al, **Molecular Biology of the Gene**, Benjamin Cummings (San Francisco) 2004.
124. Weir, B., **Genetic Data Analysis**, Sinauer (Sunderland, MA) 1990.
125. Wen, X., et al, *Large Scale Temporal Gene Expression Mapping of Central Nervous System Development*, *Proc Nat Acad Sci Neurology* 1998 Vol 95 pp 334-339.
126. Winkel-Shirley, B., *Flavonoid Biosynthesis*, *Plant Physiology*, Vol 126 June 2001 pp 485-493. <http://www.plantphysiol.org/cgi/reprint/126/2/485>
127. Yasumoto, A., T., Yahara, *Reproductive Isolation on Interspecific Backcross of F1 Pollen to Parental Species *Hemerocallis fulva* and *Hemerocallis citrina**, *Journal of Plant Research*, Vol 2008, pp .
128. Yu, O. et al, *Flavonoid Compounds in Flowers: Genetics and Biochemistry*, General Science Books, <http://www.danforthcenter.org/yu/pdf/e-flower-2006.pdf>
129. Zambryski, P., *Cell to Cell Transport of Proteins and Fluorescent Tracers via Plasmodesmata during Plant Development*, *Jour Cell Bio* Vol 161 No 2 Jan 2004, pp. 165-168.