Chromosome Research (2008) 16:919–933
DOI: 10.1007/s10577-008-1249-z

FISH mapping and molecular organization of the major repetitive sequences of tomato

Song-Bin Chang[1,2], Tae-Jin Yang[3], Erwin Datema[4], Joke van Vugt[5], Ben Vosman[4], Anja Kuipers[6], Marie Meznikova[7], Dóra Szinay[1], René Klein Lankhorst[8], Evert Jacobsen[6] & Hans de Jong[1,8]*
[1]*Wageningen University, Laboratory of Genetics, 6703 BD Wageningen, The Netherlands;* [2]*Department of Agronomy, National Taiwan University, Taipei 10617, Taiwan;* [3]*Department of Plant Science, College of Agriculture and Life Science, Seoul National University, Seoul 151-921, Republic of Korea;* [4]*Wageningen University and Research Centre, Plant Research International, 6700 AA Wageningen, The Netherlands;* [5]*Departement van Moleculaire biologie, 191 NCMLS Radboud Universiteit, Nijmegen, 6500 HB Nijmegen, The Netherlands;* [6]*Wageningen University, Laboratory of Plant Breeding, 6700 AA Wageningen, The Netherlands;* [7]*Institute of Biophysics, Czech Academy of Sciences, 61265 Brno, Czech Republic;* [8]*Wageningen University, Centre for Biosystems Genomics, 6708 PB Wageningen, The Netherlands; E-mail: hans.dejong@wur.nl*
** Correspondence*

Received 5 May 2008. Received in revised form and accepted for publication by Pat Heslop-Harrison 16 June 2008

Key words: Cot, fluorescence *in-situ* hybridization, heterochromatin, repetitive DNA, repetitive sequences, *Solanum lycopersicum*, tomato

Abstract

This paper presents a bird's-eye view of the major repeats and chromatin types of tomato. Using fluorescence *in-situ* hybridization (FISH) with Cot-1, Cot-10 and Cot-100 DNA as probes we mapped repetitive sequences of different complexity on pachytene complements. Cot-100 was found to cover all heterochromatin regions, and could be used to identify repeat-rich clones in BAC filter hybridization. Next we established the chromosomal locations of the tandem and dispersed repeats with respect to euchromatin, nucleolar organizer regions (NORs), heterochromatin, and centromeres. The tomato genomic repeats TGRII and TGRIII appeared to be major components of the pericentromeres, whereas the newly discovered TGRIV repeat was found mainly in the structural centromeres. The highly methylated NOR of chromosome 2 is rich in $[GACA]_4$, a microsatellite that also forms part of the pericentromeres, together with $[GA]_8$, $[GATA]_4$ and Ty1-*copia*. Based on the morphology of pachytene chromosomes and the distribution of repeats studied so far, we now propose six different chromatin classes for tomato: (1) euchromatin, (2) chromomeres, (3) distal heterochromatin and interstitial heterochromatic knobs, (4) pericentromere heterochromatin, (5) functional centromere heterochromatin and (6) nucleolar organizer region.

Abbreviations

BAC	bacterial artificial chromosome
PBS	phosphate-buffered saline (Na_2HPO_4 and NaH_2PO_4)
BSA	bovine serum albumin
CONSED	software for editing sequence assemblies
Cot	(= $C_0 \times t$) C_0 is the initial concentration of single stranded DNA in mol/L and t is the reannealing time in seconds. Cot is a measure of DNA complexity
DAPI	4′,6-diamidino-2-phenylindole
FITC	fluorescein isothiocyanate
GYPSODE	GYPSY-like LTR retroelement
HB50	hybridization buffer containing 50% deionized formamide
LTR	long terminal repeat
NEB	New England Biolabs Restriction Enzyme buffer
NOR	nucleolar organizer region

SSC	standard saline citrate ($1 \times SSC = 0.15$ M NaCl–0.015 M Na_3-citrate)
PHRED	software for reading DNA sequencing trace files
PHRAP	software for assembling shotgun DNA sequence data
PCRT	pericentromere repeat
TGR	tomato genomic repeat

Introduction

Tomato (*Solanum lycopersicum*) is genetically one of the best studied Solanaceous species with a fully saturated genetic map of 1270 cM based on segregating *S. lycopersicum* × *S. pennellii*/*pimpinellifolium* F2 populations (http://www.sgn.cornell.edu/; Tanksley *et al.* 1992). The species is also recognized for its eminent pachytene morphology (Ramanna & Prakken 1967, Khush & Rick 1968, Zhong *et al.* 1998) and well-differentiated synaptonemal complexes (Sherman & Stack 1995, Peterson *et al.* 1999). Its genome size amounts to 950 Mb (Arumuganathan & Earle 1991) and tomato thus ranks among plant species with middle-sized genomes. About 25% of the genome consists of single copy sequences which reside mostly in the euchromatin fraction of the chromosomes (Peterson *et al.* 1996, Van der Hoeven *et al.* 2002). Genomics has long underestimated the significance of heterochromatin because of the low proportion of coding sequences. However, recent DNA sequence analyses of centromeres and flanking pericentromeres in various plants (e.g. Budiman *et al.* 2004, Copenhaver *et al.* 1999, Nagaki *et al.* 2004) demonstrated that transcribed genes exist in these regions. The repeat fraction of the tomato genome can be divided into tandem repeat families like the 45S and 5S rDNAs, the [TTTAGGG] telomere repeat, and several species-specific repeats that are abundant in the centromeres, telomeres and other heterochromatic regions, and other dispersed repeats such as retrotransposons and transposable elements that occur both in heterochromatin and in euchromatin.

The precise way in which the repetitive sequence-rich part of the tomato genome is organized is still to be elucidated. This is partly due to difficulties in reconstituting long stretches of repetitive sequences owing to their unstable maintenance in cloning vectors (Song *et al.* 2001), but also because of the problematic assemblage of repeat arrays into longer genome scaffolds (Tilford *et al.* 2001). Although current sequencing technologies can produce sequence reads of more than a thousand bases, they are still not able to cover reasonable stretches of repeats in heterochromatin domains.

Typical of the pachytene chromosomes of tomato and other *Solanum* species are the clearly distinguishable centromeres, heterochromatin and euchromatin domains, which make them outstanding for the characterization of different chromatin classes with respect to the presence and organization of tandem and dispersed repeats. The first repetitive sequence described in tomato was the highly repetitive species-specific 452 bp *Hind*III DNA motif, THG2 (Zabel *et al.* 1985), which was identified as a member of a larger repeat complex dispersed throughout the genome. FISH showed that this repeat is present in the heterochromatin of all pericentromeres, but not in the nucleolar organizer regions (NORs) (Zhong *et al.* 1996). Three major genomic repeats, TGRI, TGRII and TGRIII, were described by Schweizer *et al.* (1988), Ganal *et al.* (1988, 1991) and Lapitan *et al.* (1989, 1991). It was shown that TGRI is located in the distal regions, and that TGRII and TGRIII are located in the pericentromere heterochromatin regions of the chromosomes. Studies of the [TTTAGGG] telomere repeat and the subtelomeric TGRI repeat (Lapitan *et al.* 1989, 1991) in pachytene and in extended DNA fibre FISH (Zhong *et al.* 1996, 1998) revealed accurate distributions and size estimates of these repeats in the chromosome ends and in the long-arms interstitial knobs. Initial low resolution *in-situ* hybridizations at metaphase complements also identified the TGRII repeat family on nearly all chromosomes except for chromosome 2, whereas TGRIII occurred in the centromere regions of only nine of the chromosomes (Ganal *et al.* 1988). The ribosomal DNA repeats, including the 45S rDNA tandem repeat family, have been identified at a single locus at the end of the short arm of chromosome 2 in the nucleolar organizer region with minor loci on the 2L, 6S, 9S and 11S arms. The 5S ribosomal RNA gene cluster was found in the short-arm pericentromere of chromosome 1 close to the centromere (Lapitan *et al.* 1991, Xu & Earle 1994, 1996a,b). The $[GA]_n$ motif is the most abundant of the 17 microsatellites tested for the tomato genome (Broun & Tanksley 1996). $[GACA]_4$ and $[GATA]_4$ were among the less abundant microsatellites in the tomato genome and were used by Vosman *et al.* (1992) to discriminate 15 cultivars on the basis of unique fingerprints. Vosman & Arens (1997) showed that these microsatellite motifs are frequently associated with one other repetitive element, which they termed U30 (cf. Ganal *et al*. 1988, Yang *et al*. 2005, Wang *et al*. 2006).

In this paper we present an updated view of the molecular organization of the major repetitive DNA sequences and chromatin types in tomato (*S. lycopersicum* cv. Cherry). We first present a simplified technique for isolating different Cot-DNA fractions, based on typical DNA renaturation of sheared heat-denatured genomic DNA and show their FISH signals on mitotic and pachytene cell complements, and the hybridization intensities on a BAC filter of the Heinz 1706 *Hind*III BAC library. Secondly, we have established the FISH distribution of various major repeats in tomato, including U30, as component of TGRII, TGRIII and a newly characterized centromere-specific retrotransposon-like repeat, termed TGRIV. Thirdly, we have studied the chromosomal distribution of the dispersed repeat Ty1-*copia* and the microsatellites $[GA]_8$, $[GACA]_4$, $[CAA]_n$ and $[GATA]_4$. The data of all repeats in this study were used for generating an overview showing the chromosomal positions of these repeats with respect to chromatin, heterochromatin regions, NORs and centromeres.

Material and methods

Tomato *Solanum lycopersicum* cv. VFNT Cherry (LA1221, 2n=2×=24) was used in all our FISH experiments. For the spread preparations of mitotic root tip and meiotic pachytene complements we followed the protocols of Zhong *et al*. (1996) and Budiman *et al*. (2004).

Isolation of genomic DNA and production of Cot fractions

Cot fractions of tomato genomic DNA were prepared according to Zwick *et al*. (1997) with some modifications. Total genomic DNA was isolated according to the CTAB method and sonicated into a fragment size of about 1 kb. We denatured 0.5 μg/μl sonicated DNA in 0.3 M NaCl at 95°C for 10 min, and then transferred it to a water bath at 62.4°C (Peterson *et al*. 1998) for single-strand DNA reassociation. Reassociation time according to Chang (2004) for Cot-1 is 11 min 18 s, for Cot-10 is 1 h 53 s, and for Cot-100 is 18 h 50 min. The remaining single-strand DNA was digested with 1 U/μg S1-endonuclease (Fermentas) for 90 min at 37°C and the double-strand DNA fraction was finally purified by ethanol precipitation.

Tandem and dispersed repeats

The list of all repeats is given in Table 1. As probe we used the 327 bp DNA fragment of plasmid pWVA322, containing the U30 element (Vosman & Arens 1997). The probe showed a FISH distribution that was identical to that of the TGRII repeat (data not shown), but with a far brighter fluorescence intensity, and we therefore considered this repeat as a component of TGRII (hereafter referred to it as TGRII*). The synthetic oligonucleotides $[GA]_8$, $[GACA]_4$ and $[GATA]_4$ (Vosman & Arens 1997)

Table 1. Overview of the tandem and dispersed tomato repeats used in this study

Repeats	Probe (origin)	Unit length	Copy number	Reference (tomato)	GenBank accession no.
45S rDNA	pW71	9.1 kb	2300	Ganal *et al*. (1988)	
5S rDNA	pW74	400 bp	1000	Lapitan *et al*. (1991)	
Telomere[TTTAGGG]	pAtT4	7 bp	N.D.	Richards & Ausubel (1988)	
TGRI		162 bp	77000	Schweizer *et al*. (1988)	X87233
TGRII*	U30 pWVA332	780 bp	4200	Vosman & Arens (1997) Yang *et al*. (2005) Wang *et al*. (2006)	X90770
TGRIII	PCR product	509 kb	2100	Ganal *et al*. (1988)	AY880063
TGRIV	(BAC 57J04)	7 kb	N.D.	This study	AC211033.1
Ty1-copia	PCR product	N.A.	N.D.	Kuiper *et al*. (1998)	
$[GA]_8$		2 bp	N.D.	Broun & Tanksley (1996)	
$[GACA]_4$		4 bp	N.D.	Vosman & Arens (1997)	
$[GATA]_4$		4 bp	N.D.	Vosman & Arens (1997)	

were end-labelled with biotin by terminal transferase (Pharmacia Biotech). The 266 bp conserved domain of the reverse transcriptase gene of the Ty1-*copia* retrotransposons was amplified from tomato cv. C31 genomic DNA according to Flavell *et al*. (1992), and cloned as described by Kuipers *et al*. (1998). The consensus sequence of the TGRIV repeat was calculated using the 'cons' program of the 'EMBOSS' package (version 5.0.0.) after clustal W (version 1.83) alignment of three different repeat units obtained from BAC 57J04.

The Gypsode1 part and the presumed LTR part of TGRIV were amplified form BAC 57J04 using the following primer sets:

Gypsode part

Gypsode-forward1: CATCTCCCATTTCCAAGTCG T_m=60.5°C
Gypsode-reversed1: TGTTGGCTGCTTGTTATGC T_m=62.1°C
Gypsode-forward2: CAAGCAGCCAACCAATC CAT T_m=63.3°C
Gypsode-reversed2: TGCACCACATCTATAAA CAGGAAG T_m=60.4°C

LTR part

Specific-forward3: AACTGCGCGTCTACACCTG T_m=60.1°C
Specific-reversed3: GGCGTGACCAGTAGCT TAGG T_m=59.9°C
Specific-forward4: GGGGGTCAGTACGAAC AATG T_m=60.2°C
Specific-reversed4: TGGAGAAATCCACTCC ATGC T_m=61.0°C

Isolation of repeat and BAC DNA

DNA from BACs and repeats in plasmids were isolated by plasmid isolation kit using the manufacturer's protocol (Roche). The Ty1-*copia* was PCR-amplified with the specific forward primer 5′-ACNGCNTTYYTNCAYGG-3′ and reverse primer 5′-ARCATRTCRTCNACRTA-3′.

BAC filter hybridization

BAC filter hybridization was based on the Southern blotting protocol described by Van Daelen & Zabel (1994) with some modification using non-radioactive dig-11-dUTP labelling (Roche). One BAC filter of the tomato Heinz 1706 BAC library (LeHba-A, http://www.genome.arizona.edu) representing 2× tomato genome equivalents was hybridized with tomato Cot-100 DNA, TGRII, and TGRIII repeats. We spotted two dots of unlabelled and denatured Cot-100 DNA at margins as 100% reference and randomly picked three points at margins as 0% reference. We repeated the hybridization in two laboratories and both gave similar results. In addition, there are the dual spots for the same BAC clones in the arrangement of the 4×4 matrix (see the inset of Figure 2). Only the right orientation of the dual spots confirmed the true hybridization of the target BAC clones. We also found that the difference of the intensity between dual spots was less than 3% in those randomly selected 800 clones.

Detection of 5-met-Cyt DNA

Cell spread preparations were incubated at 60°C for 30 min before RNaseA treatment (100 μl per slide of a 10 mg/ml, 1:100 in 2× SSC solution) at 37°C for 1 h. Three washing steps followed in 2× SSC for 5 min, 10 mM HCl for 2 min, and pepsin (200 μl per slide of a 1:100 pepsin solution in 10 mM HCl) at 37°C for 7 min. After 2×5 min PBS steps the preparations were postfixed in 1% (para)formaldehyde in PBS for 10 min, followed by two more washing steps in PBS. After dehydrating in an ethanol series (70%–90%–98%, 3 min each), and HB50 treatment (20 μl), slides were denatured at 80°C for 2 min, then cooled in ice-cold 2× SSC, 2×5 min. Slides were incubated in 1% BSA in PBS at 3°C for 30 min and washed in PBS 3×5 min before incubation in purified Anti-5-methyl-Cytosine Polyclonal Antibody (Megabase Research Products, USA) in 1:50 PBS (50 μL per slide) at 37°C for 2 h (or at 4°C overnight). Then washing steps in PBS 3×5 min, 4× SSC+0.05% Tween for 5 min and treatment with Alexafluor 488-conjugated goat anti-rabbit (Invitrogen) in 1% BSA in PBS (1:1000, 100 μl per slide) at 37°C for 40 min, followed by 3×4T (4× SSC+0.05% Tween 20). The chromosomes were counterstained with 1 μg/μl DAPI in Vectashield antifading (Vector Laboratories). For Southern hybridization, total genomic tomato DNA was digested with 3 μl *Mbo*I (NEB, 5 U/μl) and *Sau*3AI (NEB 4 U/μl) at 37°C for 3 h. Repetitive DNA sequences were labelled with

[^{32}P]dATP (ICN Biomedicals) and detected with the use of a DecaLabel DNA Labeling Kit (Fermentas). Hybridization took place at 42°C overnight.

Fluorescence in-situ hybridization

Each repeat DNA and Cot fraction DNA sample (1–2 μg) was labelled with either biotin-16-dUTP or digoxigenin-11-dUTP by nick translation using the manufacturer's protocol (Roche) and FISH was performed according to Zhong *et al*. (1996). The post-hybridization wash for microsatellite repeats was modified and followed the protocol of Cuadrado & Schwarzacher (1998). The slides were washed three times for 30 min in 6× SSC at 20°C, followed by a stringent wash in 6× SSC at the respective duplex stability temperature (T_m) – 5°C for 2 min. The T_m value was calculated for each probe following Wallace *et al*. (1981). The detection and amplification of microsatellite signals was also modified using 6T (6× SSC+0.05% Tween 20) instead of 4T. Chromosomes were counterstained with 5 μg/ml DAPI mounted in Vectashield anti-fade (Vector Laboratories). Slides were examined under a Zeiss Axioplan 2 photomicroscope equipped with epifluorescence illumination and filter sets for DAPI, FITC and Texas-Red fluorescence. Selected images were captured either by 35 mm photocamera using ISO 100 Kodak films (FISH of microsatellites), or by a Photometrics Sensys 1305×1024 pixel CCD camera and processed with Genus Image Analysis Workstation software (Applied Imaging Corporation). We sharpened the DAPI images with a 7×7 Hi-Gauss high pass spatial filter to accentuate minor details and heterochromatin differentiation of the chromosomes. Fluorescence images were displayed in gray value (DAPI) or pseudo-coloured and processed in the multichannel mode of Adobe Photoshop. We also used this program to further improve brightness, contrast and colour saturation. Selected images of pachytene chromosomes were selected and straightened in ImageJ 1.37V (http://rsb.info.nih.gov/ij) using the straighten plugin (Kocsis *et al*. 1991).

DNA sequencing and analysis

The TGRIV clones were sequenced using BigDye terminators chemistry v3.0 (ABI) with m13fow and m13rev primers. The sequences were analysed using ABI3730 automatic DNA sequencer (ABI). Sequences were base called with PHRED. Bad-quality and vector sequences were filtered with CROSS_MATCH. High-quality sequences (over PHRED 30) were assembled and edited with PHRAP and CONSED (Ewing & Green 1998, Ewing *et al*. 1998). The TGRIV sequences were compared against the genome sequence database (http://www.ncbi.nlm.nih.gov/) and newly characterized tomato transposons (Yang *et al*. 2005) using BLAST and PipMaker (Schwartz *et al*. 2000, http://bio.cse.psu.edu/cgi-bin/pipmaker?advanced).

Results

We isolated Cot-1, Cot-10 and Cot-100 fractions as representative pools of highly repetitive (HR), high+moderately repetitive (MR) and high+moderate+low copy (LC) components of the tomato genome (Peterson *et al*. 2002). Next, we established the chromosomal locations of these repeat fractions by FISH on mitotic and pachytene complements (Figure 1a–e). The Cot-1 FISH signals were most prominent in the macrosatellite and the nucleolar organizer of chromosome 2 but also displayed weaker foci in some distal and pericentromere regions (Figure 1a). This Cot fraction is therefore supposed to contain mainly the 45S rDNA and telomere repeats. The Cot-10 fraction, which contains the repeats of the Cot-1DNA pool as well, also showed fluorescing foci on most chromosome ends and on pericentromere regions (Figure 1b). FISH with Cot-100 DNA resulted in bright signals over all chromosomes, except for some interstitial and distal regions (Figure 1c). The distribution of this fraction in heterochromatin and in euchromatin was also studied using FISH on the far longer and better-differentiated pachytene complement. Hybridization with Cot-100 showed that all pooled repeats in this fraction are located on the distal and pericentromere heterochromatin regions of all chromosomes and to a lesser extent on some minor regions in the euchromatin areas (Figure 1d,e).

We also tested the abundance of repetitive DNA in tomato BACs by hybridizing Cot-100 DNA to one of the tomato genomic BAC library filters. Figure 2 shows the hybridization of Cot-100 on filter A of the Heinz 1706 tomato BAC library, of which the

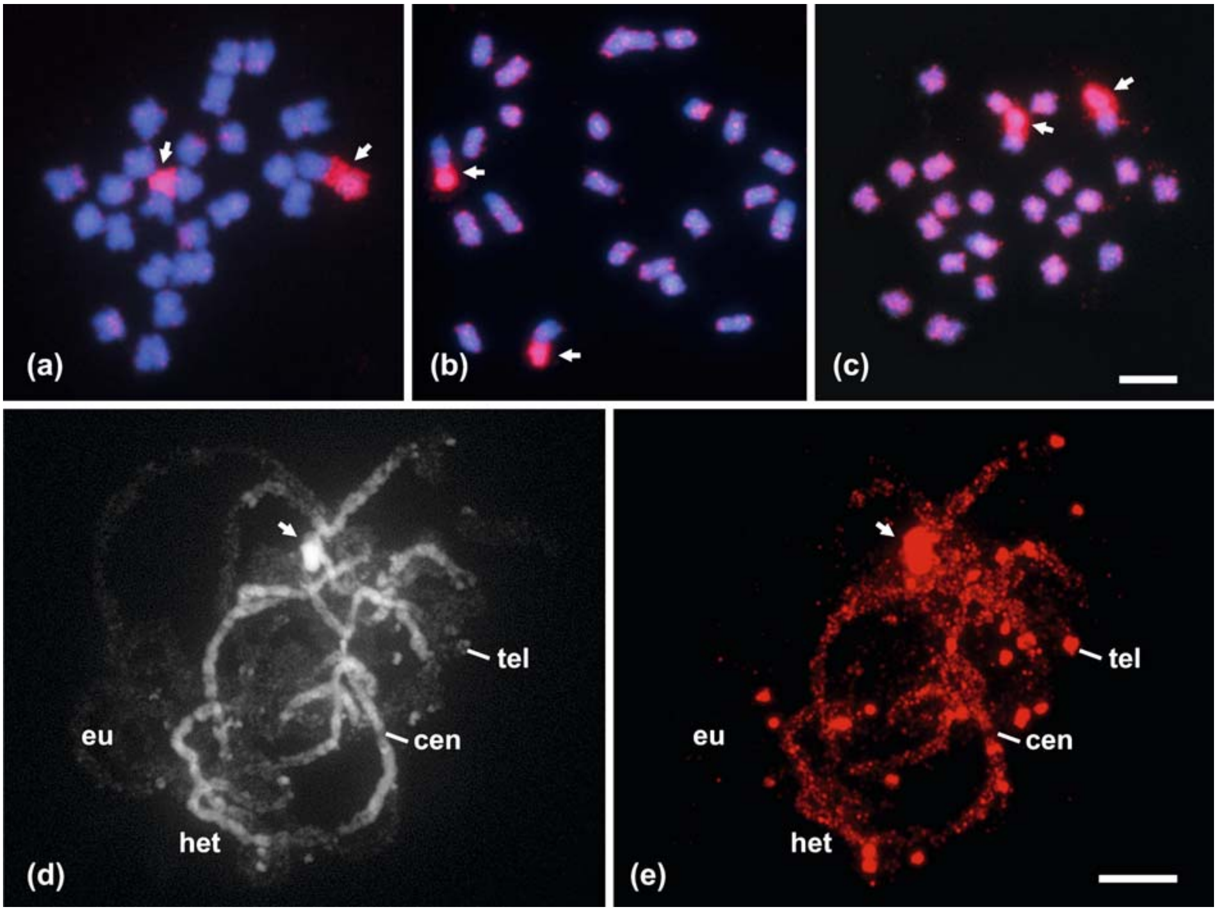


Figure 1. FISH analysis of different Cot DNA fractions on metaphase chromosomes: (**a**) Cot-1, (**b**) Cot-10, (**c**) Cot-100; and pachytene complement, (**d**) DAPI, (**e**) Cot-100. The arrows indicates the positions of the nucleolar organizer region of chromosome 2. eu=euchromatin; het=heterochromatin; cen=centromere; tel=telomere. Bar represents 5 μm.

small 4×4 matrix contains twin sets of eight different BAC clones (inset figure). We distinguished four grey levels for the hybridization signals of BAC clones as a criterion for the relative repetitive sequence content of the corresponding BAC insert. 'Dark', 'medium', 'weak' and 'null' were used as classes to define relative intensities (compared with the surrounding background) of 66–100%, 33–66%, 5–33% and <5%, respectively (Table 2). Intensities of less than 5% were considered negligible. We randomly selected 100 matrices, involving a total of 800 BAC clones, to estimate their repetitive sequence content. The 'dark' class represented 10% of the library; 'medium' represented 7.75% and 'weak' represented about 12.5%. In total, BAC clones hybridizing to the Cot-100 pool of repetitive DNA represented 30.25% of the 800 analysed BACs.

Next, we carried out a number of FISH analyses using a subset of defined tomato repeats on mitotic and pachytene cell spread preparations, see Figures 3 and 4. The fluorescence pattern of the pachytene chromosomes revealed a clear heterochromatin pattern as described in previous studies (Ramanna & Prakken 1967, Zhong *et al*. 1998, de Jong *et al*. 1999, 2000). Bright DAPI fluorescent parts were considered as heterochromatin, and weaker fluorescent parts were considered to be euchromatin. The obtained fluorescence signals of TGRII, TGRIII, TGRIV, Ty1-*copia*-like retrotransposon, and the microsatellites $[GA]_8$, $[GACA]_4$ and $[GATA]_4$ were catalogued in terms of presence/absence in (i) the distal heterochromatic knob of the short arm, (ii) the short-arm euchromatin, (iii) the short-arm pericentromere heterochromatin, (iv) the weakly fluorescing

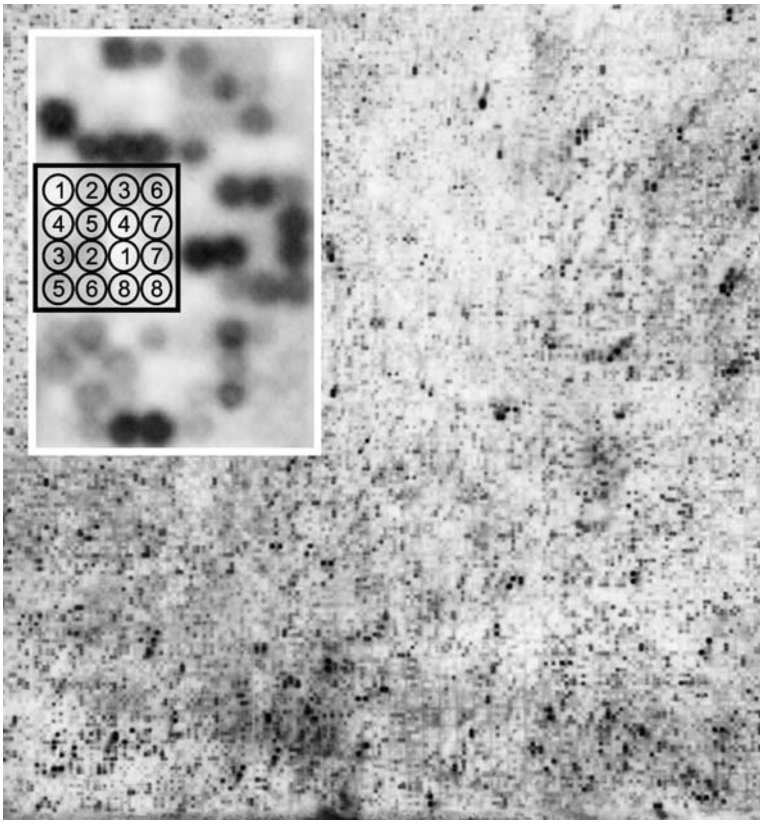


Figure 2. Filter hybridization of Cot-100 on the tomato Heinz 1706 BAC filters. Hybridization with labelled Cot-100. The inset shows an enlargement of six 4×4 matrices, each containing 8 BAC clones spotted *in duplo*. Identical figures in the matrix refer to the same clone.

centromere, (v) the long-arm pericentromere heterochromatin, (vi) the long-arm euchromatin, (vii) the long-arm interstitial heterochromatic knob, (viii) the distal heterochromatic knob, and (ix) the nucleolar organizer region. Figure 4a displays the DAPI patterns of all 12 chromosomes. Representative chromosomes from different cell complements were selected, straightened, and arranged with their centromeres on a horizontal line. As most of the borders between euchromatin and pericentromere heterochromatin areas contain several smaller heterochromatin blocks, we used the better-defined heterochromatin patterns in the carmine-stained pachytene cell complements of Ramanna & Prakken (1967) to measure heterochromatin size. The values for euchromatin and pericentromere heterochromatin were used for estimating their DNA content, using the 950 Mb for the whole tomato genome (Arumuganathan & Earle 1991) and the heterochromatin–euchromatin density factor of 5.8 (Peterson *et al*. 1995). The results are described in Table 3.

Figure 4b gives an overview of the distribution and relative intensity of the FISH signals for each repeat in the different chromosomal regions. For comparison, the hybridization patterns of Cot-1, Cot-10, Cot-100 and of the previously published tandem

Table 2. Scoring of peroxidase deposit intensities on the Heinz 1706 BAC filter A hybridized with Cot-100 in a random selection of 100 matrices (800 BAC clones)

Intensity	Dark (66–100%)	Medium (33–66%)	Weak (5–33%)	Sum of hits (5–100%)	Null (<5%)	Total
Clones	80	62	100	242	558	800
%	10	7.75	12.5	30.25	69.75	100

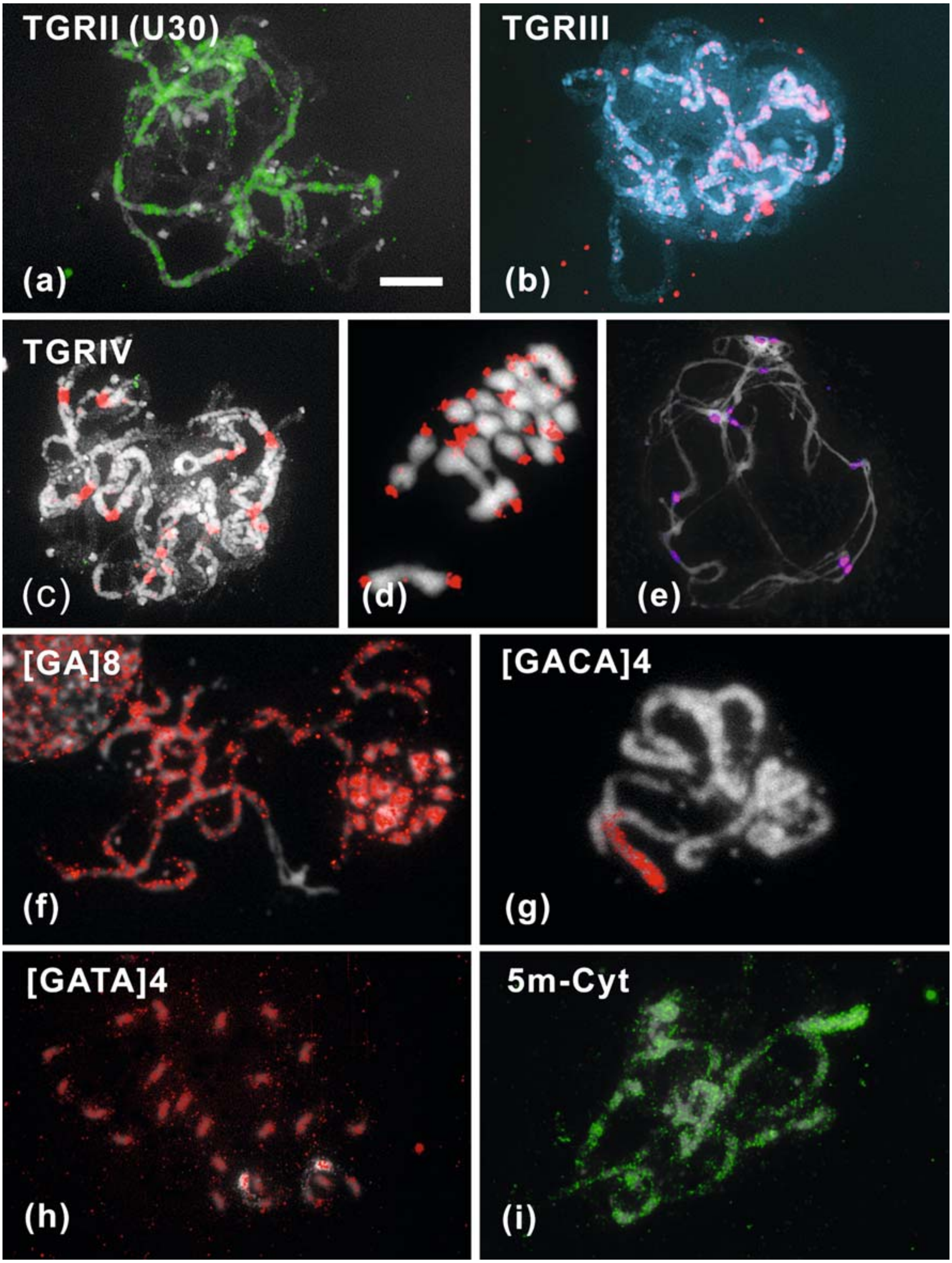

TGRII (U30)
(a)
TGRIII
(b)
TGRIV
(c)
(d)
(e)
[GA]8
(f)
[GACA]4
(g)
[GATA]4
(h)
5m-Cyt
(i)

Figure 3. FISH of repetitive sequences on chromosome complements of tomato. (**a**) Pachytene complement showing TGRII in the pericentromere of all bivalents. (**b**) TGRIII which also occurs in the pericentromere, but is weaker and less abundant as TGRII. (**c**) Pachytene complement showing the TGRIV in the centromeres of all bivalents and some weak signals in the pericentromeres of some chromosomes. (**d**) FISH of TGRIV on a metaphase I complement. (**e**) Two-colour FISH of the GYPSODE1-like part (purple) and LTR-part element (blue) of the TGRIV-element on a pachytene complement. (**f**) FISH of the $[GA]_8$ microsatellite. (**g**) $[GACA]_4$. (**h**) $[GATA]_4$ on a mitotic metaphase complement. (**i**) Immunodetection of methylated DNA detected with anti-5-methylcytosine AB-FITC.

repeats 45S rDNA, 5S rDNA, the telomere repeat TTTAGGG and TGRI (Zhong *et al*. 1998), and the [CAA] microsatellite (Yang *et al*. 2005) are also included in this figure. The TGRII repeat (and its U30 component, Vosman & Arens 1997) was detected in all heterochromatin regions except in the NOR region (Figure 3a). This repetitive element shares 82% sequence identity with a novel putative Ty3-*Gypsy* type pericentromere retrotransposon of tomato and contains the LTR-region of the previously described PCRT1a element (Yang *et al*. 2005), which is similar to the *Jinling* element (Wang *et al*. 2006). The TGRIII repeat (Figure 3b) and the tomato specific Ty1-*copia*-like retrotransposon (data not shown) were found in the pericentromere heterochromatin. The 509 bp TGRIII element is similar to a part of the LTR sequence of the PCRT2 element, which is another novel tomato Ty3-*Gypsy* type pericentromere retrotransposon, showing homology to Athila in *Arabidopsis* (Yang *et al*. 2005). Like TGRII, TGRIII was not detected in the NOR regions and, overall, TGRIII appeared to be less abundant than TGRII. In addition, neither TGRII nor TGRIII was detected in the tomato centromeres.

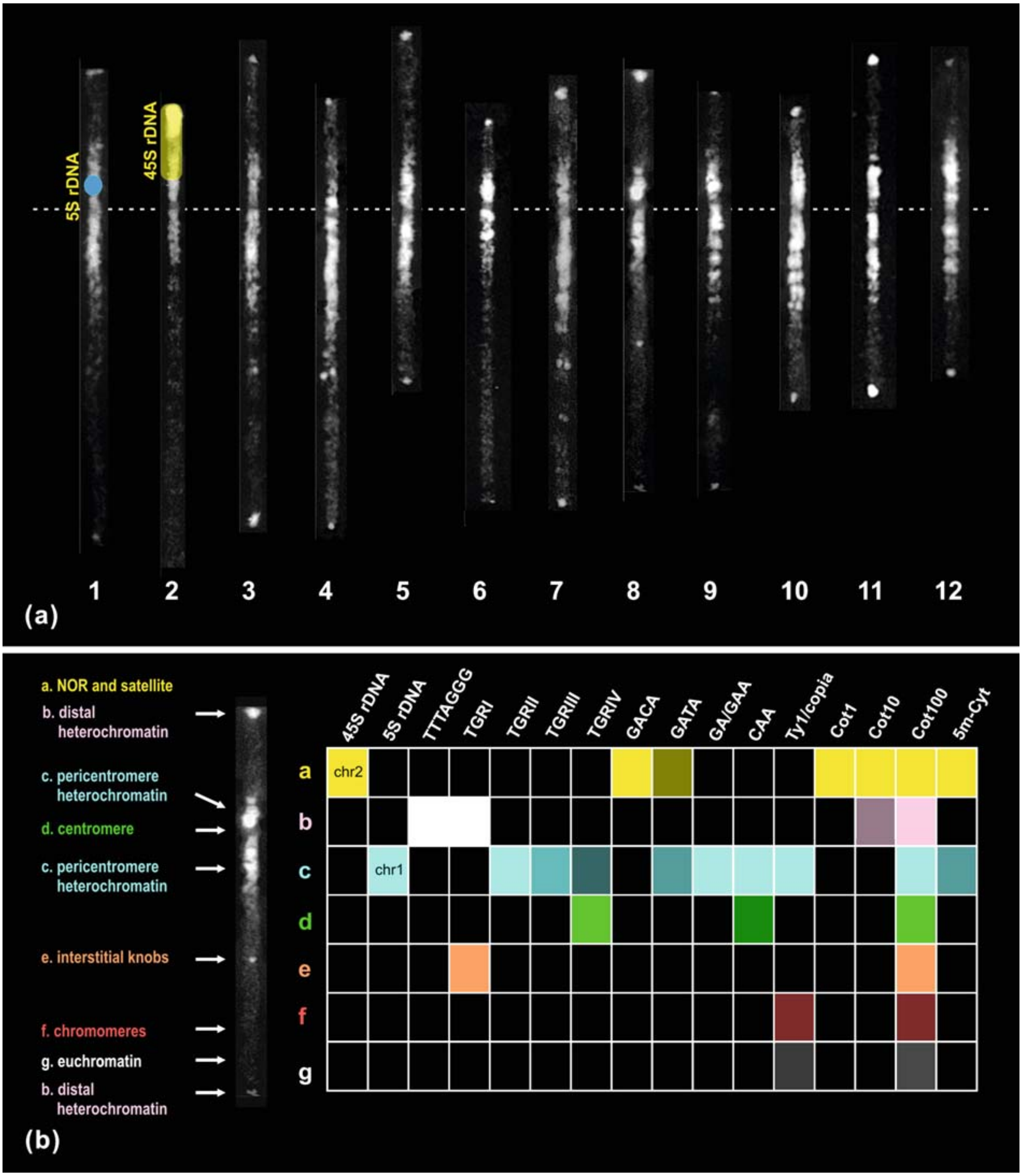
The TGRIV-element was found as a component of the structural centromere itself (Figure 3c,d). TGRIV is a repeat that was identified in the course of the Tomato Genome Sequencing project in the draft sequence of the BAC clone 57J07 (AC211033.1). The TGRIV element shows partial homology to tomato retrotransposons with the closest similarity to GYPSODE1 (Jurka 2005). When FISH experiments were carried out with the entire BAC 57J04 as a probe (results not shown), strong signals on all tomato centromeres were observed as well as weak signals in some pericentromeres. Sequence analyses showed that this BAC for the larger part is composed of retrotransposon-like sequences of the TGRIV-type and of homologues of the pericentromere retrotransposons PCRT–1, –2, –3, and –4 (Yang *et al*. 2005). Next to a part with high similarity to the GYPSODE1 element (nt 2401–7000), the TGRIV element also contains a unique part at its 5′ end (nt 1–2400) that shows no homology with any entry in the database but which could be an LTR of an ancestral TGRIV-type retrotransposon. When both these parts were labelled separately and used in a two-colour FISH assay it was confirmed that both the GYPSODE1-like part of this element and the presumed LTR sequence co-localize and occur in all tomato centromeres; see Figure 3e.

The FISH experiments with the microsatellites were performed both on metaphase and on pachytene complements. The $[GA]_8$ and $[GATA]_4$ signals overlapped entirely in the pericentromere heterochromatin (Figure 3f,h), although $[GA]_8$ was far brighter than $[GATA]_4$. Strikingly, hardly any $[GA]_8$ signals were observed in the NOR region, while the $[GATA]_4$ probe partly painted this region. $[GACA]_4$ clustered almost entirely at the NOR regions and only very few foci were found in heterochromatin regions (Figure 3g).

We also analysed the pattern of methylated DNA in the heterochromatin of pachytene complements. As shown in Figure 3i, we observed very strong fluorescence of the 5-methylated-Cyt DNA in the satellites and NORs of chromosome 2, bright fluorescence in the pericentromeres, and less or no fluorescence in the centromeres and the euchromatin regions. Southern hybridizations of *Mbo*I- and *Sau*3AI-digested genomic DNA probed with 45S rDNA and TGRII were in agreement with the microscopic observation that both the 45S rDNA NOR and pericentromere were highly methylated (results not shown).

Discussion

The FISH of Cot fractions and of various tandem and dispersed repeats in tomato have shown that most of the repeats are confined to the clearly distinguishable heterochromatin blocks at the telomeres, in the pericentromeres and in the large nucleolar organizer region (NOR). None of the repeats is diagnostic for the identification of individual chromosomes in the



	45S rDNA	5S rDNA	TTTAGGG	TGRI	TGRII	TGRIII	TGRIV	GACA	GATA	GA/GAA	CAA	Ty1/copia	Cot1	Cot10	Cot100	5m-Cyt
a	chr2							🟨	🟫				🟨	🟨	🟨	🟨
b			⬜	⬜										🟪	🟪	
c		chr1			🟦	🟦	🟦		🟦	🟦	🟦	🟦			🟦	🟦
d							🟩				🟩				🟩	
e				🟧											🟧	
f												🟥			🟥	
g												▨			▨	

Figure 4. (**a**) DAPI images of straightened pachytene complements, ordered according to the tomato chromosome nomenclature (Ramanna & Prakken 1967) with the centromeres positioned on the dotted line. Chromosomes were taken from different cell complements and their lengths adapted fit the karyotype representation. The chromocentres in the middle of the long arms are polymorphic. (**b**) Overview of the major repeat and chromatin classes, along with the Cot hybridizations and methylated DNA on the chromosomes of tomato (*S. lycopersicum*). The chromatin classes (a–g) as described in the text are displayed in different colours. Chromosome 2 is the only chromosome with a large satellite and NOR and a chromatin class unique in its methylation status and microsatellites. Chromosome 1 has the 5S rDNA locus. All other chromosomes are essentially the same for the distribution of the repeat classes. The cells in the table show a solid colour if the repeat has a strong fluorescence signal and a textured pattern if the fluorescence in weak or interrupted.

Table 3. DNA content estimates (in Mb) of euchromatin and heterochromatin regions for each tomato pachytene chromosome[a]

	Chromosome												Total
	1	2	3	4	5	6	7	8	9	10	11	12	
Short-arm euchromatin	5.7	0.0	9.2	8.4	9.7	4.1	6.2	6.2	8.1	5.5	10.0	8.0	81.2
Short-arm heterochromatin	29.3	23.5	13.3	9.1	25.1	9.6	24.0	14.9	17.6	16.0	11.7	25.6	219.7
Long-arm heterochromatin	33.6	27.2	35.2	44.8	32.0	18.7	29.9	25.6	38.4	50.1	32.0	29.9	397.3
Long-arm euchromatin	44.5	38.8	30.4	19.5	9.7	26.9	18.9	19.3	15.5	10.2	9.7	8.4	251.8
Total amount	113.1	89.5	88.2	81.7	76.4	59.2	79.0	66.0	79.6	81.9	63.5	71.8	950.0
% Heterochromatin	55.6	56.6	55.0	65.9	74.7	47.7	68.2	61.4	70.3	80.8	68.9	77.2	64.9

[a] Average length values for euchromatin and heterochromatin patterns in all pachytene chromosome of tomato were taken from Ramanna & Prakken (1967) and converted into DNA content using 950 Mb as the content of the whole tomato genome (Arumuganathan & Earle 1991) and the heterochromatin–euchromatin density factor of 5.89 (Peterson *et al.* 1995). The long and short arm euchromatin values include the distal heterochromatin blocks.

cell complement, except for the 5S rDNA that maps exclusively on the short-arm pericentromere heterochromatin of chromosome 1, and the 45S rDNA and the [GACA]$_4$ microsatellite, which paint exclusively the satellite and NOR of chromosome 2. In addition, Zhong *et al.* (1998) showed that TGRI occurs in most distal heterochromatic blocks and the small interstitial long-arm heterochromatic knobs. The repeat in BAC 57J04 that we refer to as tomato genomic repeat TGRIV hybridizes to all structural centromere regions and in small amounts to the pericentromeres. It is not clear yet whether this repeat is associated with the centromere-specific histone H3 variant CenH3 (Jiang *et al.* 2003).

The use of Cot hybridization of chromosomes and BACs provides most useful tools in molecular cytogenetics, chromosome walking and studies on the physical organization of the genome. First, painting the chromosomes with the Cot-100 in tomato reveals the distribution of all repeats along the tomato chromosomes. Zamir & Tanksley (1988) and Ganal *et al.* (1988) have already shown that about 70% of the tomato genome is composed of single-copy DNA sequences, which agrees with our filter hybridization data of about 30% of BAC clones with clear Cot-100 DNA hybridization. These consistent results, together with the FISH analysis, imply that most tomato repeats are clustered in the large pericentromere and NOR heterochromatin. However, a recent study on the number of genes in tomato indicated that the genome encodes approximately 35 000 genes, which together cover less than 25% of the nuclear DNA (Van der Hoeven *et al.* 2002). It is likely that the discrepancy between both estimates can be explained by the numerous low copy sequences, introns, and intergenic spacers, and different transposable element classes that have not yet been detected by FISH or are below detection of FISH and Southern hybridization. A second major application of Cot-100 hybridization on BAC filters is the identification of BAC clones containing detectable proportions of low, moderate or high repetitive DNA. A short study describing the Cot-100 values of all 33 000 clones from the *Hind*III Heinz 1706 BAC library is now in preparation and will soon be made available on the internet for public consultation. Providing this supplementary information will aid the ongoing Tomato Genome Sequencing program, which aims at sequencing only the euchromatin of tomato.

Our study has revealed some controversies in the FISH positions of Cot fractions on the chromosomes. Peterson *et al.* (1998) established the Cot value of a repeat of more than 1000 copies at 0.819; that of more than ten copies is 81.9 and the value for single-copy sequences is more than 819. The major repeats contain more than 1000 copies (Table 1) and therefore the Cot-1 fraction was expected to cover the 45S rDNA, 5S rDNA, and the tandem repeats TGRI, TGRII and TGRIII. However, our FISH study of Cot-1 showed fluorescence signals only on the NOR and chromosome ends, which are the domains of 45S rDNA, TGRI and telomeric repeats (Figure 4b). The Cot-100 fraction contains high-, medium- and low-copy repeats and was distributed mainly within all heterochromatin blocks and even in euchromatin regions. Further studies will be needed to determine more precisely the components and the copy number of each repeat of every Cot fraction.

The different chromatin types of tomato

Based on Bennetzen's concept of different plant heterochromatin hues (Bennetzen 2000), we propose here for tomato chromosomes a classification of six different chromatin classes, based on light-microscopic pachytene morphology and presence of specific repetitive DNA sequences:

1. *Euchromatin*. These are the chromosome regions with weak DAPI fluorescence in pachytene chromosomes and the domain of most of the coding sequences. Our FISH data showed little or no detectable amount of any of the repetitive sequences. FISH mapping of BACs located in these regions may not need Cot-100 for blocking of repeat signals.
2. *Chromomeres and other minor heterochromatin islands in the euchromatin regions*. A closer look at pachytene chromosomes stained with the classical chromosome staining solutions acetocarmine, or acetoorcein (Ramanna & Prakken 1967) or with the DNA-specific DAPI fluorochrome reveals smaller heterochromatin structures known as chromomeres or small heterochromatin knobs. Future FISH analysis of BACs covering these domains will be needed to elucidate whether these regions contain the short stretches of Ty1-*copia*, and other repeats that we incidentally detected in our study.
3. *Distal heterochromatin and long-arm interstitial heterochromatic knobs*. Studies of Ganal *et al*. (1988, 1991) and Zhong *et al*. (1998) have shown that all 24 chromosome ends contain the *Arabidopsis* type telomere repeat and all except for the arms 1L, 2S, 2L and 7L contain the TGRI repeat. In addition, TGRI was also detected in long-arm interstitial heterochromatin knobs of the chromosomes 6, 8 and 9, probably resulting from a DNA repeat copying mechanism between non-homologous chromosome alignments like the meiotic prophase I bouquet or mitotic interphase Rabl orientation (Zhong *et al*. 1998). TGRI is the diagnostic repeat for this chromatin class and can be helpful in identifying the chromosomes 1, 2, 3, 6, 8 and 9 depending on the long-arm interstitial heterochromatin polymorphisms in the cultivars or genotype under study. So far we have not been able to demonstrate whether interstitial heterochromatin knobs also contain TGRII or other repeats of the pericentromere heterochromatin class. We also observed in a recently performed BLAST analysis that TGRI is homologous to the 5′ part of a 4532 nt repeat which is classified as '*L. esculentum* microsatellite repeat DNA region' in the TIGR Solanaceae repeat database. Interestingly, it also shows a very high similarity to the 5′ part of '*L. esculentum* PCRT1a-like retrotransposon' in the the same database.
4. *Pericentromere heterochromatin*. This by far the largest chromatin class of the tomato chromosomes covers 65% of the total genomic DNA (Table 3), and contains TGRII, TGRIII (Ganal *et al*. 1988), $[GATA]_4$ and $[GA]_8$ microsatellites, and Ty1-*copia*-like retrotransposons. It is also the region where the bulk of the methylated DNA in the tomato genome is found. In our study we have shown that TGRII mainly occurs on all heterochromatin regions except for the NOR region, whereas TGRIII occurs on few specific pericentromere heterochromatin sites and clearly in different amounts between the chromosomes. The abundance of TGRII signals over TGRIII is consistent with the higher copy number of TGRII (Ganal *et al*. 1988 and Table 1). TGRII and TGRIII may be diagnostic molecular markers for heterochromatin regions in the ongoing Tomato Genome Sequencing program.

DNA sequencing showed that TGRII and TGRIII are components of Ty3-*Gypsy* type retrotransposons, and resemble two additional repeat classes recently discovered in the pericentromere heterochromatin regions (Yang *et al*. 2005). Over 2000 copies of the major pericentromere-specific retrotransposons, named PCRT1a (Yang *et al*. 2005) and *Jinling* (Wang *et al*. 2006) containing TGRII sequence, seem to have spread in the heterochromatin of tomato genome since about 5 million years ago (Wang *et al*. 2006). It is not known, however, whether the 452 bp pericentromere *Hind* III repeat THG2 described by Zabel *et al*. (1985) and Wolters *et al*. (1991) is part of the TGRII repeat, in spite of their apparent similar FISH distribution. Neither is it known where the chromosome 6 pericentromere sequences, showing homology to the human satellite III and the mammalian CENP-B binding box (Weide *et al*. 1998), are located. Other components of this heterochromatin class are the $[GA]_n$ and $[GATA]_n$

microsatellites (Figure 3f,h, Figure 4), of which $[GA]_n$ appeared much more abundant in the genome than $[GATA]_n$ (Broun & Tanksley 1996). However, $[GA]_8$ FISH was hardly seen in the euchromatin, which contradicts Broun & Tanksley (1996), who found $[GA]_n$ motifs in the GenBank database, and suggested that the $[GA]_n$ microsatellite might be located in the euchromatin. Such differences might be due to differences in detection of the two techniques; FISH detects only large arrays of tandem repeats, which may only be present in heterochromatin as was shown by Areshchenkova & Ganal (1999), while in the GenBank database short tandems may be more typical for euchromatin domains. Other more recent studies on the occurrence of microsatellites in tomato BAC end sequences shows that $[AT]_n$ is by far the most abundant microsatellite in tomato (E. Datema, unpublished).

The $[GATA]_4$ FISH patterns coincide with the genetic data of Arens *et al*. (1995) and Broun & Tanksley (1996), who mapped several clustered $[GATA]_n$ markers in the genetically defined (peri)centromeric region. One possible explanation for long tetranucleotide microsatellites in pericentromere heterochromatin is that they can only persist in areas with reduced recombination (Sherman & Stack 1995). $[GACA]_4$ mainly co-localized with 45S rDNA at NOR regions. Furthermore, Southern analysis often co-localized the GACA-repeat with the GATA-repeat (Vosman *et al*. 1992, Vosman & Arens 1997), but this could not be confirmed by FISH. Most of the bands identified with the $[GACA]_4$ probe also appeared with the $[GATA]_4$ probe, suggesting that they have a similar distribution pattern. Nine fragments containing $[GATA]_n$ mapped within clusters of markers that are believed to be adjacent to centromeres (Broun & Tanksley 1996, Arens *et al*. 1995, Grandillo & Tanksley 1995). Finally, large blocks of the [CAA] repeat were described by Yang *et al*. (2005) in the centromere of chromosome 12. FISH with this microsatellite showed abundant signals on all centromeres and pericentromere blocks but not in the distal heterochromatin blocks and the NOR. The Ty1-*copia* retrotransposon signals were dispersed over all heterochromatin blocks, weak in euchromatin regions and absent from the rDNA sites. Exclusion from (terminal) ribosomal sites is typical for Ty1-*copia*-like retrotransposons in various plant species (Brandes *et al*. 1997). Preferential heterochromatic accumulation has also been found for *Arabidopsis*, where Ty1-*copia*-like retrotransposons are clustered in the centromere heterochromatin (Brandes *et al*. 1997), and for *Allium cepa*, where accumulation occurs in the GC-rich heterochromatic regions (Pearce *et al*. 1996a). This is in contrast with *Vicia faba* (Pearce *et al*. 1996b) and *Alstroemeria aurea* (Kuipers *et al*. 1998), two species with large genomes, where Ty1-*copia*-like retrotransposons were reported to be excluded from AT-rich heterochromatic regions that contain highly repetitive tandem repeats.

Although all chromosomes have clear and conspicuous pericentromere regions, it is not easy to indicate precisely the borders of heterochromatin and euchromatin domains in the DAPI fluorescence patterns. FISH analyses with Cot-100 FISH on pachytene complements and sequence analyses of BACs around the heterochromatin euchromatin borders for the presence of pericentromere repeats are more precise and reliable tools for defining the borders more accurately (Szinay *et al*. 2008).

5. *Structural centromere heterochromatin*. The structural centromeres in tomato pachytene chromosomes have a smooth structure, are relatively uncondensed and have a DAPI fluorescence intensity comparable to or even less than that of euchromatin and always far less than that of the surrounding pericentromere heterochromatin blocks. We have shown that TGRIV is one of the major components. Yang *et al*. (2005) also showed a large block of the $[CAA]_n$ microsatellite in addition to a nested complex of various repeat sequences. There is still no indication which of these repeats are associated with centromere function.
6. *Nucleolar organizer region*. The largest heterochromatin segment of the cherry tomato is the macrosatellite and secondary constriction, which together form the nucleolar organizer region. Our FISH study showed exclusive signals of the 45S rDNA and $[GACA]_4$ for this region, whereas the bright propidium iodide fluorescence as used for genomic *in-situ* hybridization studies demonstrated a high GC ratio in this part of the genome. It is also the genomic part with high levels of DNA methylation.

Towards a bar coding of tomato repeats

This paper and several previous studies on tomato repeats have provided a plethora of repetitive sequences in this species, with which a bar-coded FISH ideogram as suggested by Schmidt & Heslop-Harrison (1998) can be constructed. However, as we show in our study and display schematically in Figure 4, most of the repeats reside in the pericentromere heterochromatin and are not diagnostic in the sense of providing unequivocal tools to identify chromosomes. With a few anchor-point BACs for each chromosome and repeats in a multicolour FISH (Szinay *et al*. 2008) it will be possible to construct repeat bar codes for every chromosome.

Acknowledgements

This research was co-financed by the European Commission (EU-SOL Project PL 016214) and by the Centre for BioSystems Genomics (CBSG) which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research.

References

Arens P, Odinot P, van Heusden S, Lindhout P, Vosman B (1995) GATA- and GACA-repeats are not evenly distributed throughout the tomato genome. *Genome* **38**: 84–90.

Areshchenkova T, Ganal MW (1999) Long tomato microsatellites are predominantly associated with centromeric regions. *Genome* **42**: 536–544.

Arumuganathan K, Earle E (1991) Estimation of nuclear DNA content of plants by flow cytometry. *Plant Mol Biol Rep.* **9**: 208–218.

Bennetzen JL (2000) The many hues of plant heterochromatin. *Genome Biology* **1**: reviews 107.1–107.4.

Brandes A, Heslop-Harrison JS, Kamm A, Kubis S, Doudrick RL, Schmidt T (1997) Comparative analysis of the chromosomal and genomic organisation of Ty1-copia like retrotransposons in pteridophytes, gymnosperms and angiosperms. *Plant Mol Biol* **33**: 11–21.

Broun P, Tanksley SD (1996) Characterization and genetic mapping in simple repeat sequences in the tomato genome. *Mol Gen Genet* **250**: 39–49.

Budiman MA, Mao L, Wood TC, Wing RA (2000) A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing. *Genome Res* **10**: 129–136.

Budiman MA, Chang SB, Lee S *et al.* (2004) Localization of *jointless*-2 gene in the centromeric region of tomato chromosome 12 based on high resolution genetic and physical mapping. *Theor Appl Genet* **108**: 190–196.

Chang S-B (2004) *Cytogenetic and molecular studies on tomato chromosomes using diploid tomato and tomato monosomic additions in tetraploid potato*. PhD thesis. University of Wageningen.

Copenhaver G, Nickel K, Kuromori T *et al.* (1999) Genetic definition and sequence analysis of *Arabidopsis* centromeres. *Science* **286**: 2468–2474.

Cuadrado A, Schwarzacher T (1998) The chromosomal organization of simple sequence repeats in wheat and rye genomes. *Chromosoma* **107**: 587–594.

De Jong JH, Fransz P, Zabel P (1999) High resolution FISH in plants – techniques and applications. *Trends Plant Sci* **4**: 258–262.

De Jong JH, Zhong X-B, Fransz PF, Wennekes-van Eden J, Jacobsen E, Zabel P (2000) High resolution FISH reveals the molecular chromosomal organisation of repetitive sequences of individual tomato chromosomes. In: Olmo E, Redi CA, eds. *Chromosomes Today*. Switserland: Birkhäuser Verlag, **13**: 267–275.

Ewing B, Green P (1998) Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res* **8**: 186–194.

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**: 175–185.

Flavell AJ, Smith DB, Kumar A (1992) Extreme heterogeneity of Ty1-*copia* group retrotransposons in plants. *Mol Gen Genet* **231**: 233–242.

Ganal MW, Lapitan NLV, Tanksley SD (1988) A Molecular and Cytogenetic Survey of Major Repeated DNA Sequences in Tomato *Lycopersicon esculentum*. *Mol Gen Genet* **213**: 262–268.

Ganal MW, Lapitan LV, Tanksley SD (1991) Macrostructure of the tomato telomeres. *Plant Cell* **3**: 87–94.

Grandillo S, Tanksley SD (1995) QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *L. pimpinellifolium*. *Theor Appl Genet* **92**: 935–951.

Jiang J, Birchler JA, Parrott WA, Dawe RK (2003) A molecular view of plant centromeres. *Trends Plant Sci* **8**: 570–575.

Jurka J (2005) GYPSODE1: Gypsy-type element from wild potato. *Repbase Reports* **5**: 246.

Khush GS, Rick CM (1968) Cytogenetic analysis of the tomato genome by means of induced deficiencies. *Chromosoma* **23**: 452–484.

Kocsis E, Trus BL, Steer CJ, Bisher ME, Steven AC (1991) Image averaging of flexible fibrous macromolecules: the clathrin triskelion has an elastic proximal segment. *J Struct Biol* **107**: 6–14.

Kuipers AGJ, Heslop-Harrison JS, Jacobsen E (1998) Characterisation and physical localisation of Ty1-*copia*-like retrotransposons in four *Alstroemeria* species. *Genome* **41**: 357–367.

Lapitan NLV, Ganal MW, Tanksley SD (1989) Somatic chromosome karyotype of tomato based on *in situ* hybridization of the TGRI satellite repeat. *Genome* **32**: 992–998.

Lapitan NLV, Ganal MW, Tanksley SK (1991) Organization of the 5S ribosomal RNA genes in the genome of tomato. *Genome* **34**: 509–514.

Nagaki K, Cheng Z, Ouyang S *et al.* (2004) Sequencing of a rice centromere uncovers active genes. *Nat Genet* **36**: 138–145.

Pearce SR, Pich U, Harrison G *et al.* (1996a) The Ty1-*copia* group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. *Chromosome Res* **4**: 357–364.

Pearce SR, Harrison G, Li D, Heslop-Harrison JS, Kumar A, Flavell AJ (1996b) The Ty1-*copia* group retrotransposons in *Vicia* species: copy number, sequence homogeneity and chromosomal localisation. *Mol Gen Genet* **250**: 305–315.

Peterson DG, Stack SM, Price HJ, Johnston JS (1995) Distribution of DNA in heterochromatin and euchromatin of *Lycopersicon esculentum* pachytene chromosomes. *Tomato Genet Coop Rep* **45**: 35.

Peterson DG, Price HJ, Johnston JS, Stack SM (1996) DNA content of heterochromatin and euchromatin in tomato (*Lycopersicon esculentum*) pachytene chromosomes. *Genome* **39**: 77–82.

Peterson DG, Pearson WR, Stack SM (1998) Characterization of the tomato (*Lycopersicon esculentum*) genome using *in vitro* and *in situ* DNA reassociation. *Genome* **41**: 346–356.

Peterson DG, Lapitan NLV, Stack SM (1999) Localization of single- and Low-copy sequences on tomato synaptonemal complex spreads using fluorescence *in situ* hybridization (FISH). *Genetics* **152**: 427–439.

Peterson DG, Schulze SR, Sciara EB *et al.* (2002) Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery. *Genome Res* **12**: 795–807.

Ramanna MS, Prakken P (1967) Structure of and homology between pachytene and somatic metaphase chromosomes of the tomato. *Genetica* **38**: 115–133.

Richards EJ, Ausubel FM (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana. Cell* **53**: 127–136.

Schmidt T, Heslop-Harrison JS (1998) Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends Plant Sci* **3**: 195–199.

Schwartz S, Zhang Z, Frazer KA *et al.* (2000) PipMaker – A web server for aligning two genomic DNA sequences. *Genome Res* **10**: 577–586.

Schweizer G, Ganal M, Ninnemann H, Hemleben V (1988) Species-specific DNA sequences for identification of somatic hybrids between *Lycopersicon esculentum* and *Solanum acaule*. *Theor Appl Genet* **75**: 679–684.

Sherman JD, Stack SM (1995) Two-dimensional spreads of synaptonemal complexes from solanaceous plants. VI. High resolution recombination nodule map for tomato (*Lycopersicon esculentum*). *Genetics* **141**: 683–708.

Song J, Dong F, Lilly JW, Stupar RM, Jiang J (2001) Instability of bacterial artificial chromosome (BAC) clones containing tandemly repeated DNA sequences. *Genome* **44**: 463–469.

Szinay D, Chang S-B, Khrustaleva L *et al.* (2008) High-resolution chromosome mapping of BACs using multi-colour FISH and pooled-BAC FISH as a backbone for sequencing tomato chromosome 6. *Plant J* (in press).

Tanksley SD, Ganal MW, Prince JP *et al.* (1992) High density molecular linkage maps of the tomato and potato genomes. *Genetics* **132**: 1141–1160.

Tilford CA, Kuroda-Kawaguchi T, Skaletsky H *et al.* (2001) A physical map of the human Y chromosome. *Nature* **409**: 943–945.

Van Daelen RAJJ, Zabel P (1994) Preparation of high molecular weight plant DNA and analysis by pulsed field gel electrophoresis. In: Gelvin SB, Schilperoort RA, eds. *Plant Molecular Biology Manual*. Dordrecht: Kluwer Academic, **H3**: 1–21.

Van der Hoeven R, Ronning C, Giovannoni J, Martin G, Tanksley SD (2002) Deductions about the number, organization and evolution of genes in the tomato genome based on analysis of a large EST collection and selective genomic sequencing. *Plant Cell* **14**: 1441–1456.

Vosman B, Arens P (1997) Molecular characterization of GATA/GACA microsatellite repeats in tomato. *Genome* **40**: 25–33.

Vosman B, Arens P, Rus-Kortekaas W, Smulders MJM (1992) Identification of highly polymorphic DNA regions in tomato. *Theor Appl Genet* **85**: 239–244.

Wallace RB, Johnson MJ, Hirose T, Miyake T, Kawashima EH, Itakura K (1981) The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequences to rabbit β-globin DNA. *Nucleic Acids Res* **9**: 879–894.

Wang Y, Tang X, Cheng Z, Mueller L, Giovannoni J, Tanksley SD (2006) Euchromatin and pericentromere heterochromatin: comparative composition in the tomato genome. *Genetics* **172**: 2529–2540.

Weide RJ, Hontelez J, van Kammen A, Koornneef M, Zabel P (1998) Paracentromeric sequences on tomato chromosome 6 show homology to human satellite III and to the mammalian CENP-B binding box. *Mol Gen Genet* **259**: 190–197.

Wolters AMA, Schoenmakers HCH, van der Meulen-Muisers JJM *et al.* (1991) Limited DNA elimination from the irradiated potato parent in fusion products of albino *Lycopersicon esculentum* and *Solanum tuberosum*. *Theor Appl Genet* **83**: 225–232.

Xu J, Earle ED (1994) Direct and sensitive fluorescence *in situ* hybridization of 45S rDNA on tomato chromosomes. *Genome* **37**: 1062–1065.

Xu J, Earle ED (1996a) High resolution physical mapping of 45S (5.8S, 18S and 25S) rDNA gene loci in the tomato genome using a combination of karyotyping and FISH of pachytene chromosomes. *Chromosoma* **104**: 545–550.

Xu J, Earle ED (1996b) Direct FISH of 5S rDNA on tomato pachytene chromosomes places the gene at the heterochromatic knob immediately adjacent to the centromere of chromosome 1. *Genome* **39**: 216–221.

Yang TJ, Lee S, Chang SB, Yu Y, de Jong JH, Wing RA (2005) In depth sequence analysis of the centromeric region of tomato chromosome 12: Identification of a large CAA block and characterization of centromeric retrotranposons. *Chromosoma* **114**: 103–117.

Zabel P, Meyer D, van de Stolpe O *et al.* (1985) Towards the construction of artificial chromosomes for tomato. In: van Vloten-Doting L, Groot GSP, Hall TC, eds. *Molecular Form and Function of the Plant Genome*. New York: Plenum, 609–624.

Zamir D, Tanksley SD (1988) Tomato genome is comprised largely of fast-evolving, low copy-number sequences. *Mol Gen Genet* **213**: 254–261.

Zhong XB, De Jong JH, Zabel P (1996) Preparation of tomato meiotic pachytene and mitotic metaphase chromosomes suitable for fluorescence *in situ* hybridization (FISH). *Chromosome Res* **4**: 24–28.

Zhong XB, Fransz PF, Wennekes VEJ *et al.* (1998) FISH studies reveal the molecular and chromosomal organization of individual telomere domains in tomato. *Plant J* **13**: 507–517.

Zwick MS, Hanson RE, McKnight TD *et al.* (1997) A rapid procedure for the isolation of Cot-1 DNA from plants. *Genome* **40**: 138–142.