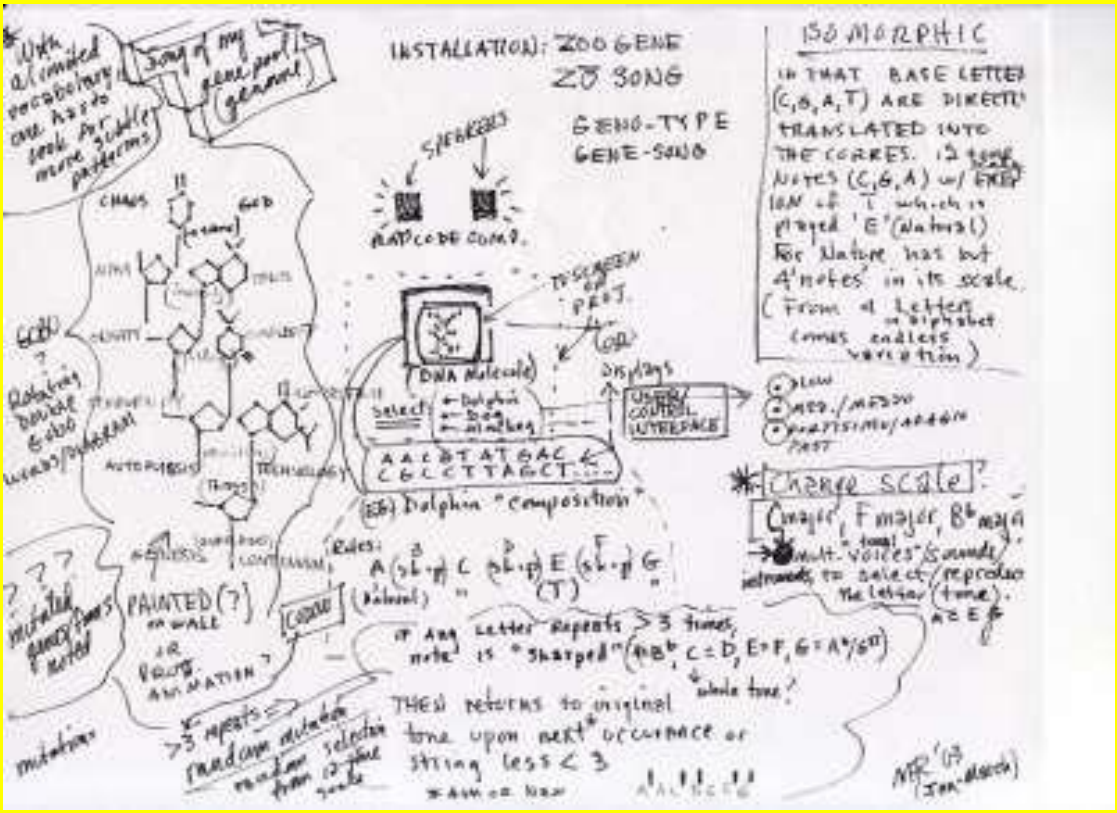
[This document contains concept designs, promotional and lecture material, and DNA>Music coding-conversion scheme information for the GENODEON project]

GENODEON – DNA Music Translation Code

The original DNA trans-coding method/concept design by Michael A. Ricciardi; conceived circa 1995, conceptual drawing/formalization, 2003, and demo'd/performed in 2010. [see page 8 for other Genodeon designs]



Details of the Coding Scheme:

There are 4 Nucleotide "letters" comprising DNA: A (adenine), T (thymine), C (cytosine), G (guanine)

Each letter is converted into a "letter-note" as follows:

A = A (natural) **T = E** (natural) **C = C** (natural) **G = G** (natural)

Each letter is played (initially) as a single note, or key, on the piano.

This conversion scheme is intuitive and naturally harmonious as each pairing is based upon musical 3rd s and 5th s.

Expanding the Conversion Scheme:

If a letter-note repeats itself (e.g., AAA...) three or more times, then on the *last* repeat, the letter-note is "sharped" (indicating a "reading error potential"), or raised half a step. Such *monomer runs* (or "tandem repeats") occur frequently in all genomes. Often such repeats/strings (usually occurring in 'triplets') can be quite long and are often the site of 'reading errors' (see next paragraph) by the polymerase molecule. Thus, in this conversion scheme, we both seek to *indicate* this repeat through 'sharping' the third occurrence of the note, and to *imitate* this 'mutation potential', which adds musical/harmonic variation to the genomic composition (note: this is a performance option; if one were solely converting DNA sequences to music, then one would follow the 'stricter' conversion scheme). [see: IMPORTANT NOTE, page __]

As an additional performance option, if that letter-note repeats a 4th, 5th, or 6th time, etc., then each new repeat of the letter-note may be substituted (with another letter-note) or duplicated (played in double time), or deleted altogether, representing a "reading error" as it might occur in Nature.

> **NOTE 1**: it is the RNA polymerase II molecule "reads" or transcribes one strand of DNA {ssDNA} into "messenger" RNA (*This function is performed by the piano player*). Reading errors by this molecule typically occur when it encounters long, repeating strings of the same letter ("tandem repeats"). An error can be a substitution, duplication, or deletion, and these can be passed on in the replicating cell and cause mutations. In Nature, such single letter substitutions (or deletions or repetitions) are known as Single Nucleotide Polymorphisms (SNPs) and can have profound impact on the expression of a given gene sequence (e.g., disease, or altered trait). We represent this in our code, as described above, and thus also further expand the musical potential of the demonstration).

Performing with the GENODEON DNA Conversion Scheme:

The general approach: Two, *single strand* DNA sequences (from two different animals) are to be played simultaneously (creating a *chimera*, or "mutant" double strand DNA composition*) in single note form. Single letter-notes can be converted into appropriate chords by playing each letter-note (A, T/E, C, G) as a major chord (with the specific letter-note as the "root" of its respective chord). In this way, broader harmonics of dissonances can be expressed and heard (and thus, in theory, the relatedness or non-relatedness of the genomes can be perceived, which is the functional purpose of this conversion scheme).

> * In "normal" (non-recombinant) DNA (which is composed of 2 spiral-bound nucleotide strands or sequences of billions of letters), each letter on each strand is bonded to its compliment on the opposing strand, like so:

NUCLEOTIDES: Cytosine [pairs with] Guanine ; Adenine [pairs with] Thymine

(C natural =) C + G = [perfect 5th]
(G natural =) G + C
(E natural =) T + A
(A natural =) A + T/E = [perfect 5th]

NOTE 2: The pianist may *substitute*, *duplicate* or *delete* (or 'sharp'/'flat') any note (letter) that follows a tandem repeat that is 3 or more letters long (e.g., ttt…, aaa…, ccc…, ggg…), as this is typically where *DNA copy errors* occur; the change in musical tone/pitch representing a "mutation".

Recall that the letter C always bonds with G (and G with C), and A always bonds with T (played as E natural). This natural pairing creates, in our music coding system, a natural harmony (E is the 'perfect 5th' of A, G is the 'perfect 5th' of C). But, in this methodology—due to the fact that we are pairing up ("mashing up"), or "recombining" unlike genomes (combining say, a hummingbird ssDNA sequence with an elephant ssDNA sequence)—this natural bonding rule (C only with G, A only with T) will very likely be broken routinely, with non-complimentary (and thus non-harmonic, or minor harmonic) pairings emerging, like so:

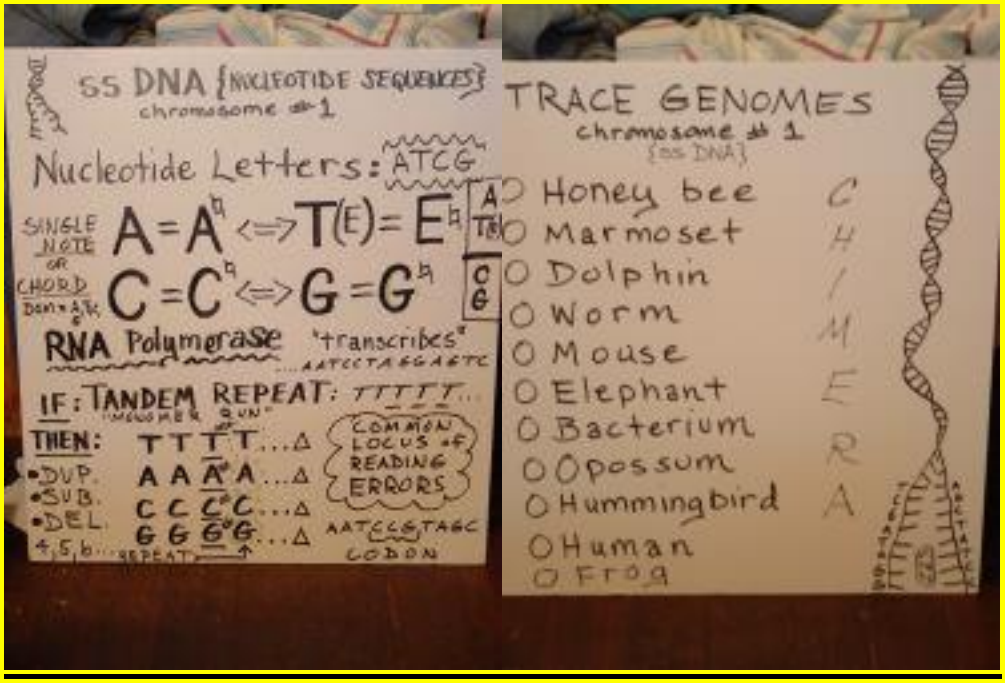
C + A
G + T (E)
A + G
(E) T + C

This possibility adds enormously to the auditory experience of the DNA music through creating new/'random' harmonic (or dissonant) pairings. **It also permits comparison/contrasting of related and unrelated animal genomes, which has scientific utility.**

IMPORTANT NOTE: Given a complete, *double-stranded* DNA 'read' of a given animal's genome (the sequence being read by a sequencing machine, such as an *Illumina* sequencer) one would acquire both the 'sense' and 'anti-sense' sequences of the DNA double helix which would follow the natural/normal base-pairing rule (A with T, C with G, etc.), and, if one were to play this dsDNA read (one with the Left hand, one with the Right) on a piano or keyboard, one should be able to immediately pick out an 'mismatch' between base pairs (e.g., C with T, or A with G) which can happen in reality. These mismatches would indicate (likely) mutation loci.

~~~~~~~~~~~~~~~~

VISUAL AIDS FOR PERFORMANCE (January 14, 2010):
~~~~~~~~~~~~~~~~



A Chimerical Composition with DNA Music / Paired DNA sequences (exercise # 1)

This first pairing (human and marmoset) is used in the first demonstration. With the marmoset being the "secret", or unknown, genome being paired with the (known) human one. After playing (first, one at a time, in single note form, then both together, in single note form, then both together in chord form*), the audience is queried to see if they can guess which genome (from a provided list) is being combined with the human. Humans and marmosets are both primates and are thus related, phylo-genetically (on the Tree of Life). This pairing is quite intriguing and harmonious and is used to "kick off" the further explorations (more fanciful pairings or chimeras) of this DNA Music experiment.

* The pianist is playing or "translating" one genome with his left hand, and the other with his right hand; "one hand, one strand".

Here is a sample chimerical composition (both animals are primates), pre-selected to demonstrate harmonic/genomic) similarity (in theory, there should be more 'matches' in letters between the two):

Homo Sapien (Human)
gatcacaggt ctatcaccct attaaccact cacgggagct ctccatgcat ttggtatttt

Marmoset
gaattccagg tgctagatgg atgtacagga atggaaacta atagggtctt ttcttcttca

{continued - 2}
Homo Sapien
cgtctggggg gtgtgcacgc gatagcattg cgagacgctg gagccggagc accctatgtc

Marmoset
tttcacccat tattgctgaa taatgtaatt agcaagcagt ttctattttc cttgaagcac

Homo Sapien {continued – 3}
gcagtatctg tctttgattc ctgcctcatc ctattattta tcgcacctac gttcaatatt

Marmoset
ataccttttgt tgtggtcggc cagattaaaa gacttttttag attatcagct gtggagaaca

Homo Sapien {continued – 4}
acaggcgaac atacttacta aagtgtgtta attaattaat gcttgtagga cataataata

Marmoset
gttactccct caaaatcttt ctttctgcca aaacccgaca agatttaaaa actatctatc

Promotional Flyer

Note: due to a technical glitch, an audio recording of the 2010 performance could not be retrieved.

GENODEON – DNA Music – Single Strand DNA Sequences

For simplicity of execution, it is advised that the following animal sequences (1st chromosome, ssDNA) be shortened to no more than two lines of code (for a total of 120

letter-notes). However, one may play any number of letter-notes one wishes, or as many as are provided, in this case the first 360 letters of each genome (Note: DNA "bar-coding" of animal species utilizes just 649 mitochondrial nucleotide letters to identify/distinguish one animal from another; the smaller nucleotide sample sizes used in the GENODEON are meant to be *representative* of each animal), and the more notes that are heard, the more "information" one can perceive about the relatedness or non-relatedness of the paired animals or organisms. But the point is to have fun and explore the auditory potential of these musical chimeras.

~~~~~~~~~~~~~~~~~~~

The following ssDNA sequences (all form the 1st chromosome; 11 animals in all) are taken from the **NCBI Sequence Read Archive at: http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?cmd=table&f=sample&m=data&s=sample**

Honey Bee (*Apis mellifera*)

GAAAACACAA CAATTAAATA TATAAAACAA GAAATTAAAA ATTATATCTC GTTTGGAATA

AACTAAAAAG CATTAAATCT ACTAATTAAG TATTTTCAAA GAATAAAAAA AGAGTCATAG

ATATAAAATA GAATCATAAG ATATAAAATA TATAAATATC ATTGAATATT TATTGAACTT

CTGATGTGTT ACACAATTGA TTAGAAGTTT AAACATACTT CGTTTTATTG ATTTTGGATT

Asiatic elephant (*Elephas maximus*) "Chendra"

gggggagatg ttgtcgagca cccatttcca agttgcaact ctccacaccg gccgctaccc

gtgggggact cgatcgcttc cgctaggctg aattgcgcga tgcgctggtc ctggttgggg

gtggccaggc agcgcgtgag aactgctggg gtgcgtggcc ccggtgtgcg cgcacccaga

gtctgggtgc agcttcgctc cgccaggttc tcctcccact cactctctcc catctcctcc

House Mouse (*Mus musculus*)

gaattcaatg tcctggtgat gggcttttgc ttgcaggtag ttctgtaaaa ggcagctatg

ctcacctctg acccatacac agtgcaacaa gcagggcatg gtttggggca gtagagagag

gagacttgaa aagattattg ttgggctttg agggggggttg acaagtccac atttcaatgg

aagccctgtt gctcctgtga aatacttgga aggcattcaa cacagttgac gaacatgaat
~~~~~~~~~~~~~~~~~~~

Frog – western clawed frog (*Xenopus Silurana tropicalis*)

gaaagacaga aaaaggggcg aagtgagcta cgctaagaga tgagcaacaa cagtaataag

agagctccaa caactgccac tcagaggtta aaacaggatt acctaaggat caagaaagac

ccagtgcctt acatttgtgc agaacccctt ccttcaaata ttcttgaatg gcactatgtg

gtgcgaggac ctgagatgac accatatgaa ggtggttatt accatggcaa acttgttttt

Bacillus Bacterium (*Escherichia coli*)

gttaactgtg gtggttgtca ccgcccatta cacggcatac agctatatcg agccttttgt

acaaaacatt gcgggattca gcgccaactt tgccacggca ttactgttat tactcggtgg

tgcgggcatt attggcagcg tgattttcgg taaactgggt aatcagtatg cgtctgcgtt

ggtgagtacg gcgattgcgc tgttgctggt gtgcctggca ttgctgttac ctgcggcgaa

Dolphin - bottle-nose dolphin (*Tursiops truncates*)

atgcaaggag atgatcccaa cagcagtgat gagtctaatg gcagtgatga taccaattct

gaaggtgaca acaaccacag tagccgagga gatgcttctt ataactctga tgaatcaagt

gataatggca atgacagtga ctcaaaagga ggagaagaag gtgatagtga taacacatca

gatgctaatg atagtggtgg tgatggcaat ggtgacatgg ggagtgataa gaatggaaaa

Hummingbird - ruby-throated hummingbird (*Archilochus colubris*)

gtccatgtag cttaatcaaa gcatagcact gaagatgcta agacgacacc aataaaatgt

ccagggacaa aagactcagt cctaacctta ccgttgattg tcgccaaaca tatacatgca

agtatccgcg ctccagtgta aatgccctca atcaccttac caagacaaaa ggagcaggta

tcaggcacac tacaactgta gcccaaaaca ccttgctcag ccacaccccc acgggtactc

Opossum - gray short-tailed opossum (*Monodelphus domestica*)

gaattctatc aaacattcag agaacagtga atcccaatac aagacaaact atttgacata

ataagcaaag agggagttct accaaactcc tttcatgaca caaacatggt gctgattcca

aaaccaggca ggtcaaaagc ggagaaagaa aactatacac caatctccct aatgaatata

gatgcaaaaa tcttaaatag gatactagca aaaagactcc agcaagtgat cagaagggtc

Nematode worm (*Caenorhabditis elegans*)

gatcaatatg aatcttgctt gaatgaccaa ctttttgaga aactataata gtgtgtttta

attttgaatc agaaacatgt ttaaaacatt ttgtagatta gattcagagg attttccaaa

catcttgaat ttttgaattc caattcctaa attcaaatgg gtgaaaataa acattttaga

attagcaaga actgatcatt tcatcaaaca tgttattttc acacacacac acacacacac

Additional GENODEON Concept Designs (UI design refinements of the original design published at the top of this document):