CHROMOSOMAL CONTROL OF WHEAT ENDOSPERM PROTEINS
A Critical Review

F. GARCIA-OLMEDO, P. CARBONERO,
C. ARAGONCILLO, G. SALCEDO
Departamento de Bioquímica,
ETS Ingenieros Agrónomos,
Madrid, Spain

Abstract

CHROMOSOMAL CONTROL OF WHEAT ENDOSPERM PROTEINS: A CRITICAL REVIEW.
Progress made in the chromosomal location of structural genes for wheat endosperm proteins, and in the study of the regulation and quantitative expression of these genes, by using aneuploids and by related techniques, is critically evaluated. Recommendations for future work are proposed.

INTRODUCTION

The development of aneuploid stocks and chromosome substitution and addition lines of hexaploid wheat (Triticum aestivum L.), pioneered by Sears [1—3], opened a new possibility of genetic analysis, which has been extensively exploited by many workers to investigate the chromosomal location of genes affecting a great variety of characters. Morris [4] has periodically catalogued advances made in this area and Konzak [5] has recently reviewed the pertinent literature. Their compilations include the location of genes affecting quite a number of enzyme systems and endosperm proteins. The integration of these data into a univocal and non-redundant body of knowledge presents many problems. The purpose of this paper is to evaluate the present status of our knowledge about the chromosomal location of genes coding for endosperm proteins and about the regulatory and quantitative aspects of their expression. For practical reasons, we will focus our attention on the major endosperm proteins, characterized as such, and will only mention some enzyme systems for the purpose of illustrating some points. A survey of future developments will complete this review.
LOCATION OF STRUCTURAL GENES FOR ENDOSPERM PROTEIN COMPONENTS

General considerations

The assignment of the structural gene for a given protein to a particular chromosome using aneuploids has to be based on the observation that the protein is absent only in stocks lacking that chromosome but is present in stocks lacking any of the remaining chromosomes. However, even in this most favourable case alternative explanations, such as the possibility of triplicate genes coding for the same protein and a fourth gene affecting the expression of the triplicate set, cannot be excluded. Only when homoeologous chromosomes affect genetic variants of a given biochemical system (isozymes or isoproteins) can this possibility be reasonably excluded. When there is more than one chromosome whose absence suppresses a given protein, the assignment of the structural genes cannot be directly inferred and, as will be discussed later, additional information is required.

To ascertain the absence of a protein the analytical procedure has to be selective for that protein. It can be generally stated that enough resolution is not attained for most protein components of most endosperm extracts by using only one of the fractionation methods (electrophoresis, SDS-electrophoresis, electrofocusing, etc.). Only when a selective extraction procedure or staining method is available will one-dimensional separation suffice. Two-dimensional separations are more efficient in this connection.

When more than one protein is included in a given fraction (for example, in an electrophoretic band), quantitative changes in that band associated with the lack or the increased dosage of a chromosome can only give clues about the possible location of structural genes.

Characterization of protein fractions and individual proteins are required to interpret the genetic data, to overcome the possible lack of resolution of the analytical procedures, and to be able to compare the findings of different research groups.

It is often the case that assignments of structural genes for components of a given endosperm protein class reported by different groups are in conflict. In our opinion, these discrepancies are due to differences in the extraction procedure, the fractionation method, or the staining procedure, as well as to the lack of characterization of the individual components. In the following paragraphs we will deal specifically with the chromosomal location of structural genes for globulins, albumins, low molecular weight hydrophobic proteins, gliadins and glutenins.
### TABLE I. CHROMOSOMAL LOCATION OF GENES THAT CONTROL GLOBULINS, ALBUMINS, AND LOW MOLECULAR WEIGHT HYDROPHOBIC PROTEINS

<table>
<thead>
<tr>
<th>Protein class</th>
<th>Chromosomes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purothionins (apoprotein)</td>
<td>1AL 1BL&lt;sup&gt;a&lt;/sup&gt; 1DL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[6-8]</td>
</tr>
<tr>
<td>(digalactosyl diglyceride)</td>
<td>5A 5B 5DS</td>
<td></td>
</tr>
<tr>
<td>Globulins (fastest ‘doublet’)</td>
<td>1AL 1BL&lt;sup&gt;a&lt;/sup&gt; 1DL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[7]</td>
</tr>
<tr>
<td>Albumins</td>
<td>5AL 5BL 5DL</td>
<td>[9]</td>
</tr>
<tr>
<td>Albumins (PCS)</td>
<td>- - 3D</td>
<td>[10]</td>
</tr>
<tr>
<td>(Mb 0.19)</td>
<td>- - 4D</td>
<td></td>
</tr>
<tr>
<td>Albumins&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3AS 3BS 3D&lt;sub&gt;B&lt;/sub&gt;</td>
<td>[11]</td>
</tr>
<tr>
<td>(soluble in Tris buffer pH 8.7)</td>
<td>4A&lt;sup&gt;a&lt;/sup&gt; 4B&lt;sup&gt;a&lt;/sup&gt; -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6A&lt;sup&gt;a&lt;/sup&gt; -</td>
<td></td>
</tr>
<tr>
<td>Albumins</td>
<td>3A 3B 3D</td>
<td>[12]</td>
</tr>
<tr>
<td>(extract, 0.8M salt; 3M salt)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM-proteins (CM1, CM2)</td>
<td>- 7B 7D</td>
<td>[13]</td>
</tr>
<tr>
<td>(CM3)</td>
<td>4A&lt;sup&gt;a&lt;/sup&gt; -</td>
<td>[14]</td>
</tr>
<tr>
<td>70% EtOH extract</td>
<td>- 3B 3D&lt;sub&gt;A&lt;/sub&gt;</td>
<td>[15]</td>
</tr>
<tr>
<td>(non-gliadins)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4A&lt;sup&gt;a&lt;/sup&gt; 4B&lt;sup&gt;a&lt;/sup&gt; 4D&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- - 7D</td>
<td></td>
</tr>
<tr>
<td>70% EtOH extract (albumins)</td>
<td>- 3BS 3D&lt;sub&gt;B&lt;/sub&gt;</td>
<td>[16,17]</td>
</tr>
<tr>
<td></td>
<td>4A&lt;sub&gt;B&lt;/sub&gt; -</td>
<td>4D</td>
</tr>
<tr>
<td></td>
<td>- - 5D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 6B -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4A&lt;sub&gt;B&lt;/sub&gt; -</td>
<td>?D</td>
</tr>
<tr>
<td></td>
<td>- 7B 7D</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Absence of chromosome only decreases band or spot; extra doses enhance.

<sup>b</sup> Designation proposed by the reviewers.

Globulins, albumins and low molecular weight hydrophobic proteins

A rather confusing situation exists concerning the chromosome assignments of genes coding for endosperm proteins other than gliadins and glutenins. The chromosomes implicated by different workers in the control of components of these fractions are listed in Table I. There is considerable overlapping of the proteins included under each protein class entry. The equivalence of proteins studied by the different authors is obvious in some cases but in other cases further experimental work will be needed to ascertain the identity. A discussion of this problem will be attempted here.
The identity of \( \alpha \) and \( \beta \)-purothionins, extracted with petroleum ether, with the two globulins with greater electrophoretic mobility at pH 3.2 was first suspected from strictly biochemical data [18] and later confirmed by genetic analysis [6, 7]. This case is a good example of how the purification and characterization of the protein can overcome the lack of resolution of the analytical procedure used: \( \alpha_p \) and \( \alpha_D \) purothionins could not be separated electrophoretically, but their structural genes could be located by aminoacid analysis of the \( \alpha \)-fraction in different genetic stocks [7].

The remaining proteins listed in Table I are either albumins, extracted with water, or CM-proteins, low molecular weight hydrophobic components that are extracted with chloroform:methanol (2:1). All the CM-proteins and many of the albumins are also soluble in 70% ethanol [17].

The following discussion deals with the possible identity of proteins of these classes assigned to the same chromosome (Abbreviations: electrofocusing = EF; electrophoresis = EPH):

### Chromosome 3A


### Chromosome 3B

Equivalence of components 14—15 (EF X EPH) of Aragoncillo et al. [16] with component 76 (EPH) of Waines [15] and possibly with some component of complex band 3 (EF) of Noda and Tsunewaki [11] and component 6 or 7 (EPH) of Cubadda [12]. The uncertainty is ascribed to differences in extraction and separation methods used by the last two groups.

Less certain are the possible equivalences of components 6—7 (EF X EPH) of Aragoncillo et al. [16] with component 7 or 6 (EPH) of Cubadda [12] and even less, with component 4 (EF) of Noda and Tsunewaki [11], which has a lower isoelectric point. Waines [15] did not detect any components equivalent to 6—7 because these were probably included in the major complex band 90—97 of his one-dimensional electrophoretic pattern.

It is likely, from isoelectric point data, that components 6—7 and 14—15 of Aragoncillo et al. [16] are members of the 12 000 molecular weight family of \( \alpha \)-amylase inhibitors [19—20], which seems to include also the albumin of Ewart [21], albumin 13B of Feillet and Nimmo [22] and inhibitor \( \text{Am}^1 \) of Shainkin and Birk [23].
**Chromosome 3D**

Component 5 (EF X EPH) of Aragoncillo et al. [16] is identical with band 7 (EF) of Noda and Tsunewaki [11], is included in the major complex band 90–97 (EPH) of Waines [15], and is possibly equivalent to one of the components 1, 2, 3 or 5 (EPH) of Cubadda [12].

Bozzini et al. [10] assigned albumin PCS to this chromosome, but its isoelectric point is higher than that of component 5 of Aragoncillo et al. [16]. They also assigned albumin Mb 0.19 to chromosome 4D, but more recent evidence seems to indicate that this protein is actually the component 5 of Aragoncillo et al. [16] and therefore controlled by 3D [24—25]. Component Mb 0.19 seems also to be equivalent to albumin 13A of Feillet and Nimmo [22], to the albumin purified by Fish and Abbot [26], to the α-amylase inhibitor Aml₂ of Shainkin and Birk [23], to a component purified by O’Donnell and McGeeney [27] and possibly also to inhibitor 1 of Saunders and Lang [28]. Albumin Mb 0.19 seems to be the main component of a family of closely related α-amylase inhibitors (for a detailed discussion see Ref. [25]).

**Chromosome 4A**

Components 12–13 (EF X EPH) of Aragoncillo et al. [16] are included in band CM3 (EPH) of Aragoncillo [14] and in band 83 (EPH) of Waines [15]. Components with the same mobility as 12–13 are also associated with the D genome [16, 17].

Component 16 (EF X EPH) of Aragoncillo et al. [16] is included in band 69 (EPH) of Waines [15].

**Chromosome 4B**

Waines [15] suspected the location of structural genes for proteins included in bands 69 and 83 (EPH) in this chromosome. This was not confirmed by Aragoncillo et al. [16], using a higher resolution two-dimensional method (EF X EPH).

**Chromosome 4D**

Component 17 (EF X EPH) of Aragoncillo et al. [16] is included in band 65 (EPH) of Waines [15].

**Chromosome 7B**

Components 8–9 (EF X EPH) of Aragoncillo et al. [16] are included in band CM2 (EPH) of García-Olmedo and Carbonero [13] and are not detected
<table>
<thead>
<tr>
<th>Protein class</th>
<th>Chromosomes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gliadins (lactate buffer, 2M urea)</td>
<td>–</td>
<td>1DS [30, 31]</td>
</tr>
<tr>
<td>Gliadins (70% EtOH extract)</td>
<td>1A(7D) 1B(2B) –</td>
<td>[32]</td>
</tr>
<tr>
<td>– – 2D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>– – 6D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gliadins</td>
<td>– – 1D</td>
<td>[33]</td>
</tr>
<tr>
<td>Gliadins (70% EtOH extract)</td>
<td>2A^a 2B^a 2D^a</td>
<td>[15]</td>
</tr>
<tr>
<td>6A^a – 6D^a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gliadins (2M urea, 20% sucrose)</td>
<td>1AS 1BS 1DS</td>
<td>[34–36]</td>
</tr>
<tr>
<td>6AL 6BL 6DL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gliadins (8.5 mM lactate buffer)</td>
<td>1A 1B 1D</td>
<td>[37]</td>
</tr>
<tr>
<td>A-gliadin</td>
<td>6A 6B –</td>
<td></td>
</tr>
<tr>
<td>Glutenins</td>
<td>– 1DL</td>
<td>[38]</td>
</tr>
<tr>
<td>Glutenins</td>
<td>– 1DL</td>
<td>[39]</td>
</tr>
<tr>
<td>– 4DL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gluten components (unclassified)</td>
<td>1AS^a 1BS^a 1DS^a</td>
<td>[40]</td>
</tr>
<tr>
<td>2A^a 2B^a 2D^a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3AS^a – 3D^a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5AS^a 5BS^a 5DS^a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6AL^a 6B^a 6D^a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^a Absence of chromosome only decreases band or spot; extra doses enhance.

^b Designation proposed by reviewers.

by Waines [15] because they are included in band 90–97 (EPH), which is very complex, and do not stain well with Amido Black [29].

**Chromosome 7D**

Components 3–4 (EF X EPH) of Aragoncillo et al. [16] are included in band CM1 (EPH) of García-Olmedo and Carbonero [13] and in band 105 (EPH) Waines [15].

Gliadins and glutenins

The chromosomes implicated by different workers in the control of glutenins and gliadins are listed in Table II.
The chromosomal location of genes coding for components of the gliadin fraction is now well established. Wrigley and Shepherd [36], using a high-resolution fractionation by combined electrofocusing and electrophoresis, have fully confirmed previous assignments of the gliadin genes to homoeologous chromosome groups 1 and 6 [15, 30–37]. The possible involvement of homoeologous group 2 chromosomes in the control of gliadins will be discussed later.

The chromosomal control of glutenins has been studied by Orth and Bushuk [38] and by Bietz et al. [39]. Genes coding for five glutenin subunits have been located [39]. However, other gluten components remain to be assigned because either they are coded by triplicate (or duplicate) genes or, more likely, bands formed in (sodium dodecyl sulphate) SDS-electrophoresis include the products of more than one gene.

**Linkage maps**

Little effort has been made to map further the structural genes within chromosomes or chromosome arms. In this respect, Solari and Favret [41] undertook linkage studies of the gliadin genes. More recently Qualset and Wrigley [42] have made a more complete study using a two-dimensional fractionation method.

A different approach to this problem is based on the analysis of alien translocation lines. For example, *Agropyron elongatum-Triticum aestivum* transfers obtained by Sears [43] have been used to approximately map genes coding for endosperm proteins by Rodríguez-Loperena et al. [44] and by Hart et al. [45].

**REGULATORY AND QUANTITATIVE ASPECTS**

**Regulatory effects**

It is outside the scope of this review to discuss in detail the extensive literature on chromosomal effects on total endosperm protein and its aminoacid composition [46–52 and others]. However, a brief comment is warranted.

It can be said that no major regulatory effect on total protein level or composition has been ascribed so far to a chromosome, or chromosome arm, using aneuploids. The results of Bozzini and Giacomelli [53], concerning a drastic change in the proportion of albumins brought about by suppression of one arm of chromosome 2A, have not been confirmed by other workers using the same genetic stock.

Additive effects on total protein have been reported for chromosomes belonging to the seven homoeologous groups. There is considerable disagreement
with respect to the chromosomes involved between the various authors. This is probably due to two main causes: differences in the material investigated and lack of the needed corrections for yield factors. It is well known that there is a negative correlation between number of grains per spikelet (sterility) and protein. The most reliable data are those obtained with intervarietal and interspecific substitution lines, corrected for yield. Thus, chromosomes of groups 2 and 5 seem to account for a good part of the intervarietal differences in protein content [46-52].

The search for genes regulating the expression of one or a few genes coding for endosperm proteins has not been as successful as that for structural gene locations.

Shepherd [34] reported that four doses of chromosome 2A seemed to suppress a gliadin component and to promote the presence of a new band. This matter was apparently not pursued and no mention of it is made in the more recent report of Wrigley and Shepherd [36]. Waines [15] also implicated chromosomes of group 2 in the control of gliadins, but the low resolution for gliadins of the separation method used did not allow to check whether specific gliadin components were missing when each of the chromosomes of this group was absent.

Orth and Bushuk [38] reported the repression of the synthesis of some glutenin subunits by four doses of either chromosome 2B, 3B or 6B. The more recent report by Bietz et al. [39] clearly shows that no such repression takes place.

Aragoncillo et al. [16] concluded that the structural gene for non-gliadin component 2 was located in chromosome 6B and that its expression was apparently repressed by four doses of chromosome 7B in the absence of chromosome 7D. Sears suggested later that this result could be due to segments of the 6B chromosome from cv. Hope, still present in the nulli 7D-tetra 7B stock used in our study, and supplied a new stock that had undergone two further backcrosses to Chinese Spring wheat. Analysis of the new stock confirmed Sears' hypothesis.

Studying the chromosomal control of lipopurothionins, García-Olmedo et al. [6] found that the short arm of chromosome 5D was apparently required for the expression of the structural genes for the apoprotein, which were located in chromosomes 1AL, 1BL and 1DL. Further genetic and biochemical evidence [8, 54] demonstrated that chromosome 5D (short arm) actually affected the level of digalactosylglyceride, which was required for solubility of the lipopurothionins in the extraction solvent.

A similar effect to that initially found in the case of lipopurothionins [6] has been recently reported for phosphodiesterase by Wolf et al. [55]. A dosage-dependent regulatory effect has been ascribed to group 5 chromosomes that affects the structural genes for the enzymes, which are located in group 3 chromosomes. Characterization of the enzymes and further genetic evidence will
be needed to discriminate a true regulatory effect from alternative explanations of the reported results.

Genome interactions

It should be mentioned in this review that aneuploids have been of great help in clarifying inter-genome interactions. A alloplloid is in fact a 'permanent heterozygote' in which positive and negative heterotic interactions are effectively fixed. In the case of alloploids inter-genome heterosis is difficult to assess because it is superposed with positive and negative effects associated with the change of ploidy level. Therefore, both effects have to be considered jointly. Although in many cases the expression of homoeoalleles seems to be additive, García-Olmedo et al. [6] have proposed different modes of inter-genome complementation: enzyme subunit complementation, holoprotein completion complementation, homoeoallelic dosage compensation, and complementation at the level of metabolic pathways.

Chromosome dosage responses and dosage compensation

Although dosage effects have been repeatedly mentioned in connection with the investigation of chromosome-protein associations in wheat [12, 15, 16, 34, and others] and dosage compensation has been suspected to occur in a few of the cases [6, 12, 14, 56], these effects had not been investigated in a quantitative way. Such an investigation was undertaken by us in connection with the coordinated research programme on the use of aneuploids for wheat protein improvement.

Chromosome dosage responses for a group of six wheat endosperm proteins were investigated, using nulli-tetrasomic lines of Chinese Spring [57]. The output of each locus was investigated as a function of its own dose and of that of its homoeologue. Quasi-linear gene dosage responses were observed for all the proteins studied. However, for three of them, dosage compensation occurs: for a given dosage of its structural gene, the amount of the protein is 30–80% higher when the chromosome carrying the homoeologous gene is absent.

In a related study [58] two allelic proteins from the endosperm of the tetraploid wheat were investigated. The net output of protein molecules was measured for each of the alleles at 1, 2 and 3 doses. A linear dosage response was observed for both of them, but the output of one of the alleles was about double that of the other. These effects were observed for the parental material, the reciprocal F₁ generations and the segregating F₂ generation.
FINAL CONCLUSIONS

From the above discussion the following general conclusions can be drawn:

(a) It seems urgent that doubts about the possible equivalence or identity of proteins whose structural genes have been assigned to the same chromosome be resolved with the co-operative effort of the authors involved. Reports of new assignments should include an effort to face this problem.

(b) Further biochemical research is needed, in most cases, to translate electrophoretic bands, or spots, into well-characterized molecules. This is essential for the correct interpretation of most of the genetic results.

(c) Structural genes accounting for 50—70% of the endosperm protein have been located so far. Further research is needed to locate the genes coding for the remaining major endosperm proteins.

(d) Chemists working on the purification and characterization of endosperm proteins should consider the possibility of checking their distribution in Chinese Spring and its aneuploids.

(e) Regulatory and quantitative aspects of the expression of genes coding for endosperm proteins should be actively investigated in order to find more effective ways of genetic manipulation to increase protein quantity and quality.

ACKNOWLEDGEMENTS

This review is dedicated to Dr. E.R. Sears. A similar discussion was presented at a symposium on Genetic Aspects of Cereal Proteins organized in his honour by the American Association of Cereal Chemists, in San Francisco, California, 23—27 October 1977.

REFERENCES

[44] RODRIGUEZ-LOPERENA, M.A., ARAGONCILLO, C., TORRES, J.V., CARBONERO, P.,
    (1971) 9.
    Genetics 74 (1973) 5239 (Suppl. 2).
    EUCARPIA (1977) 9.
    (Suppl. 2).
    (in press).
    Columbia, Mo. (1973) 843.
[57] ARAGONCILLO, C., RODRIGUEZ-LOPERENA, M.A., SALCEDO, G., CARBONERO, P.,
[58] SALCEDO, G., ARAGONCILLO, C., RODRIGUEZ-LOPERENA, M.A., CARBONERO, P.,