VARIATION IN DNA CONTENT AND REDUNDANCY-POSSIBLE SIGNIFICANCE TO FOREST TREE BREEDING'

Richard B. Hall, Gerald R. Stairs,
and Jerome P. Miksche [2]

Research in the early 1940's led to the acceptance of deoxyribose nucleic acid (DNA) as the active genetic material in the living cells of all plants and animals. Today it is well known how DNA is organized into genes capable of coding for and controlling the many different enzymes which determine cellular activities. Much is also known about how this same DNA serves as the bridge of genetic information from one generation to the next. However, two relatively recent discoveries have demonstrated that we still have much to learn about DNA and its total role in cell biology. The first discovery was that not all normal individuals of the same species have the same DNA content/cell; indeed there often is a trend of increasing DNA content with increasing latitude of origin (Miksche, 1968, 1971; El-Lakany and Sziklai, 1971). Secondly, not all of the DNA of a cell is devoted to unique genetic messages; in fact, some DNA sequences are present in as many as a million copies per cell (Britten and Kohne, 1968). These discoveries have led to many new hypotheses of DNA organization and function in the genomes of higher organisms. This paper will summarize some of these new hypotheses and emphasize the implications they may have for forest tree genetics and tree improvement.

DISCUSSION

The Concept of DNA Constancy and its Refutation

Soon after the affirmation of DNA as the genetic material a general concept of DNA constancy became established in the dogma of molecular biology. This constancy concept held that, barring ploidy exceptions, the amount of DNA/cell was a constant between and within individuals of the same species (Boivin, *et al.,* 1948; Mirsky and Ris, 1949). On theoretical grounds it was argued that all individuals of the same species should have the same number of genes, and since genes are made of DNA, all individuals of the same species should have the same DNA content/cell. However, there was no systematic attempt to experimentally verify the hypothesis.

Today the constancy concept need no longer be accepted. Though it is still believed that all individuals of the same species have the same number of genes, recent evidence indicates that multiple and variable numbers of copies of some of these genes can be present. Other, non-genic, functions of DNA may also exist, and these could lead to further fluctuations in DNA content. Comparisons of DNA contents of both closely and distantly related organisms show there is no constant relationship between DNA content/cell and the complexity of an organism. *Pinus resinosa* Ait. has twice as much DNA/cell as *Pinus banksiana* Lamb. (Miksche, 1967), and the latter species has ten times as much DNA/cell as does the most complex organism—man (Shapiro, 1971). This suggests that the total amount of DNA present in the cell does not directly reflect the total number of unique genes in the genome.

Experimental evidence for :.'the refutation of the constancy concept is now available. Miksche has demonstrated statistically significant variations of 1.5, 1.6, and 1.7 fold in DNA content/cell for provenances of *Pinus banksiana, Picea glauca* (Moench) Voss, and *Picea sitchensis* (Bong.) Carr. respectively (Miksche, 1968, 1971). In addition, El-Lakany and Sziklai (1971) found 1.9 fold variation in both coastal and inland forms of *Pseudotsuga menziesii* (Mirb.) Franco, and Dhir (1971) found greater than two-fold variation in *Pinus resinosa.* These findings are compatible with and help to explain other results which have described intraspecific variation in interphase nuclear volume (Mergen, 1967; Miksche, 1968; EI-Lakany and Sziklai, 1971) and total length of the metaphase chromosomes (Burley, 1965).

In several of these cases a moderate to strong correlation was found between increasing latitude of seed origin and increasing DNA content, nuclear volume, or chromosome length. It should be emphasized that each of these studies was conducted on embryo or few-day-old seedling root tip material. At best, the time span over which the provenances were grown in a common environment before analysis was quite short. Work with other plants has indicated that there can be environmental modifications of DNA content over a single generation (Evans, 1968; Durrant and Jones, 1971; LaCour, *et al.,* 1956). Consequently, additional long-term common environment studies are needed to determine if provenance differences in DNA content are permanent.

Evidence for Redundancy

If, as the previous paragraphs have suggested, not all the DNA is devoted to unique sequences, then some of the DNA sequences must be repeated in the genome. Experimental evidence of this was first obtained in 1964 using techniques to assess the rate of reannealing of denatured DNA (Britten and Kohne, 1968). The pres-

ence of redundant sequences in the DNA isolated from an organism leads to reannealing rates that are much faster than would be expected if all the DNA sequences were unique. To date, every higher organism studied in this way has been found to have some portion of its genome composed of redundant DNA sequences; these sequences, up to several hundred nucleotides in length, are repeated many times in the DNA of each cell's nucleus. The magnitude of redundancy is variable among species. Estimates in different *Drosophila* species have ranged from 5 to 15 percent (Laird, 1969); in calves a value of 40 percent has been determined (Britten and Kohne, 1968); and for wheat somewhat greater than 50 percent redundancy has been found (Bendich and McCarthy, 1970). Estimates of redundancy in conifer genomes indicate that between 70 and 80 percent of the DNA of each cell is redundant (Miksche and Hotta, 1971).

There are indications that at least a portion of the inter- and intra-specific differences in DNA content/cell can indeed be attributed to changes in the amount of redundant sequences in specific genomes. The two-fold difference in DNA content of two *Drosophila* species is accompanied by a corresponding shift in redundancy (Laird, 1969). DNA variations in seven coniferous genera and three *Picea glauca* provenances are associated with changes in percent redundancy, such that the higher the DNA content per cell, the greater the proportion of repeated sequences (Miksche and Hotta, 1971). However, differences in DNA content were not associated with differences in DNA base composition (Miksche and Hotta, 1971, 1972). This could be due to large-scale redundancy of only a few sequences whose average base composition was close to the ever-all average base composition for a species. Alternatively, if most of the sequences of a genome were involved in redundancy to approximately the same extent, then the resulting increases in DNA content would not be accompanied by shifts in average base composition. Reannealing data favor the latter explanation.

Physiological and Evolutionary Implications of Redundant Genes

It must be assumed that redundancy serves some positive function(s) in the genome. Otherwise, one would expect this phenomenon, which places a heavy demand on cell nucleotide resources, would be selected against and eliminated in evolution. One likely explanation is there is a redundancy of genes which code for products needed in large amounts by the cell. There is experimental evidence for large-scale redundancy in the genes coding for animal hemoglobin (Williamson *et al.,* 1970) and histones (Kedes and Birnstiel, 1971). The hundred to many thousand-fold redundancy of the genes which code for ribosomal RNA is well-established (Watson, 1970; Hotta and Miksche, 1972). In line with the often-observed correlation between increasing nuclear volume or DNA content/cell and increasing latitude of origin, Stebbins (1966) postulated that redundancy in certain gene systems may be involved in adaptations to cold climates.

Redundancy of genes coding for specific products might exist as tandemly repeated sequences along a single molecule of chromosomal DNA, as single units present at many loci throughout the genome, or as identical multiple strands of DNA molecules traversing portions of the chromosomes. However, each of these alternatives pose the same theoretical problem. If independent mutation and recombination of the many copies of a single gene are not prevented by some mechanism, the gene may with time evolve to have more than two alleles per diploid genome and thus cease to behave as a simple Mendelian factor in inheritance. To overcome this objection, Callan (1967) proposed that a master-slave organization existed for redundant genes. Under this hypothesis each gene consists of a master sequence and its copies are slaves. It is only the master sequence that can recombine with the homologous allele, and it is only in the master sequence that mutation is of genetic importance, because all of the slave sequences are copied afresh or repaired each generation, using the master sequence as a template.

It may well be that such a mechanism operates for some genetic systems in higher organisms. Alternatively, genes exhibiting simple Mendelian inheritance may belong to the unique fraction of the genome, while the genes existing as redundant copies are subject to independent mutation and recombination. For such duplicated genes, natural selection would tend to maintain a sufficient number of copied genes capable of producing all functional, but not necessarily homogeneous, products needed by the cell. Such gene systems may be partially responsible for what we know as quantitative genetic inheritance. Just as additive, dominance, and epistatic relationships between genes can affect breeding techniques and results, it is possible that interactions between non-identical copies of a single gene may also have an effect. Hence, information concerning possible ties between redundancy and quantitative inheritance is important, particularly to breeders working with the highly redundant conifers.

Gene Control and Genecology

A second function which has been attributed to redundant DNA is the control and co-ordination of the many gene systems present in the eukaryotic genome. Recent models for gene control include redundant segments of DNA as central features in integrating the functioning of the various enzyme-coding genes (Britten and Davidson, 1969; Crick, 1971). With the evolutionary appearance of new redundancy and the mutation-directed divergence of old redundant sequences, new patterns of gene regulation could arise to adapt a species to new or changing environments.

The habitats of many tree species cover a wide ecological breadth. Rehfeldt and Lester (1969) have suggested that the particularly variable environments to which pioneer tree species are subjected have led to the evolution of genetic systems with a high degree of developmental homeostasis—fairly constant physiological and morphological response to a range of environments. Such homeostasis could be due in part to

redundant DNA control systems which activate the same developmental gene systems in response to varied stimuli.

Eventually the importance of redundant product and control genes in environmental adaptation should be determined in detail. In the meantime it would be well to determine what effect species variability in redundancy has on the progeny of intra- and interspecific hybrids. Do differences in the number of gene copies or control elements of the parental genomes cause difficulties in the offspring, or might it be possible to create progency with new and useful redundancy characteristics?

Control of Chromosome Structure and Behavior

A final category of speculated functions for redundant DNA is the control of chromosome structure and behavior. The constancy of metaphase chromosome morphology suggests that a recognition system must be built into the chromatin which directs the same set of folding and "packaging" steps at each nuclear division. This might depend on certain DNA sequences which are repeated at critical points along the DNA strand (Comings and Riggs, 1971; Sutton, 1972). Since chromosome morphology shows so little between- and within-species variability in most of the conifers, any redundancy in their genomes which is devoted to the control of chromosome folding must be very evolutionarily stable.

Redundant sequences have also been suggested as being important in mechanisms for the meiotic pairing of chromosomes (King, 1970; Comings and Riggs, 1971) and the attachment of the centromeres to the spindle fibers (Walker, 1971). Here again, it may be some time before cell biologists have worked out the details of such mechanisms. However, some of the practical questions they pose for tree breeding can be examined now. Will crosses between parents with different redundancy levels lead to chromosome pairing difficulties and possible sterility in the offspring? Also, do differences in centromeric redundancy affect meiotic and mitotic behavior in any way?

A Fortuitous Byproduct of Redundancy Research

Whatever their functions in the genome eventually prove to be, redundant sequences have already yielded benefits to the study of chromosomes. In the past few years several techniques have been developed which preferentially stain portions of the chromosome having concentrations of redundant sequences. The resulting visualization of the distribution of redundancy has not only aided speculation on the functions of redundancy; the distinctive banded appearance of individual chromosomes has led to new methods of karyotyping. In human genetics it is now possible to identify nearly all of the chromosomes by these methods (Dress and Shaw, 1971), and the results have been applied to a variety of medical, legal, and research problems. Further work is needed to adapt these staining procedures to tree species, but the benefits could be equally rewarding. The lack fo substantial morphological differentiation in conifer chromosomes has long stiffled their intensive study. If relatively simple techniques can be worked out for identifying each of the conifer chromosomes, they could be applied to such problems as the study of evolutionary relationships, the detection of genetic migration between populations, the verification of hybrids, and the correlation of specific chromosome abnormalities with morphological or physiological traits.

SUMMARY

As we have tried to indicate in this brief presentation, the study of variations in DNA content and redundancy has potentially great contributions to make to the understanding of coniferous genomes—including mechanisms of adaptation to the environment, influences on quantitative inheritance, determinants of chromosome structure, aspects of meiosis and mitosis, and new means of karyotype analysis. In a general sense, the more we know about the genomes of the species we deal with, the broader our opportunities are for their beneficial manipulation. To ignore the study of redundancy and its variation in coniferous genomes would be to ignore some 80 percent of the genetic story potentially available to us. Variation is the basic raw material in any type of breeding work; we have found that the organisms we work with have considerable variation in DNA content and redundancy. Now the task is to understand and use that variation.

LITERATURE CITED

Bendich, Arnold J., and Brian J. McCarthy. 1970. DNA comparisons among barley, oats, rye, and wheat. Genetics 65: 545-565.

Boivin, Andre, Roger Vendrely, et Mme. Colette Vendrely. 1948. L'acide desoxyribonucleique du noyau cellulaire, depositaire des caracteres hereditaires, arguments d'ordre analytique. Comptes rendus de L'Acadmie Sciences. 226: 1061-1063.

Britten, R. J., and D. E. Kohne. 1968. Repeated sequences in DNA. Science 161: 529-540..

Britten, Roy J., and Eric H. Davidson. 1969. Gene regulation for higher cells: a theory. Science 165: 349-357.

Burley, Jeffery. 1965. Karyotype analysis of Sitka spruce, *Picea sitchensis* (Bong.) Carr. Silvae Genetica 14: 127-132.

Callan, H. G. 1967. The organization of genetic units in chromosomes. J. Cell. Science 2: 1-7.

Comings, David E., and Arthur D. Riggs. 1971. Molecular mechanisms of chromosome pairing, folding and function. Nature 233: 48-50.

Crick, Francis. 1971. General model for the chromosomes of higher organisms. Nature 234: 25-27.

Dhir, Narinder K. 1971. Study of some seedling and nuclear characteristics of red pine *(Pinus resinosa* Ait.). Masters Thesis. University of Minnesota.

Drets, Maximo and Margery W. Shaw. 1971. Specific banding patterns of human chromosomes. Proc. Nat. Acad. Sci. 68: 2073-2077.

Durrant, A., and T. W. A. Jones. 1971. Reversion of induced changes in amount of nuclear DNA in *Linum.* Heredity 27: 431-439.

El-Lakany, M. H., and Oscar Sziklai. 1971. Intraspecific variation in nuclear characteristics of Douglas-fir. Advancing Frontiers of Plant Sciences 28: 363-378.

Evans, G. M. 1968. Nuclear changes in flax. Heredity 23: 25-38.

Kedes, Laurence H., and Max L. Birnstiel. 1971. Reiteration and clustering of DNA sequences complementary to histone messenger RNA. Nature New Biology 230: 165-169.

King, Robert C. 1970. The meiotic behavior of the *Drosophila* oocyte. Int. Rev. Cytology 28: 125-168.

LaCour, L. F., E. M. Deeley, and J. Chayen. 1956. Variations in the amount of Feulgen stain in nuclei of plants grown at different temperatures. Nature 177: 272-273.

Laird, Charles D., and Brian J. McCarthy. 1969. Molecular characterization of the *Drosophila* genome. Genetics 63: 865-882.

Hotta, H., and J. P. Miksche, 1972. Hybridization of coniferous ribosomal RNA to DNA. Eighth Central States Forest Tree Improvement Conference Proceedings. Columbia, Missouri, October, 1972.

Mergen, F., and B. A. Thielges. 1967. Intraspecific variation in nuclear volume in four conifers. Evolution 21: 720-724.

Miksche, Jerome P. 1967. Variation in DNA content of several gymnosperms. Can. J. Genet. and Cytol. 9: 717-722.

Miksche, Jerome P. 1968. Quantitative study of intraspecific variation of DNA per cell in *Picea glauca* and *Pinus banksiana.* Can. J. Genet. and Cytol. 10: 590-600.

Miksche, Jerome P. 1971. Intraspecific variation of DNA per cell between *Picea sitchensis* (Bong.) Carr. provenances. Chromosoma 32: 343-352.

Miksche, J. P., and Y. Hotta. 1971. DNA base composition and repetitious DNA in several conifers at interspecific and intraspecific levels. Paper presented at 11th Annual Amer. Soc. for Cell Biology Meeting, New Orleans.

Miksche, J. P., and H. Hotta. DNA base composition and repetitious DNA in several conifers. Chromosoma. In press.

Mirsky, A. E., and Hans Ris. 1949. Variable and constant components of chromosomes. Nature 163: 666-667.

Rehfeldt, G. E., and D. T. Lester. 1969. Specialization and flexibility in genetic systems of forest trees. Silvae Genetica 18: 118-123.

Shapiro, Herman S. 1971. Deoxyribonucleic acid content per cell of various organisms. pp. H-104 to H-114 in Handbook of Biochemistry, 2nd Ed. Herbert A. Sober, editor. The Chemical Rubber Co., Cleveland.

Stebbins, G. Ledyard. 1966. Chromosomal variation and evolution. Science 152: 1463-1469.

Sutton, W. D. 1972. Chromatin packing, repeated sequences and gene control. Nature New Biology 237: 70-71.

Walker, P. M. B. 1971. Origin of satellite DNA. Nature 229: 306-308.

Watson, James D. 1970. Molecular biology of the gene. pp. 541-543. 2nd ed. W. A. Benjamin, Inc., New York.

Williamson, R., M. Morrison, and J. Paul. 1970. DNA-RNA hybridization of 9S messenger RNA for mouse globin. Biochem. Biophys. Res. Comm. 40: 740-745.