Chromosoma (2009) 118:391–400
DOI 10.1007/s00412-009-0205-9

RESEARCH ARTICLE

2C or not 2C: a closer look at cell nuclei and their DNA content

Johann Greilhuber · Jaroslav Doležel

Received: 1 December 2008 / Revised: 28 January 2009 / Accepted: 3 February 2009 / Published online: 26 February 2009

Abstract The life cycle of animals and plants involves changes in chromosome number (nuclear phase) and sometimes even the karyotype, and consequently the DNA content of a nuclear genome is not static in time. Thus, in order to interpret DNA content data, it is important that the status of the materials from which DNA content is estimated be precisely defined. The previously proposed distinction between "holoploid" (C) and "monoploid" (Cx) genome size covers the most frequent states of plant and animal nuclear genomes. However, restricting nomenclature to just C and Cx still leaves a number of unresolved problems. Here, we propose an extension of the C-value terminology to handle a range of cytogenetic conditions, life cycle segments, and nuclear phases. A set of superscripts and subscripts are used in a formal way to identify life cycle segments and to express the quantitative relationship between these segments. A revision of the current usage of the holoploid chromosome number *n* was necessary to maintain the intimate link between *n* and C-value and between the monoploid chromosome number *x* and Cx-value. In this revision, haplophase individuals (i.e., "haploid" animals and "haploid" spontaneous or experimentally induced land plant sporophytes) have chromosome number *n* (not 2*n*, as is the current tradition) and thus nuclear DNA contents based on 1C. However, to avoid an unlimited progression of *n* levels due to generative polyploidy, zygotic individuals are assigned as 2*n* starting from the zygote, whatever their ploidy level. Their ploidy is indicated by multiples of the basic chromosome number *x*. The extended terminology for genome size should eliminate ambiguities in reporting DNA contents in both plants and animals.

Communicated by I. Schubert

J. Greilhuber (✉)
Department of Systematic and Evolutionary Botany,
University of Vienna,
Rennweg 14,
1030 Vienna, Austria
e-mail: johann.greilhuber@univie.ac.at

J. Doležel
Institute of Experimental Botany,
Laboratory of Molecular Cytogenetics and Cytometry,
Academy of Sciences of the Czech Republic,
Sokolovska 6,
77200 Olomouc, Czech Republic
e-mail: doleze@ueb.cas.cz

Introduction

The more that is known about living organisms, the more diversity is discovered at the population, organismal, and cellular levels. At the same time, there is a pressure to rationalize observations to make sense of the complexity which surrounds us. Over 50 years ago, the measurement of DNA content in a range of animal tissues led to the recognition that it follows a set of simple multiples of a basic quantity, i.e., onefold, twofold, fourfold, etc. (Swift 1950). These classes are now referred to as "C-values" (2C, 4C, 8C...), as coined by Swift (1950), with the "C" referring to the word "constant" (Greilhuber et al. 2005). The C-value terminology has been retained to the present day, even in the face of difficulties, particularly in the context of polyploid genomes. The major problem which has arisen is that there is no clear relationship between C-value and genome size, a fact which has become ever more apparent as whole-genome sequencing is applied ever more widely. To avoid some ambiguities, an updated terminology for genome size was developed, in which the specific terms "monoploid genome size" and "holoploid genome size" and their abbreviations Cx-value and C-value were suggested to denote the DNA content of, respectively, the monoploid chromosome set (*x*)

and the holoploid chromosome complement (*n*; Greilhuber et al. 2005). The literal meaning of the neologism "holoploid" is "complete with regard to ploidy level." This improved terminology offers the following advantages: (1) the terms "genome" and "genome size" can retain their broad meaning; (2) the term "holoploid" is meaningful and is complementary to "monoploid"; and (3) the symbol Cx is self-explanatory, thanks to its reference to the chromosome number symbol *x*. It is particularly useful in the context of comparisons between closely related taxa having differing levels of ploidy, such as in reproduction mode screening (Matzk 2007).

The suggestion of Greilhuber et al. (2005) marked an improvement in genome size terminology. However, for the sake of simplicity and accessibility, certain scenarios were deliberately ignored—for instance, the presence of allosomes, which cause genome size to be both sex and gamete dependent in animals and bryophytes and sporophyte and spore dependent in some angiosperms. In human, for example, the large size difference between the X and the Y chromosomes results in a significant difference between the DNA content of the X- and the Y-containing sperm and between the male and the female individual's genome size. There is a wealth of diversity in the life cycles of eukaryotes. Some of these variants are associated with unusual nuclear DNA contents, which are hard—or even impossible—to describe using the current C- and Cx-value terminology. It is not obvious, to give just a single example, what is (or are) the sporophytic and the gametophytic 1C-value(s) of the pentaploid complex *Rosa canina* nor to establish the base C-value from which the endosperm develops.

In this paper, we detail a number of life cycle variants which result in DNA contents not readily characterizable using the genome size terminology of Greilhuber et al. (2005). We use this opportunity to characterize these situations and to extend the existing genome size terminology to allow it to cover these cases in a way which leaves less room for ambiguity.

Extended genome size terminology

We introduce here a set of superscripts and subscripts to C and Cx (Table 1) and apply these in a formal way to

Table 1 List of terms and symbols used as superscripts and subscripts

Term	Symbol	Explanatory note
Superscript		
Dikaryon	Dk	In higher fungi (Dikarya), extended life cycle segment between plasmogamy and karyogamy
Diplophasic	d	With zygotic or unreduced chromosome number
Endosperm, secondary	E	Embryo-nourishing tissue in angiosperms
Female individual	F	In man and animals, an egg-producing individual
Gamete	G	Germ cell with the capacity to fuse with another germ cell of opposite sex
Gametophyte	Gph	Sexual generation in plants with alternation of generations
Germ line	Ge	In animals, cell line giving rise to gametes, which is separated early in development from the soma
Haplophasic	h	With meiotically (or otherwise) reduced chromosome number
Male individual	M	In man and animals, a sperm-producing individual
Megagamete	meG	Egg, egg cell
Megagametophyte	meGph	Female gametophyte; in seed plants, the embryo sac
Megaspore	meS	Female gametophyte-forming spore in heterosporous plants
Megasporophyte	meSph	Megaspore-forming asexual generation in seed plants
Microgamete	miG	Spermium, sperm cell
Microgametophyte	miGph	Male gametophyte; in seed plants, the pollen grain and pollen tube
Microspore	miS	Male gametophyte-forming spore in heterosporous plants
Microsporophyte	miSph	Microspore-forming asexual generation in seed plants
Monokaryon	Mk	In higher fungi (Dikarya), haplophasic hyphae before plasmogamy
Soma	So	Body of an organism as distinct from the germ line cells
Spore	S	Asexual germ type
Sporophyte	Sph	Asexual generation in plants with alternation of generations
Zygote	Z	Fusion product of male and female gamete, sexually formed germ
Subscript		
Chromosome	chr	chromosome
Individual genomes	A, B, C...	In allopolyploids, the constituent monoploid genomes

identify life cycle segments and to express the quantitative genome size relationships between these segments. Thus, the general structure of the modified C- and Cx- descriptors is: ^{superscript}level$C_{subscript}$ and ^{superscript}level$Cx_{subscript}$, the "level" being an integer.

The following set of rules is proposed:

1. A superscript is added to C or Cx, indicating the *generation* or *life cycle* segment in plants (sporophyte, gametophyte, endosperm), type of animal (e.g., a worker or a queen in ants, etc.), and its sex (male, female) and if reduced or non-reduced (haplophasic or diplophasic). The relevant chromosome constitution can be added in parentheses. Individual elements within the superscript are separated by a dot (·).
2. Where a superscript is attached to the C- or Cx-value of a given organism, the genome size may differ from that described by C- or Cx-value lacking any or carrying a distinct superscript. This measure is required to accommodate the observation that genomic DNA content can be dependent on which life cycle segment is being considered.
3. In allopolyploids, the subscript specifies the constituent monoploid genomes, by employing a relevant set of capital letters. This can be useful when individual Cx-values are being quantified. The number of letters in the subscript indicates the ploidy level of the life cycle segment. For example, $1C_{AB}$ indicates the haplophasic (gametic) unreplicated DNA content of an allotetraploid having A and B as its parental genomes.
4. A "C" lacking any superscript indicates the mean genomic DNA content (e.g., averaged over both male and female animals) and/or that no specification has been attempted. If a polyploid's Cx-value lacks a subscript, this can be understood as representing the mean of all its constituent monoploid genomes.
5. The proposed terminology demands a strict arithmetical relationship between the chromosome numbers *n* for haplophase and 2*n* for diplophase. In this regard, we choose not follow the long-standing cytogenetic tradition, in which haplophasic animals and sporophytes are given 2*n* chromosome numbers (see below). However, to avoid an unlimited progression of *n* levels (*n* is generally called the haploid chromosome number, although we prefer to characterize it as the "holoploid haplophasic chromosome number") due to generative polyploidization, zygotic individuals should always be assigned as "2*n*" whatever their ploidy level (explained in more detail below). This, of course, reflects current usage.

Symbols within a superscript can be combined. For example, d·meGph indicates a diplophasic megagametophyte or embryo sac, h·Sph a haplophasic sporophyte (a "haploid" plant), d·meSph a diplophasic megasporophyte (such as a "female" *Melandrium album* plant). The terms haplophasic, diplophasic, triplophasic, pentaplophasic, etc. are consistently used here to characterize *life cycle* segments (Fig. 1) and are associated with, respectively, 1*n*, 2*n*, 3*n*, 5*n*,.... The terms haploid, diploid, triploid, pentaploid, etc. refer to the number of monoploid chromosome sets present in the holoploid genome (1*x*, 2*x*, 3*x*, 5*x*,...). A formal list of relevant definitions can be found in Greilhuber et al. (2005).

A dip into the exotic waters of genome size

Allosomes Allosomes or heterosomes, i.e., the heteromorphic sex chromosomes of animals and bryophytes and the allosomes of dioecious plants which determine megaspory and microspory (Rieger et al. 1991), can generate genome size differences between particular life cycle segments. Since dioecious plant sporophytes are either microsporous or megasporous, their designation is, respectively, ^{miSph} and ^{meSph}. Their chromosome constitution may be added to this superscript. The diplophasic genome size of human presented in this way as measured by Greilhuber et al. (1983) would be ^{M}2C=6.08 pg DNA (male) and ^{F}2C= 6.38 pg DNA (female). The female chromosome constitution can be indicated as part of the superscript; thus, $^{F(2n=44+XX)}$2C = 6.38 pg DNA or $^{F(2n=44+XX)}$1C = 3.19 pg DNA. For the male, ^{M}1C=3.04 pg DNA or half of the ^{M}2C-value. The mean 1C-value of sperm cells is ^{miG}1C=..pg DNA and the 1C-values of male- and female-determining sperm cells are, respectively, $^{miG(n=22+Y)}$1C = ..pg DNA and $^{miG(n=22+X)}$1C = ..pg DNA.

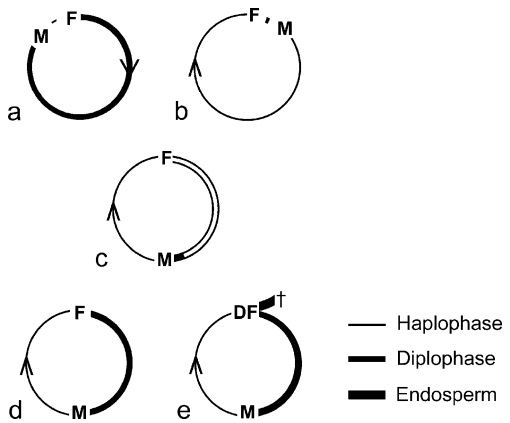

Fig. 1 The alternation of nuclear phases in sexual eukaryotes. **a** Gametic in diplonts (e.g., animals), **b** zygotic in haplonts (e.g., many green algae), **c** zygotic in dikaryohaplonts (fungi), **d** intermediate in diplohaplonts (e.g., embryophyta without secondary endosperm), **e** intermediate in diplohaplonts with secondary endosperm (angiosperms). Extended lines indicate vegetative growth of the life cycle segment via mitotic division, as opposed to those (zygote and gametes) in which mitosis does not occur. *DF* double fertilization, *F* fertilization, *M* meiosis, † exit of endosperm. The figure is modified from Widder (1967, Fig. 5)

As a second example, the dioecious plant *M. album* has different 2C-values for the "female" (2*n*=24, XX), "male" (2*n*=24, XY), and "supermale" (2*n*=24, YY). The corresponding genome sizes under the proposed terminology are $^{meSph(2n=22+XX)}2C = 5.73$ pg DNA, $^{miSph(2n=22+XY)}2C = 5.85$ pg DNA, and $^{miSph(2n=22+YY)}2C = 5.97$ pg DNA (Vagera et al. 1994).

B chromosomes B (or supernumerary or accessory) chromosomes are present in many plant and animal species. They are not essential for viability and, as a result, their number can vary both among individuals and even among cells within a single individual (Jones and Rees 1982). The presence of B chromosomes affects the nuclear DNA content, so C-values need to indicate the chromosome constitution as part of the superscript. Thus, $^{Sph(2n=6+B)}2C = ..$pg DNA would indicate a genome composed of six A chromosomes and one B chromosome per diploid sporophyte nucleus.

Aneuploidy Aneuploid cells or individuals have a chromosome complement which is not an exact multiple of *x*. The scheme allows for the inclusion of chromosome constitution, as for example, $^{Sph(2n=2x+1=13)}2C = ..$pg DNA, which indicates the genome size of an aneuploid sporophyte carrying an extra dose of one chromosome (trisomy). Since the zygotic chromosome number 2*n* includes the trisomic chromosome, aneuploidy is indicated by expressing 2*n* as 2*x*+1. As *x* itself can never be aneuploid, Cx cannot be based on an aneuploid genome. Note that aneuploidy differs from dysploidy, which refers to different *x* among related taxa (Rieger et al. 1991).

Haploidy Haplophasic plant sporophytes and animals originate either from the vegetative or somatic growth of a meiotically reduced cell (such as *in vitro* generated pollen embryogeny, parthenogenesis of the egg cell, or synergid embryos) or following chromosome elimination in the zygote during embryogenesis (or failure of karyogamy; Nogler 1984). These individuals have a chromosome number *n*, identical to the gametophyte (see, e.g., Rutishauser 1969; Rieger et al. 1991). The DNA content of a haplophasic non-replicated somatic nucleus is 1C. In a monohaploid sporophyte, *n*=*x* and therefore the C- and Cx-values are identical: $^{h\cdot Sph}1C=1Cx$. Note that spontaneous polyhaploid sporophytes are haplophasic but not haploid. Thus, in a dihaploid sporophyte with *n*=2*x* chromosome number, $^{h\cdot Sph}1C=2Cx$.

The dikaryon of higher fungi Nuclear-phase alternation is commonly described as being either gametic in diplonts (Fig. 1a; animals), zygotic in haplonts (Fig. 1b; green algae), or intermediate in diplohaplonts (Fig. 1d, e; gymnosperms and angiosperms). However, in the Basidiomycetes and Ascomycetes fungi, a considerable time elapses between plasmogamy and karyogamy (Fig. 1c), during which the mycelium grows as a dikaryon. After karyogamy, meiosis follows. As the indexed C-value terminology is concerned with nuclear DNA content and not with cellular DNA content, it seems justified to indicate the non-replicated and replicated holoploid genome size of a nucleus during the monokaryon stage as, respectively, $^{Mk}1C$ and $^{Mk}2C$ and during the dikaryon stage as $^{Dk}1C$ and $^{Dk}2C$. If required, the individual genomes can be indicated by a subscript. The genome size of the organism at this dikaryotic stage is given by the mean of the two contributory genomes. While the fungus during the dikaryophase may be called genetically diploid, it is nevertheless haplophasic from the viewpoint of nuclear genome size.

Do "haploids" have 2n chromosome number and 2C holoploid genome size? It has long been common practice for cytogeneticists to assign 2*n* to haplophasic ("haploid") animals (e.g., male Hymenoptera) or haplophasic higher plant sporophytes (see Levan and Müntzing 1963; John 1990; Singh 2003). Thus, John (1990) in the context of animals states that "In haploid organisms 2*n*=*x* so that whereas *x* is an algebraic term, *n* is not." However, we cannot ascertain the basis for this customary nomenclature. The departure from the algebraic relationship prioritizes the *life cycle* segment over the *nuclear phase*. Note that, in contrast, convention has it that the non-reduced embryo sac in an apomictic plant is given the chromosome number 2*n* (Nogler 1984). Thus, usage of the symbol 2*n* is inconsistent.

The justified use of 2*n* in polyploid series at first sight supports the assignation of 2*n* to haplophasic individuals as well since the polyploid sporophyte is logically described as 2*n* irrespective of its ploidy level. Thus, each new cycle of reproduction starting with the zygote begins with 2*n*, and ploidy level is specified by the multiplier of *x*. However, the potential complication of applying 2*n* to haplophasic individuals can and should be avoided because either they do not reproduce sexually (although exceptionally, polyhaploids with an even number of *x* may do so) and are therefore a dead-end in the reproductive cycle or they do so by means of a modified meiosis, whereby gametes with *n* chromosomes are produced, as in male hymenoptera (John 1990). Given the tight relationship between the C-value and *n* on the one hand and between the Cx-value and *x* on the other (Greilhuber et al. 2005), we suggest a change in convention, according to which haplophasic sporophytes and animals are assigned a chromosome number *n*, with 1C being given as the non-replicated holoploid genome size. This applies equally to monohaploids and polyhaploids. However, for polyhaploid individuals, which can reproduce and are members of a new population (biospecies), a 2*n* sporophytic chromosome number and a 2C non-replicated holoploid genome size should be assigned.

Generative polyploidy Members of a polyploid series are all assigned the chromosome number $2n$, with the ploidy level indicated by x (e.g., $2n = 2x = 14$, $2n = 3x = 21$, $2n = 4x = 28$, etc.). As a result, C-values are multiples of Cx-values. This represents a tidy solution, as long as the taxa do not intercross with one another, but rather represent distinct reproductive communities. However, where the parent(s) and offspring differ in their ploidy level, assigning a C-level is not simple. For example, does a triploid progeny ($2n = 3x = 21$) of a $2x$ by $4x$ cross have a 3C or a 2C genomic DNA content? Similarly, does a tetraploid individual ($2n = 4x = 28$) obtained from backcrosses between a triploid $3x$ hybrid and its diploid parent have a 4C or a 2C genomic DNA content? The answer should be that both have their genomes in 2C because the chromosome number correctly assigned to these plants is always $2n$. Following the proposal of Greilhuber et al. (2007), therefore, we suggest that nuclear DNA content histograms should label the peaks of the various polyploid taxa with the Cx symbol, as this provides a means of avoiding confusion between *generative polyploidy* and *nuclear phase* (Fig. 2).

Apomixis and parthenogenesis Some plant and animal species can reproduce asexually *via* apomixis or parthenogenesis. Offspring of such a mother individual has an unreduced chromosome number (i.e., they are diplophasic)

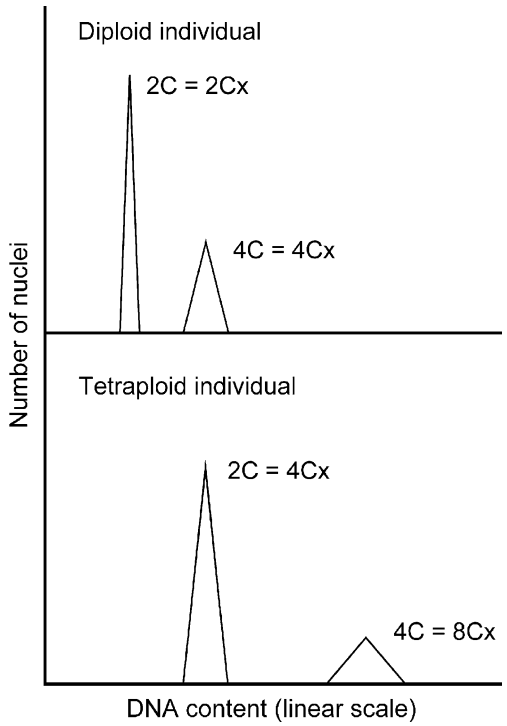

Fig. 2 A schematic nuclear DNA content histogram of a polyploid series of plants ($2x$ and $4x$), illustrating the use of the symbols C and Cx to label peaks of DNA content

which is identical to that of the mother. In angiosperms, apomixis (i.e., reproduction by seeds without sexual origin of the embryo) can be generated by apomeiosis (i.e., diplospory and apospory) or adventitious embryony (Nogler 1984). The non-replicated nuclei of a non-reduced (diplophasic) embryo sac are primarily in 2C. For instance, if genome size is determined in a megagametophyte, it can be specified as $^{d\cdot meGph}$2C=..pg DNA or $^{d\cdot meGph}$1C=..pg DNA—the latter being one half of the former. The superscript in both formulae is the same because it indicates merely the nuclear phase and generation in which the C-value was determined.

Interspecific hybrids and complex heterozygotes The genome constitution of interspecific hybrids and complex heterozygotes can be described using capital letters to indicate component (monoploid) genomes (e.g., AABB in an allotetraploid), and this is readily incorporated into the C-value formula. Thus, the non-replicated unreduced holoploid genome size of a diploid ($2n=2x$, AB) sporophyte is $2C_{AB}$=..pg DNA. A non-replicated reduced averaged holoploid genome size measured in young spore nuclei of a complex heterozygote would be assigned $^{miGph}1C_{(AB)/2}$=..pg DNA. The size of individual monoploid genomes can be estimated during and after meiosis in complex heterozygotes, such as in certain species of *Oenothera* (Onagraceae) and in *Rhoeo spathacea* (Commelinaceae). Thus, reduced non-replicated pollen nuclei have $1C_A$=..pg DNA and $1C_B$=..pg DNA, and dyad cells have $2C_{2A}$=.., $2C_{2B}$=..pg DNA. In particular cases, it may be possible to determine the size of the individual parental genomes in an allopolyploid. The genome symbol may then be added to the C-value term—so for example, in a tetraploid allopolyploid angiosperm sporophyte of genomic constitution AABB, $1Cx_A$=..pg DNA would indicate the size of the monoploid non-replicated A genome and $1Cx_B$ the size of the monoploid non-replicated B genome.

Anorthoploids and hybrids between different ploidy levels Anorthoploid organisms (see Levan and Müntzing (1963) for the confusing history of the term "anorthoploid") have an odd number of monoploid chromosome sets ($3x$, $5x$, etc.) in diplophase and are partly or completely sterile due to meiotic irregularity. Hybrid combinations between parents of differing ploidy level and spontaneous polyploids are not necessarily anorthoploid but may suffer from analogous meiotic disturbances. Due to these disturbances, any gametophytes formed tend to have irregular chromosome numbers, so ambiguity surrounding the 1C-value of such organisms is inevitable, given that no stable gametic 1C-value exists (see Bennett and Smith's (1976) original definition of 1C DNA content as gametic DNA content). A 1C-value would then be given by $^{d\cdot Sph}$2C/2=..pg DNA,

which is both arithmetically and biologically correct, because it represents the mean DNA content of a population of meiotic products or gones (Greilhuber et al. 2005).

"Permanent anorthoploids" or asymmetrically compensating allopolyploids A special type of fixed hybridity is associated with "permanently anorthoploid" organisms (Grant 1971) or rather more correctly "asymmetrically compensating allopolyploids." These include the *R. canina* complex (mostly 5*x* but also 4*x* and 6*x*), *Leucopogon juniperinus* (3*x*; Smith-White 1948; see also Grant 1971) and some members of the *Onosma helvetica* group (3*x*; Teppner 1971). In these organisms, specific meiotic mechanisms have developed to ensure that functional pollen and sperm cell nuclei all have the basic chromosome number *x* and that functional embryo sacs and egg cells have (2*n*−*x*). Upon fertilization, the 2*n* chromosome number is restored (Fig. 3). Of the several parental genomes present in the hybrid genome, only one is present in two copies and its chromosomes associate with one another during meiosis to form bivalents. The remaining genomes are present in single copy, so their chromosomes fail to pair. Only one set of the bivalent-forming genomes contribute to the functional microgametophyte. The complementary process occurs during megasporogenesis, where one set of the bivalent-forming chromosomes and all other unpaired chromosome sets contribute to the functional gametophyte. The term "permanent anorthoploids" is not fully justified because the meiotic mechanism (a type of meiotic drive) does not depend on and is not in all cases associated with anorthoploidy. For instance, *Rosa pomifera*, a tetraploid with 2*n*=28, forms seven bivalents and 14 univalents (7″ + 14′), while the hexaploid *Rosa agrestis* forms 7″ +28′ (Klásterská 1969). We suggest that the term "asymmetrically compensating allopolyploid" is a better descriptor than "permanently anorthoploid" because both megaspore and microspore meiosis generate distinct chromosome sets and this imbalance is corrected upon fertilization.

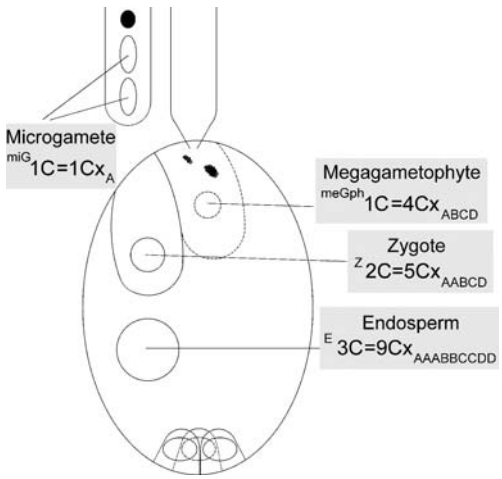

Fig. 3 The fertilized embryo sac of a *Rosa-canina*-type allopentaploid. The male gametes are overreduced (genome A) with genome size $^{miGph}1C=1Cx_A$. The embryo sac is partially reduced (genomes ABCD). The genome size of the pre-fertilization megagametophyte is $^{meGph}1C=4Cx_{ABCD}$. The zygote is diplophasic (2C) and pentaploid (5Cx), the endosperm is triplophasic ($^{E}3C$) and nonaploid (9Cx). For further explanation, see text

If a single 1C-value is to be given for an asymmetrically compensating allopolyploid, the formula 1C = 2C/2 = ..pg DNA can be used. For example, from somatic tissue of *R. canina* (2*n* = 5*x* = 35; assumed genome composition AABCD), the 2C-value was estimated by Greilhuber (1988) to be 2.86 pg DNA, so 1C = 2C/2 = 1.43pg DNA. The functional gametophytes are "overreduced" on the male side and "partly reduced" (but nevertheless still reduced) on the female side. So the functional *R. canina* microgametophyte is $^{miGph}1C = 1Cx = 0.57pg$ DNA and the megagametophyte $^{meGph}1C = 4Cx = 2.29pg$ DNA. However, even this represents an oversimplification, as it is likely that the individual chromosome sets differ in their DNA content. The suggested formulae are thus $^{miGph}1C = 1Cx_A$ = ..pg DNA and $^{meGph}1C = 4Cx_{ABCD}$ = ..pg DNA.

Somatic polyploidy (including endopolyploidy) Many nuclei in somatic cells amplify their DNA content during differentiation via endoreduplication (repeated DNA replication not followed by mitotic chromosome contraction and division, resulting mostly in decondensed interphase chromosomes or loose chromosome bundles with 2^{r+1} chromatids, where the exponent *r* is the number of endoreduplication cycles). Polyteny (formation of banded giant chromosomes) is a special form of endopolyploidy restricted to Diptera, Collembola, and some Ciliata). In specialized tissues of seed plants, such as for example the anther tapetum, modified mitotic divisions are common and these result in the production of one nucleus rather than two. In other situations, nuclei can fuse during ana-telophase to form nuclei with multiplied DNA amounts. Nuclei resulting from such processes are somatically polyploid. Somatic polyploidization can arise during both diplophase and haplophase but in angiosperms also in the secondary endosperm. A description of the DNA content in an (endo)polyploid nucleus is based on C, and not on *n*, because *n* as a chromosome number symbol does not indicate the replication state of the chromosome complement. The nuclear DNA content is then given on basis of the holoploid genome size (1C) of the corresponding segment of the life cycle, which may be specified by a superscript. The number of multiples of the holoploid genome size is specified by a number placed in front of the C-value formula. For example, a haplophasic antipodal nucleus of *Scilla bifolia* (Hyacinthaceae) has after the last embryo sac mitosis a $^{h\cdot meGph}1C$ DNA content. The

same nucleus during S-phase of the third endoreduplication cycle may have 12.7C DNA content (i.e., between 8C and 16C); its complete C-value term in this case would be $12.7^{h\cdot meGph}1C$. For another example, a triplophasic endopolyploid endosperm nucleus in maize may fall into the 12C class; its DNA content would be $12^{E(3n)}1C$ or simply $^{E(3n)}12C$.

Chromatin diminution In a few organisms (*Cestodes*, *Copepoda*, *Diptera*), particular chromosome segments (typically heterochromatin) are deleted in cells destined to become somatic, and/or complete chromosomes are eliminated. This "chromatin diminution" (Boveri 1899) results in a difference in DNA content between the germ line and somatic tissue. From a formal view, these cases could be treated in an analogous but inverted way to somatic polyploidy. Thus, C-values could be based on the germ line nuclei and the somatic nuclear DNA content given by the appropriate C-levels. However, the diminution process which occurs very early in ontogeny seems to be strongly regulated and is probably highly significant for the determination of cell and body size (see Gregory et al. 2000). Moreover, endopolyploidization may be superposed on this process. A more elegant solution is to distinguish the two genome sizes by superscripts. For example, in the copepod *Cyclops strenuus*, the DNA content of a non-replicated germ line nucleus is 4.28 pg. About 56% of the germ line DNA is missing in somatic nuclei and the "somatic genome size" is 1.87 pg DNA (Beermann 1977). We suggest that genome size in this species be specified separately as $^{So}2C=1.87$ pg DNA and $^{Ge}2C=4.28$ pg DNA. The relationship between the genome sizes of the soma and germ line can be given by $^{So}2C = 0.437^{Ge}2C = 1.87$ pg DNA and $^{Ge}2C = 2.289^{So}2C = 4.28$ pg DNA.

Endosperm—more than a nourishing tissue The endosperm surrounds the embryo in the seed of most flowering plants and provides the germinating seedling with nutrition. While the "primary endosperm" in a gymnosperm is merely the body of the megagametophyte and as such is a segment in the alternation of nuclear phases, in angiosperms the "secondary endosperm" is derived from a fertilization process and represents a nuclear phase of its own. However, as a terminal tissue, it is not *alternating*. Typically, in the course of "double fertilization," the central cell of the embryo sac is fertilized by one of the two sperm cells which are released from the pollen tube. Before fertilization, the central cell contains, according to embryo sac type, one or several nuclei (polar nuclei), which, together with the incoming sperm nucleus, form the primary endosperm nucleus. In exceptional cases (in apomicts with autonomous endosperm development), fertilization of the central cell does not occur, but nevertheless an endosperm develops (Nogler 1984). Thus, endosperm can be in a different nuclear phase (commonly, but imprecisely called "ploidy level"), such as diplophase, triplophase, pentaplophase, etc., according to the embryo sac type and the fertilization process involved. The *Polygonum* and *Allium* embryo sac types generate a triplophasic endosperm (chromosome number 3*n*, Fig. 4a, b), the *Fritillaria* type usually a pentaplophasic endosperm (5*n*), and the *Oenothera* type a diplophasic endosperm (2*n*). As the chromosome number and the nuclear phase in endosperm is specified using the symbol *n*, holoploid genome size is given on the basis of C, i.e., 3C, 5C, 2C, 4C, etc. (see in particular Greilhuber et al. 2005). For instance, in *Allium cepa* (2C=33.5 pg DNA, Bennett and Smith 1976), the non-replicated size of the endosperm genome can be given as 3C=50.25 pg DNA or $^{E}3C=50.25$ pg DNA. In the endosperms of diploids and polyploids, the same rules apply for the use of Cx as in the sporophyte. So, for example, in a tetraploid plant species (2*n*=4*x*) with a triplophasic endosperm (3*n*=6*x*), the DNA content of the non-replicated endosperm nuclei is 3C=6Cx.

The endosperm of hybrids formed from parents having different ploidy levels Given that the fertilization process involves the fusion of *reduced* gametes, it is logically consistent to use the number of gametic holoploid genomes combined in the hybrid zygote to determine the levels of *n* and C. In a *Polygonum*-type embryo sac, the endosperm will be 3*n*, in an *Oenothera* type 2*n*, and in a *Fritillaria* type 5*n*, irrespective of whether the ploidy levels of the

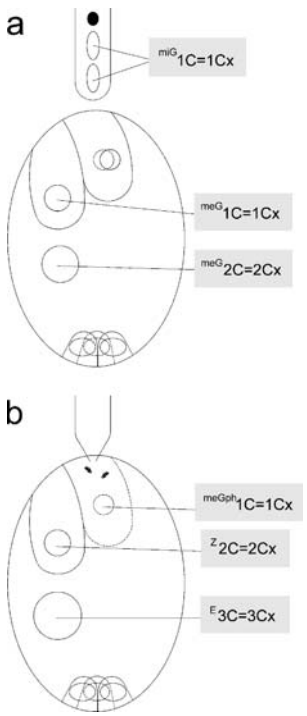

Fig. 4 Nuclear phases, ploidy, and genome sizes in a reduced *Polygonum*-type embryo sac, **a** before and **b** after fertilization with a reduced male gamete, where both the nuclear phase level and ploidy level are the same (*n*=*x*) for both. See text for details

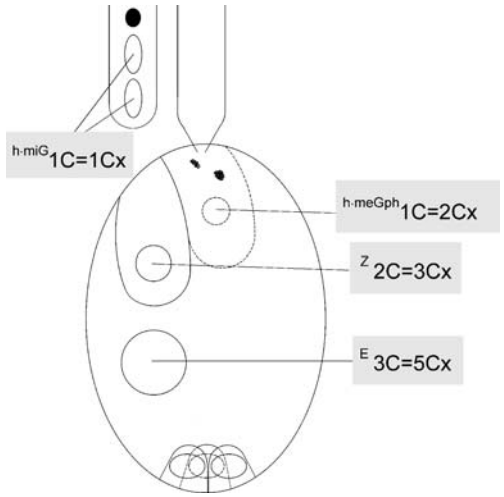

Fig. 5 Nuclear phases, ploidy, and genome sizes in a reduced embryo sac of a tetraploid plant ($2n=4x$) fertilized by a reduced male gamete from a diploid plant ($2n=2x$). The embryo sac before fertilization is haplophasic (n, 1C) and diploid ($2x$, 2Cx). After fertilization, the zygote becomes diplophasic ($2n$, 2C) and triploid ($3x$, 3Cx). The primary endosperm nucleus, which represents the fusion of three reduced genomes, is triplophasic ($3n$, 3C) and pentaploid ($5x$, 5Cx). For further explanation, see text

gametes are identical or not. Thus, these endosperm nuclei have unreplicated $^{E}3C$, $^{E}2C$, and $^{E}5C$. Consider a hybrid arising from a cross between a $4x$ maternal and a $2x$ paternal plant. Here, three haplophasic holoploid genomes (each with n chromosomes) fuse in the central cell of a *Polygonum*-type embryo sac (Fig. 5). As the endosperm nuclei have $3n$ and non-replicated the $^{E}3C$ condition, their genome size can be represented by $^{E}3C = 5Cx = ..pg$ DNA. The endosperm is triplophasic ($3n$) but pentaploid ($5x$). The complementary cross ($2x \times 4x$) would also produce a triplophasic ($3n$) but tetraploid ($4x$) endosperm with $^{E}3C=4Cx$.

Endosperms formed from unreduced gametes Some endosperms form from unreduced gametes, either in a programmed way (as during the apomixis) or as a by-product of irregular meiosis. Unreduced gametes are diplophasic ($2n$), and their ploidy depends on the number of monoploid chromosome sets present (Fig. 6). For example, in diploid barley ($2n = 2x = 14$) fertilized by unreduced pollen tube with sperm cells having $2n = 2x = 14$, the endosperm is tetraplophasic and tetraploid ($4n = 4x = 28$; $^{E(4n)}4C = 4Cx = 20.08$ pg DNA), while the embryo is diplophasic and triploid ($2n = 3x = 21$; $^{Z}2C = 3Cx = 15.06$ pg DNA; C-value from Doležel et al. 1998). An analogous example based on hexaploid wheat ($2n = 6x = 42$; $2C = 6Cx = 34.66$ pg DNA; C-value from Bennett and Smith 1976) produces a tetraplophasic dodecaploid endosperm ($4n = 12x = 84$; $^{E(4n)}4C = 12Cx = 207.96$ pg DNA) and a diplophasic nonaploid embryo ($2n = 9x = 63$; $^{Z}2C = 9Cx = 155.97$ pg DNA).

Endosperm of apomictic angiosperms The cytogenetic constitution of the endosperm of an apomict depends on the mode of formation of the seed. In adventitious embryony (where the embryo originates from a somatic nucellus or integument cell), the central cell of a haplophasic embryo sac is usually fertilized. Where the embryo sac is either of diplosporic or aposporic origin, the development of the endosperm can start either autonomously (independent of whether or not fusion of the polar nuclei has occurred) or pseudogamously after fertilization. Hybrid endosperms can arise when pseudogamy prevails (Nogler 1984). Here, also, the endosperm nuclear phase depends on the number of chromosome complements involved in the fusion, and the ploidy level on the number of monoploid genomes or chromosome sets involved. Thus, for example, in the case of the fertilization of a pseudogamous tetraploid apomictic plant (with the holoploid chromosome number $2n_a$) *via* a non-reduced pollen tube of a closely related hexaploid individual (with holoploid chromosome number $2n_b$), the sperm cell has $2n_b=6x$, and both the chalazal and the micropylar polar nucleus are $2n_a=4x$. The endosperm then has $2n_b+4n_a$ chromosomes and is therefore hexaplophasic and 14-ploid ($14x$). The endosperm nuclei thus have non-replicated an $^{E}6C$ condition, but the value of their $^{E}1C$ ($=^{E}6C/6$) is clearly non-identical with the 1C-value of the mother plant (the function of the superscript is to indicate this difference). Their genome size can be given as $^{E(2nb+4na)}6C = 14Cx = ..pg$ DNA.

The endosperm of asymmetrically compensating allopolyploids The *R. canina* ($2n=5x$) endosperm forms by the

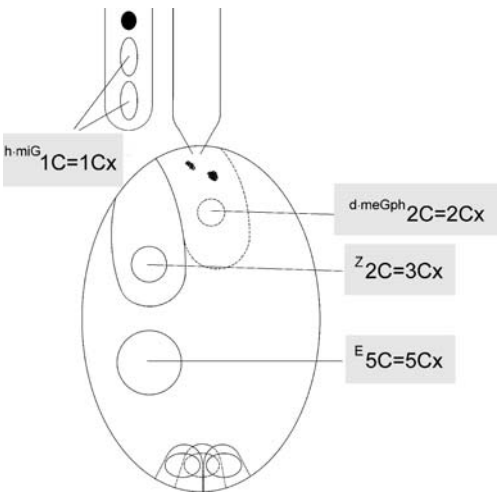

Fig. 6 Nuclear phases, ploidy, and genome sizes in a non-reduced embryo sac of a diploid plant ($2n=2x$) fertilized by a haplophasic pollen tube ($n=x$). The embryo sac before fertilization is diplophasic ($2n$, 2C) and diploid ($2x$, 2Cx). After fertilization, the zygote is diplophasic ($2n$, 2C) and triploid ($3x$, 3Cx); the primary endosperm nucleus is pentaplophasic ($5n$, 5C) and pentaploid ($5x$, 5Cx). For further explanation, see text

fusion of one 1*x* sperm nucleus carrying a copy of the A genome with two tetraploid polar nuclei, each carrying a copy of the A, B, C, and D genomes (Fig. 3). The gametes are reduced in a modified way, with the sperm cell nuclei being "over-reduced" and the polar nuclei "underreduced." Thus, the former can be described as $^{miG}1C$, while each of the two polar nuclei are $^{meG}1C$. In the triple fusion, three "haplophasic" nuclei are combined, so that the non-replicated endosperm nuclei become $^{E}3C$, which are thus in triplophase and nonaploid. The endosperm genome size is expressed by $^{E}3C = {}^{miG}1C_A + 2\left({}^{meG}1C_{ABCD}\right) = 9Cx_{AAABBCCDD} =$ ..pg DNA (Fig. 3). If, for simplicity, it is assumed that each of the four distinct genomes are of the same size (1Cx=0.572 pg DNA), then $^{E}3C$= 5.148 pg DNA.

The treatment of endosperm and embryo zygotes should account for their nuclear phase A zygote is the progenitor cell of the embryo and should therefore be assigned 2*n* as its diplophasic chromosome number and 2C as its (non-replicated) holoploid genome size, irrespective of both ploidy (*x*-level) and nuclear phase (*n*-level) of the contributing gametes. As the endosperm is also a product of fertilization and thus represents a kind of zygote, one may ask if it is possible to assign it the nuclear phase characteristic of embryo sac type (3*n* for the *Polygonum* type, 5*n* for the *Fritillaria* type, etc.), irrespective of both the number of monoploid sets and the nuclear phase of the nuclei. However, there is a fundamental difference between the embryo and the endosperm. The 2*n* level for the zygotic embryo is justified and required to avoid the complication of an unlimited progression in nuclear phase in the context of generative polyploidy. On the other hand, the endosperm is a reproductive dead-end, so it is no longer necessary to guard against this progression. As a result, the nuclear phase of the endosperm can remain the sum of the nuclear phases of the nuclei which fused to generate the primordial endosperm cell.

Following a similar line of reasoning, treating the nuclear phases of the endosperm as a kind of somatic polyploidy is not helpful. In the hypothetical example above where a pseudogamous tetraploid apomictic plant is fertilized by a non-reduced male gamete of a related hexaploid, it would be illogical to propose that the endosperm nuclei have unreplicated a 7C state (the maternal parent contributing two doses of 4*x* and the paternal parent one dose of 6*x*, so that the endosperm is 14*x* and therefore 7C) since the DNA content of a 6*n* constitution is not 7C but 6C. However, it is possible to write $^{E(14x=2na+2nb+2nb)}6C = {}^{Sph(2nb=4x)}7C = 14Cx =$ ..pg DNA, provided it is understood that "Sph" refers to the maternal plant where the endosperm develops. Somatic polyploidy (e.g., endopolyploidy in endosperm nuclei) must not be confused with generative polyploidy and the nuclear phase.

And what about the chromosomes? While the extended C-value terminology describing cell nucleus DNA content avoids ambiguities, the DNA content of individual chromosomes has not yet been considered under similar aspects. Yet, frequently, the size of an individual chromosome is important, such as in DNA-based quantitative karyotyping and genome sequencing projects or in classical aneuploidy studies. The need for a chromosome-based DNA content terminology may not seem obvious, but a logical system would clarify whether a given DNA content refers to either a non-replicated or a replicated chromosome. The use of the C notation combined with a subscript avoids such ambiguities. For instance, $1C_{chr}$=..pg DNA indicates the mean DNA content of a non-replicated chromosome of a complement; $1C_{chr3}$=..pg DNA specifies the DNA content of a non-replicated chromosome 3 (consisting of one double-stranded DNA molecule), while $2C_{chr3}$=..pg DNA specifies the DNA amount in the same chromosome after replication, and $16C_{chr3}$=..pg DNA the DNA content of a polytene chromosome with 16 chromatids.

Overview and conclusions

This interpretation of genome size in eukaryotic organisms rests on the fundamental discoveries of cytogenetics, specifically mitosis and meiosis, the alternation of generations in plants, the fusion of gametes at fertilization, chromatin diminution, and the existence of allosomes. As organisms differ so widely in their life cycles, the genomic constitution of individual life cycle segments can vary according to nuclear phase (reduced or haplophasic and unreduced or diplophasic and dikaryophasic (restricted to fungi)), type of generation (where alternation of generations occurs), sex and mega/microspory, gamete type, spore type, and in angiosperm endosperms according to endosperm type. This variation is overlaid by generative polyploidy, apomixis, chromosome variation, and hybridization. Our solution to account for this variation is to describe the changes in nuclear DNA content during the life cycle by introducing a flexible but consistent terminology.

Building on the initial proposal of Greilhuber et al. (2005), we have developed an indexed C/Cx terminology, which can account for cytogenetically complicated situations, in which the basic C/Cx terminology fails. We anticipate that its main application will be to describe C- and Cx-values in the presence of allosomes. However, some rather less obvious applications can also be envisaged. For example, we have proposed a set of guidelines for specifying C- and Cx-values in hybrids between parents differing in ploidy level and for spontaneous polyploids formed by unreduced gametes. We have also dealt with the

endosperm in hybrids and apomicts. The system copes with the specific problem caused by "asymmetrically compensating allopolyploidy" which occurs, for example, in the *R. canina* species complex. Our attempt to formalize genome size terminology has led to a revision of the conventional number symbols *n* and 2*n* for haplophasic animals and plant sporophytes, tying them to the alternation of nuclear phases. Given the complexity of cytogenetic mechanisms, no global system of terminology can be expected to be simple. Nevertheless, we believe that what we have suggested here is convenient enough for the presentation of DNA content.

The extended terminology on genome size and refined information on C-values may be easily integrated into the existing genome size databases (Gregory et al. 2007). It should be noted that the changes would concern mainly the organisms with allosomes. A possible modification would involve the column for chromosome number, in which the superscript, eventually including chromosome constitution, would be included (Cx-values need not to be given). We recommend that the authors describe their values on genome size in detail, so that it is understood in which life cycle segment the data were obtained.

Acknowledgements We thank Dr. Eva Maria Temsch for the artwork. J. D. is supported by the Czech Republic Ministry of Education Youth and Sports (grant award no. LC06004).

References

Beermann S (1977) The diminution of heterochromatic chromosomal segments in *Cyclops* (*Crustacea*, *Copepoda*). Chromosoma 60:297–344

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Phil Trans Roy Soc London B 274:227–274

Boveri T (1899) Die Entwicklung von *Ascaris megalocephala* mit besonderer Rücksicht auf die Kernverhältnisse. Festschrift für V. von Kupffer

Doležel J, Greilhuber J, Lucretti S, Meister A, Lysák MA, Nardi L, Obermayer R (1998) Plant genome size estimation by flow cytometry: inter-laboratory comparison. Ann Bot 82(Suppl A):17–26

Grant V (1971) Plant speciation. Columbia University Press, New York

Gregory TR, Hebert PDN, Kolasa J (2000) Evolutionary implications of the relationship between genome size and body size in flatworms and copepods. Heredity 84:201–208

Gregory TR, Nicol JA, Tamm H, Kullman B, Kullman K, Leitch IJ, Murray BG, Kapraun DF, Greilhuber J, Bennett MD (2007) Eukaryotic genome size databases. Nucl Acids Res 35:D332–D338 Database issue

Greilhuber J (1988) "Self-tanning"- a new and important source of stoichiometric error in cytophotometric determination of nuclear DNA content in plants. Plant Syst Evol 158:87–96

Greilhuber J, Volleth M, Loidl J (1983) Genome size of man and animals relative to the plant *Allium cepa*. Can J Genet Cytol 25:554–560

Greilhuber J, Doležel J, Lysák M, Bennett MD (2005) The origin, evolution, and proposed stabilization of the terms 'genome size' and 'C-value' to describe nuclear DNA contents. Ann Bot 95:255–260

Greilhuber J, Temsch EM, Loureiro JCM (2007) Nuclear DNA content measurement. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells. Wiley, Weinheim, pp 67–101

John B (1990) Meiosis. Cambridge University Press, Cambridge

Jones RN, Rees H (1982) B chromosomes. Academic Press, London

Klásterská I (1969) Cytology and some chromosome numbers of Czechoslovak roses I. Folia Geobot Phytotax (Praha) 4:175–189

Levan A, Müntzing A (1963) Terminology of chromosome numbers. Portug Acta Biol 7:1–16

Matzk F (2007) Reproduction mode screening. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells. Wiley, Weinheim, pp 131–152

Nogler G (1984) Gametophytic apomixis. In: Johri BM (ed) Embryology of angiosperms. Springer, Berlin, pp 475–518

Rieger R, Michaelis A, Green MM (1991) Glossary of genetics and cytogenetics. Springer, Berlin

Rutishauser A (1969) Embryologie und Fortpflanzungsbiologie der Angiospermen. Springer, Vienna

Singh RJ (2003) Plant cytogenetics, 2nd edn. CRC Press, Boca Raton

Smith-White S (1948) Polarised segregation in the pollen mother cells of a stable triploid. Heredity 2:119–129

Swift HH (1950) The deoxyribose nucleic acid content of animal nuclei. Physiol Zool 23:169–198

Teppner H (1971) Cytosystematik, bimodale Chromosomensätze und permanente Anorthoploidie bei *Onosma* (Boraginaceae). Oesterr Bot Z 119:196–233

Vagera J, Paulíková D, Doležel J (1994) The development of male and female regenerants by in vitro androgenesis in dioecious plant *Melandrium album*. Ann Bot 73:455–459

Widder FJ (1967) Der Generationswechsel der Spermatophyten. Aquilo Ser Bot 6:273–296