



## Chromosomal dynamics of human neocentromere formation

Peter E. Warburton

Dept. of Human Genetics, Box 1498, Mount Sinai School of Medicine, 1425 Madison Ave, East Bldg 14-52A, New York, NY 10029, USA; Tel: 212-659-6728; Fax: 212-849-2508; E-mail: peter.warburton@mssm.edu

*Key words:* acentric, aneuploidy, centromere, chromatin

### Abstract

Neocentromeres are rare human chromosomal aberrations where a new centromere has formed in a previously non-centromeric location. The emergence of new centromeres on a chromosome that already contains an endogenous centromere would be a highly deleterious event which would lead to dicentricity and mitotic instability. Nonetheless, neocentromere formation appears to provide a mechanism for the acquisition of a new centromere. Neocentromeres are most often observed on chromosomal arm fragments that have separated from an endogenous centromere, and therefore actually lead to mitotic stability of what would have been an acentric fragment. Neocentromeres have recently also been observed on apparently unrearranged chromosomes where the endogenous centromere has been inactivated. Furthermore, the process of centromere repositioning during primate chromosomal evolution may depend on the acquisition and subsequent fixation of neocentromeres. This remarkable plasticity in the position of centromeres has important implications for human cytogenetics and chromosome evolution, and provides an opportunity to further our understanding of the process of centromere formation and structure.

Human neocentromeres have been identified primarily by clinical cytogenetics labs in patients that are dysmorphic or developmentally delayed. In most cases, routine molecular cytogenetic chromosome analysis reveals a supernumerary chromosome with unusual characteristics. These neocentric chromosomes are often 'marker' chromosomes, which are defined as supernumerary chromosomes whose origin cannot be identified by conventional banding techniques. When informative, the banding pattern of neocentric chromosomes suggests derivation from a chromosomal arm, often as a distal symmetrical inverted duplication, which can be corroborated using whole chromosome painting FISH probes.

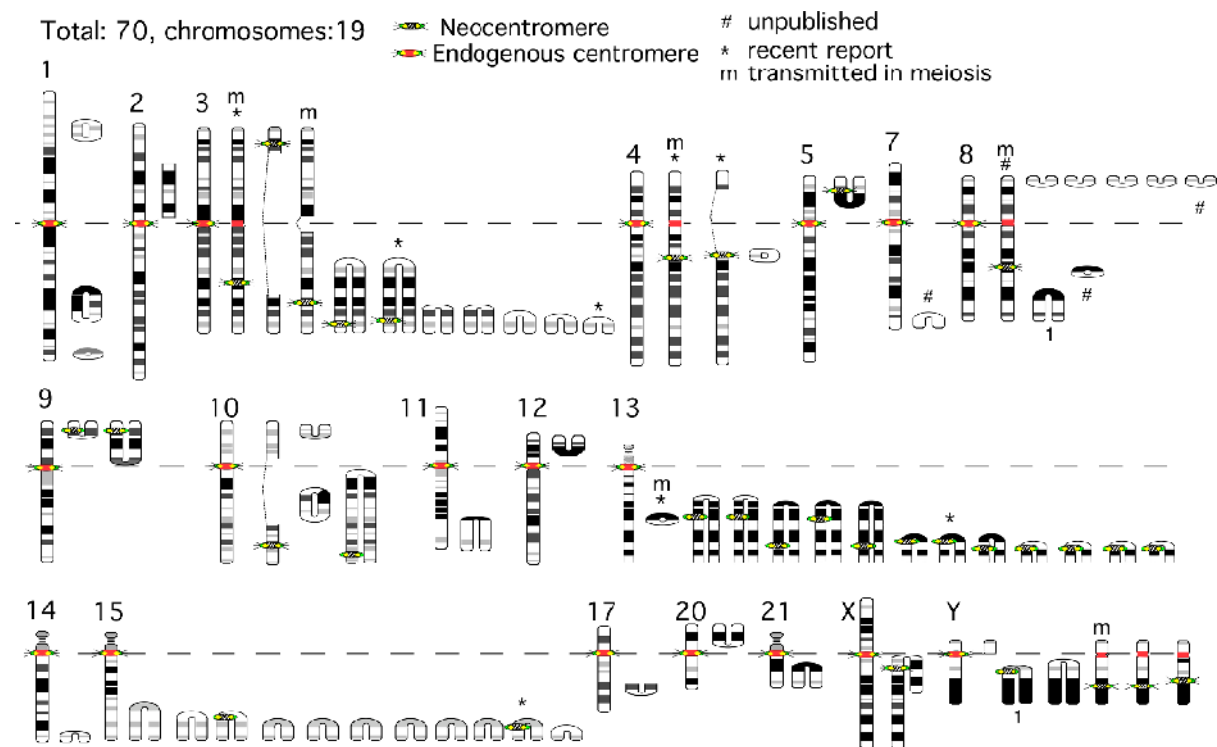
These neocentric chromosomes contain a normal looking centromeric constriction. However, they do not show any hybridization using chromosome-specific centromeric alpha satellite DNA probes to investigate their chromosomal origin, confirming that they do not contain endogenous centromeres. At this point, a rare neocentric chromosome is indicated, which can be confirmed using cytogenetic techniques that are generally not standard in most clinical labs. The chromosomal origin of these neocentric markers can be precisely determined by microdissection and hybridization back to normal metaphase chromosomes (Knecht *et al.* 2002), or by Comparative Genomic Hybridization if there is aneuploidy (Levy *et al.* 2000). The presence of a

neocentromere can be confirmed by immunofluorescence using antibodies to CENTromere Proteins (CENPs) e.g. CENP-C (Warburton *et al.* 2000).

### Human neocentromere cases

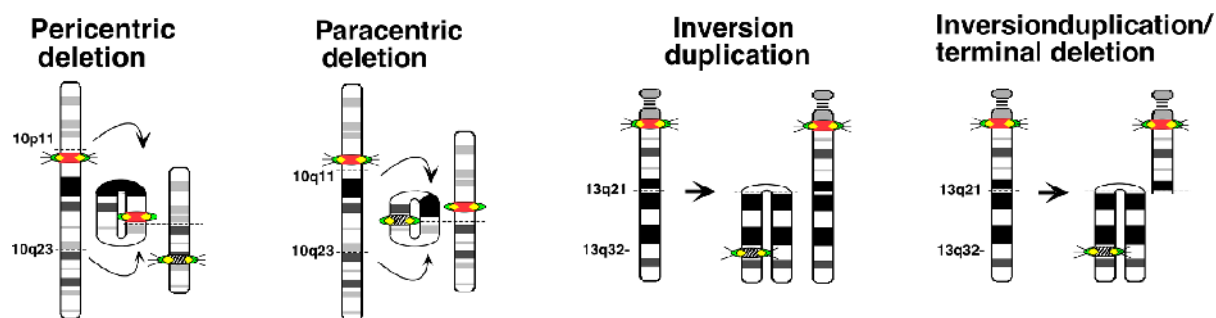
At the time of writing this review, 70 neocentric chromosomes have been identified, including several unpublished cases that have been characterized in my laboratory in collaboration with

clinical cytogenetic labs (Figure 1) (Amor & Choo 2002). These fall into several classes depending on their origin. Neocentromeres have arisen by interstitial chromosomal 'deletions', either pericentric or paracentric (Figure 2). In most cases, the corresponding deleted chromosome was also present, resulting in a balanced karyotype. This deletion and ring excision process is consistent with intrachromatid exchange between direct repeats at the breakpoints (Stankiewicz & Lupski 2002). The most common class of neocentric chromosomes are supernumerary



**Figure 1.** All chromosomes containing neocentromeres thus far described. A total of 70 neocentric chromosomes have been described on derivatives of 19 different human chromosomes. A normal ideogram for each of the 19 chromosomes is shown. To the right of these are ideograms showing the type and position of each neocentric chromosome derivative, and when known the position of the neocentromere. Paracentric deletions give rise to derivative chromosomes contained wholly in one arm (e.g. examples from chromosomes 1 and 2), while pericentric deletions give rise to fused chromosome fragments spanning the endogenous centromere (e.g. examples from chromosome 3 and chromosome 10). Inversion duplications of distal arms of chromosomes are also observed (e.g. examples from chromosome 15). The neocentric chromosomes 4 and 8 are shown.

A complete list of cases up to 2002 was described in Amor & Choo, 2002. Seven additional cases have been published since that review, marked with an asterisk (\*) (Knegt *et al.* 2002, Barbi *et al.* 2003, Papenhausen *et al.* 2003, Ventura *et al.* 2003, 2004, Amor *et al.* 2004, Barwell *et al.* 2004, Warburton *et al.* 2004). Four unpublished cases characterized in our laboratory have also been included in this list, marked by a number sign (#) (P. Papenhausen, J. Tepperberg, H. Mak-Tam, J. Engelen, P. Warburton, unpublished/in preparation). An 'm' indicates that the neocentric chromosome has been transmitted in meiosis. A '1' indicates neocentromeres reported twice in Amor & Choo, 2002 due to two published reports on the same chromosome.



**Figure 2.** Chromosomal rearrangements associated with neocentromere formation. Pericentric deletions involve excision of a centric fragment, usually a ring chromosome, which contains the endogenous centromere, and fusion of the two acentric arms, which is mitotically rescued by neocentromere formation. Paracentric deletions involve excision of an acentric fragment, which is mitotically rescued by neocentromere formation, and fusion of the centric and acentric fragment. Inversion duplications involve an inversion and duplication of a distal portion of a chromosome. These are sometimes accompanied by complementary terminal deletions.

inverted duplications (invdups) of the distal arm of a chromosome. These cases either have two normal corresponding chromosomes, resulting in tetrasomy for the duplicated region, or an apparently complementary deletion giving rise to trisomy for the duplicated region (Figure 2).

An additional class of novel neocentric chromosomes has recently been described on chromosomes 3, 4 and 8 (Figure 1). In these cases, an apparently unrearranged autosome was observed to contain an active neocentromere in a position well removed from the endogenous centromere, which is inactivated, giving rise to a neodicentric chromosome where the functional centromere has shifted to a new location on a chromosome arm (Amor *et al.* 2004, Engelen & Warburton, in preparation, Ventura *et al.* 2004). Similar neodicentric chromosomes have been previously described for the Y chromosome (Tyler-Smith *et al.* 1999). However, the centromeres of human Y chromosomes contain that smallest arrays of alpha satellite DNA, which do not contain the binding site for CENP-B, and thus may represent a 'weak' centromere that might be easily inactivated. The large amount of heterochromatin found in the Y chromosome long arm may support neocentromere formation. Thus, it seems possible that the Y chromosome would be more prone to neodicentricity than autosomes (Warburton 2001, Amor *et al.* 2004).

There is an ascertainment bias of observed neocentric chromosomes for compatibility for recognized pregnancy or fetal survival, which probably

explains the relatively small size of most detected neocentric chromosomes, and/or the balanced karyotype of most interstitial deletion cases. The severity of the phenotype of neocentromere cases is dependent on the degree of aneuploidy resulting from the stability of the acentric fragment, and can range from severe neonatal lethal in cases of tri- or tetrasomy to phenotypically normal in balanced cases.

The detection of neocentric chromosomes in phenotypically normal adults (Knegt *et al.* 2002, Amor *et al.* 2004), usually ascertained due to routine chromosome screening for fertility problems or affected offspring, raises the question of how common neocentromeres may be in the population. While neocentromeres are no doubt exceedingly rare, an unbiased assessment of neocentromere frequency and associated risk of abnormal phenotype, as performed for human marker chromosomes (Warburton 1991), would help to elucidate their potential impact on human health and chromosome evolution.

Thus far, 6 of the 70 neocentric chromosomes described (Figure 1) have been observed in more than one generation of a family, although it seems reasonable to assume that all neocentromeres would be competent in meiosis if given the chance. In the only case of meiotic transmission through the mother, a phenotypically normal woman with a small 13q21 neocentric ring chromosome and complementary deleted chromosome 13 (Knegt *et al.* 2002) passed the neocentric ring to a fetus, which was electively aborted (Knegt &

Warburton, unpublished). In such cases, the ring chromosome, with no meiotic pairing partner, would probably randomly segregate with either the normal or the deleted homologous chromosomes (Knecht *et al.* 2002). In a similar case, a phenotypically normal father passed a large neocentric chromosome 3 to his daughter, but a complementary small centric chromosome 3 did not cosegregate, resulting in partial monosomy in a dysmorphic daughter, for whom chromosome analysis was indicated (Wandall *et al.* 1998).

The remaining four meiotically transmitted neocentric chromosomes consisted of the three neodicentric autosomes (Amor *et al.* 2004, Engelen & Warburton, in preparation, Ventura *et al.* 2004) and a neodicentric Y chromosome (Tyler-Smith *et al.* 1999). In each case, the neodicentric chromosome was ascertained fortuitously due to routine chromosome analysis, usually amniocentesis for advanced maternal age, and subsequently identified in the phenotypically normal father. Meiotic pairing of the neodicentric chromosome and its homologue should occur relatively normally, although the apparent meiotic success of these chromosomes suggests that homologous centromeres need not be directly opposed for proper segregation to occur in meiosis I. However, recombination between the endogenous centromere and the neocentromere would lead to a dicentric and an acentric chromosome, which would probably result in spermatogenic failure in meiosis II or early embryonic loss. Interestingly, inactivation of the neocentromere in such a dicentric chromosome would restore the chromosome to its normal state.

#### **Neocentromere hotspots: chromosomal rearrangements and/or neocentromere formation**

Neocentromeres have been observed on 19 different human chromosomes, and it is reasonable to assume that they will eventually be found on the remaining chromosomes as well (Figure 1). Nevertheless, there is a striking non-randomness to the distribution of neocentromeres in the genome. Several genomic regions have been observed to form a disproportionate number of neocentromeres, including 3q, 13q and 15q.

Indeed, as more total cases of neocentromeres have been described, additional examples from these 'neocentromere hotspots' have been added to the list, and consistently account for about one half of all neocentromere cases. However, additional regions of interest are also emerging. For example, three neocentromeres have been described on derivatives of chromosome 4, all in approximately band 4q21 (Figure 1) (Grimbacher *et al.* 1999, Amor *et al.* 2004, Warburton *et al.* 2004), which would be highly unlikely if neocentromeres are equally likely to form across chromosome 4. In addition, five invdup 8p23-pter neocentric chromosomes have been observed.

These observations raise the question of whether neocentric chromosome formation results from two independent rare events, chromosomal rearrangement and neocentromere emergence, and is therefore a very rare event (the product of the frequencies of both events). Alternatively, these events may be mechanistically dependent on each other, which might increase the observed frequency of neocentric chromosomes. The three examples of neodicentric autosomes as well as the neodicentric Y chromosomes support the idea that neocentromeres can arise ectopically in otherwise normal unrearranged chromosomes (Amor *et al.* 2004). The three neocentric chromosome 4 derivatives that have been observed illustrate three potential mechanisms for stabilization of a chromosome on which a neocentromere has arisen: (1) inactivation of the endogenous centromere, with no rearrangement; (2) pericentric deletion of a ring containing the endogenous centromere; or (3) paracentric deletion of a ring containing the neocentromere (Figure 1). Two of these mechanisms may also be illustrated by the neodicentric chromosome 8 and paracentric 8q21 ring chromosome. The interstitial deletion events in these cases may actually be facilitated by the emergence of a neocentromere, which results in initial functional dicentricity of these chromosomes and chromosome breakage events.

The paucity of observed neodicentric chromosomes thus far may suggest that inactivation of endogenous centromeres is a less likely event than chromosomal rearrangements for stabilization of neocentromeres. However, these neodicentric chromosomes are likely to be overlooked by

clinical labs more often than supernumerary chromosomes, because: (1) they have an identical banding pattern to the normal chromosome and have a normal karyotype, except for a repositioned centromeric constriction, and (2) they thus far do not present a phenotype and need to be ascertained fortuitously. It is remarkable that the first three examples have been ascertained independently within approximately the last year, and it will be of interest to see how many more are observed by clinical labs in the future.

Other neocentric chromosomes suggest that rearrangement may be a driving force in their formation. For example, five neocentric invdup8(p23-pter) chromosomes have been observed (Ohashi *et al.* 1994, Giglio *et al.* 2001, Voullaire *et al.* 2001, Papenhausen, Tepperberg and Warburton, unpublished). Several direct and inverted low-copy repeats and a common inversion polymorphism at human chromosome 8p23 appears to predispose this region to chromosomal rearrangements (Giglio *et al.* 2001, Sugawara *et al.* 2003). Thus, an increase in the frequency of chromosomal rearrangements leading to acentric fragments on a background of rare neocentromere formation may increase the likelihood of observing these neocentric chromosomes. This model supposes that most neocentromeres that arise in the population are lost due to resulting chromosomal instability, unless they arise on, or result in the generation of, rearranged chromosomal fragments that are compatible with life.

It is possible that some chromosomal rearrangements themselves provide the means to acquire a neocentromere. Although small neocentric invdup marker chromosomes are usually asymmetrical with respect to the position of the neocentromere, this may actually reflect an asymmetry in the composition of the chromosome itself. FISH mapping of the breakpoints of several small invdup chromosomes from both 8p and 15q have suggested that between the two duplicated symmetrical arms is found an unduplicated region containing proximal sequences contiguous with one of the arms (Giglio *et al.* 2001, Ventura *et al.* 2003). Such an arrangement is consistent with meiotic recombination between chromosomes that are heterozygous for a polymorphic inver-

sion flanked by low copy inverted repeats, with the single copy region representing a portion of the polymorphic inverted region. However, the data as presented in these reports is difficult to distinguish from the simple loss of resolution and merging of FISH signals as a symmetrical breakpoint is approached in these small highly condensed chromosomes.

In two neocentric invdup15q chromosomes with different breakpoints in 15q24.1 and in 15q25.3, the neocentromeres have been localized to the breakpoint boundary between the single copy and the duplicated DNA, to a remarkably high resolution for such a FISH based analysis (Ventura *et al.* 2003). At face value these studies suggest that, in some neocentric invdup chromosomes where the neocentromere is near the breakpoint junction, even if it appears off center, neocentromeres may have formed at the position of chromosomal rearrangement, perhaps because the rearrangement and repair provide the opportunity to acquire centromeric chromatin. It must be stressed, however, that in the majority of cases neocentromere formation has occurred far removed from the chromosomal rearrangement.

Two examples of experimentally-derived neocentromere formation have been described in *Drosophila* (Williams *et al.* 1998) and in animal cell culture (Shen *et al.* 2001). In both of these examples, a region that had been in close proximity to an endogenous centromere on a minichromosome acquired a neocentromere after breakage and rearrangement, consistent with a localized cis-spreading of centromeric chromatin into the neocentromere. However, all human neocentromeres thus far described are well removed from endogenous centromeres, and so do not support a cis-acting spreading mechanism. A trans-spreading mechanism between transiently associated neo- and endogenous centromeres could be hypothesized for the initial acquisition of centromeric chromatin in humans.

#### Analysis of 'neocentromere DNA'

The high number of observed 13q neocentromeres cases (Figure 1) may reflect an ascertainment bias for viability of chromosome 13 trisomies, although the paucity of similar

cases on chromosomes 18 and 21 and the X chromosome, which are equally well tolerated in trisomy, suggest that 13q has a genuinely higher propensity for formation of neocentric derivatives (Warburton *et al.* 2000). Seven independent neocentromeres have been cytologically localized to chromosome band 13q32. Each of these is found on an invdup13q chromosome, which had between them at least five distinct inversion breakpoints in bands from 13q32 to 13q21 (Figure 1; Warburton *et al.* 2000, Li *et al.* 2002). Thus, these relatively common 13q32 neocentromeres are more consistent with the hypothesis that there is a 'hotspot' for neocentromere formation in 13q32, which is subsequently stabilized by rearrangement events across the chromosome. One possible explanation for this neocentromere hotspot is the presence of a 'neocentromere DNA' sequence in 13q32 that has formed multiple independent neocentromeres.

This 'neocentromere DNA' hypothesis was tested by identifying the underlying DNA at three independent 13q32 neocentromeres, using chromatin immunoprecipitation (ChIP) with antibodies to CENP-A, the centromere-specific histone H3 homologue. This CENP-A ChIP DNA was hybridized to a genomic microarray containing 14 Mbp of contiguous BACs spanning band 13q32 (Alonso *et al.* 2003), permitting the identification of the BAC(s) that contain CENP-A chromatin. This experiment showed that the CENP-A chromatin domains for each of the three 13q32 neocentromeres tested was in a distinct location, occupying a maximum of 215 kb, 130 kb and 275 kb. All three CENP-A domains were within ~6.5 Mbp in 13q32, and were separated by ~5 Mbp and ~1 Mbp, respectively. Thus, these results do not support the presence of a particular DNA sequence within 13q32 that accounts for the neocentromere hotspot.

The analysis of three 13q32 neocentromeres doubles the number of CENP-A domains that have been localized within neocentromeres, added to two that were previously identified by CENP-A ChIP in 10q25 and 20p12 (Lo *et al.* 2001a, 2001b), and one identified using only immunofluorescence in 9p23 (Satinover *et al.* 2001). Extensive sequence analysis and comparison of these six regions failed to reveal any conserved motifs, homologies to centromeric sequences such as

alpha satellite DNA, or tandemly arrayed DNAs, that may have suggested an explanation for neocentromere formation. Indeed, the only characteristics these DNA regions had in common were a slightly elevated AT richness and a corresponding decrease in density of SINEs, characteristics that are also shared with alpha satellite DNA arrays at endogenous centromeres (Lo *et al.* 2001a, 2001b, Alonso *et al.* 2003).

This analysis of the DNA sequence at a total of six independent CENP-A binding domains seems to close the book on the existence of a 'neocentromere DNA' that has a high propensity for CENP-A binding and account for the observed human neocentromeres. Nonetheless, it remains possible that common sequence motifs may be involved in specifying other centromeric chromatin domains which lead to neocentromere formation. In this light, the ~1 Mbp separation of two 13q32 neocentromere CENP-A domains is well within the range where they may have overlapping heterochromatin domains that share a common DNA sequence (Alonso *et al.* 2003).

#### **Epigenetics of chromatin domain structures at neocentromeres**

If neocentromeres are not determined by sequence, than a strictly epigenetic mechanism for their formation and propagation must be considered. A well-accepted model for the propagation of centromeres in general is the epigenetic templating of CENP-A, where the presence of CENP-A at a chromosomal location marks the location for deposition of additional CENP-A, which propagates the centromere through subsequent cell cycles (Sullivan *et al.* 2001). This model is independent of the primary DNA sequence on which the CENP-A is found. Thus, once the CENP-A domain has been ectopically established at neocentromeres, it can propagate itself using the same epigenetic mechanism as a normal endogenous centromere.

The determining factor in neocentromere formation is the epigenetic acquisition of centromeric chromatin, including CENP-A nucleosomes, as well as a host of other centromere and mitotic kinetochore associated proteins (Saffery *et al.* 2000, Craig *et al.*

2003a). Presumably, this chromatin contains the structural 'blueprint' which allows the centromeric chromatin to fold correctly in three dimensions and assemble the large centromere/kinetochore complex (Sullivan *et al.* 2001). For proper centromere function, this region must contain chromatin domains for kinetochore assembly on the outside faces of the sister chromatids, which is probably designated by the CENP-A chromatin in the inner kinetochore plate. It must also contain domains for sister chromatid cohesion located between the sister chromatids, which in *S. pombe* has been associated with heterochromatin (Bernard *et al.* 2001). Neocentromeres represent a unique opportunity to map the organization of these centromeric chromatin domains onto the underlying single copy DNA, which is not possible at endogenous centromeres due to the highly homologous repetitive DNA found there (Choo 2001).

As described above, the CENP-A chromatin domains have been mapped by ChIP and BAC array hybridization for a total of five neocentromeres. In each case, the resolution of this assay is limited by the size and degree of overlap of the BACs used to span the region, which is the likeliest explanation for the relatively small size of the chromosome 13q32 CENP-A domains compared to the ones from 10q25 and 20p12 (Lo *et al.* 2001a, 2001b, Alonso *et al.* 2003). However, extended-fibre chromatin analysis of human and *Drosophila* centromeres suggests that CENP-A chromatin is interspersed with histone H3 containing chromatin in subdomains of 10–40 kb (Blower *et al.* 2002). Further ChIP mapping of the size and organization of CENP-A chromatin subdomains within the several hundred kb domains that are currently mapped should provide considerable insight into the three-dimensional structural models of the path of DNA between the inner kinetochore plate and heterochromatin domains.

The domain organization of the 10q25 neocentromere has been further investigated using ChIP and BAC arrays, as well as other techniques, and compared to the corresponding region in a normal chromosome 10 (Saffery *et al.* 2003). An increased density of scaffold/matrix attachments (S/MAR) has been observed across a 3.5 Mbp domain that contains the CENP-A chromatin, which may be related to the primary constriction

at this neocentromere. Overlapping one edge of this S/MAR domain is found a 900 kb domain that is associated with CENP-H. Interestingly, this CENP-H domain is separated from the CENP-A domain by ~1 Mbp, even though CENP-A is required for localization of CENP-H to centromeres (Goshima *et al.* 2003). These results suggest an interaction between these widely spaced CENP-A and CENP-H chromatin domains as part of assembly of the kinetochore (Saffery *et al.* 2003).

### Heterochromatin at neocentromeres

On the other side of the high-density S/MAR domain at the 10q25 neocentromere is found a surprisingly small domain (~100 kb) of heterochromatin protein 1 (HP1) associated chromatin, although this is consistent with reports of a decreased immunofluorescence signal for HP1 at neocentromeres relative to endogenous centromeres (Saffery *et al.* 2000, Craig *et al.* 2003a). Suv39H1, which is associated with HP1, has been localized to a chromosome 13 neocentromere (Aagaard *et al.* 2000) but a formal demonstration at neocentromeres of histone H3 methylated at Lys 9 remains elusive. Other more classical cytogenetic hallmarks of centromeric heterochromatin e.g. constitutive heterochromatin banding (C-banding) and Dapi/distamycin brightness are absent from neocentromeres. One main difference from endogenous centromeres is the absence from neocentromeres of alpha satellite DNA and CENP-B, which may contribute considerably to centromeric heterochromatin structure by the dimerization of CENP-B bridging alpha satellite DNA repeat units (Ando *et al.* 2002).

The amount and extent of centromeric heterochromatin that is present at neocentromeres appears to be reduced relative to endogenous centromeres. A reduction in sister chromatid cohesion at the neodicentric chromosome 4 may reflect this reduced amount of heterochromatin by reduction in cohesin binding at this neocentromere (Bernard & Allshire 2002, Amor *et al.* 2004). A molecular analysis of the extent of heterochromatin and HP1 domains at endogenous human centromeres has not been performed due

in part to the repetitive DNA structure. Thus it is not clear to what extent the HP1 domain structure at neocentromeres reflects the situation at endogenous centromeres, although the ~1 Mbp separation from the CENP-A domain is consistent with the separation across the chromatid of the inner kinetochore plate and the sister chromatid cohesion. These neocentromere results may suggest that only a small portion of 'centromeric heterochromatin' at endogenous centromeres contains HP1 and is necessary to form a sister chromatid cohesion domain, with the remainder intervening 'heterochromatin' composed of alpha satellite DNA/CENP-B chromatin.

The relatively small size of the HP1 domain was somewhat unexpected, but is consistent with an even more unexpected result – that the transcriptional activity of genes that reside in the 10q25 neocentromere remains unchanged relative to the normal chromosome 10 (Saffery *et al.* 2003). This is true in the region of increased S/MAR density and the CENP-H domain. A gene found in the HP1 domain could not be evaluated due to absence of expression in the somatic cell hybrid lines used, but HP1 is known to usually be a powerful transcriptional repressor. Remarkably, a gene that spans the CENP-A domain was also expressed at normal levels. It may be of significance that the 5' end and promoter region of this gene fell outside the CENP-A domain, given the importance of histone H3 N-terminal tail modifications in transcriptional activation, but transcription can at least proceed through CENP-A chromatin. The idea that centromeres are not necessarily a transcriptional graveyard for genes is further supported by a recent analysis of the rice chromosome 8 centromere, where at least 4 genes that reside in the kinetochore domain are expressed, although they appear to correlate with regions of discontinuity for CENH3 (rice CENP-A) chromatin (Nagaki *et al.* 2003). None of the five additional human neocentromeres contained genes within their mapped CENP-A domains. In fact, the size and position of two of the chromosome 13q32 CENP-A domains precisely fit between genes in a relatively gene-dense region, suggesting an avoidance of active chromatin (Alonso *et al.* 2003).

### Neocentromere positioning

Analysis of the replication timing around the 10q25 neocentromere revealed that the CENP-A domain resided in the latest replicating region within the surrounding 5 Mbp (Lo *et al.* 2001a), although an adjacent region was actually slightly later to replicate. These data suggest that CENP-A chromatin may favour locally late-replicating DNA within a particular cluster of replication domains, but that other factors may fine tune this affinity and affect the final position. Interestingly, when these cells were treated with TSA, a histone deacetylase inhibitor, the CENP-A domain shifted 320 kb onto the latest replicating DNA (Craig *et al.* 2003b). This shift may have occurred because the fine-tuning of the affinity of CENP-A for the original region was decreased by hyperacetylation of core histones, and it now moved to the neighbouring late-replicating region. A lowered affinity of CENP-A for chromatin with hyperacetylated core histones would be consistent with its tendency to localize in regions without genes. The observed unidirectional shift of the CENP-A only towards the late-replicating region may have to do with the fact that, in the other direction, there is an abrupt change to earlier replication timing. Upon removal of the TSA, the CENP-A domain shifted back to its original position, presumably because the fine tuning of the affinity of CENP-A for this region is restored. Thus, the position of the CENP-A domain appears to be affected by combinations of several factors, such as local replication timing, core histone modifications and/or chromatin activity. This remarkable pliability of the position of the CENP-A domain may be akin to the cis-spreading of neocentromere chromatin observed in *Drosophila* and animals (Williams *et al.* 1998, Shen *et al.* 2001).

One area of study that may shed light on the question of neocentromere positioning is the examination of evolutionary centromere repositioning. Comparison of chromosomes 9 and X in humans and primates have revealed a conservation of marker order on these chromosomes between species, which however does not include the position of the centromere (Montefalcone

*et al.* 1999, Ventura *et al.* 2001). These results are explained most parsimoniously by a process of centromere emergence on the otherwise unarranged chromosome, which is equivalent to neocentromere formation and fixation in the different species. Indeed, the neodicentric chromosomes recently described may represent this neocentromere repositioning in progress, especially since they have been observed to go through meiosis (Amor *et al.* 2004, Engelen & Warburton, in preparation). Recently, the position of two chromosome 15q human neocentromeres were found to be 8 Mb proximal and 1.5 Mb distal to the centromere on the equivalent chromosome in Old World monkeys (Ventura *et al.* 2003). Interestingly, this region contains a high density of chromosome-specific duplicons, which were postulated to be involved in the high frequency of centromere emergence in this region. It will be of interest to determine if other ancestral centromeres in primates correspond to the position of neocentromeres that have been observed in humans.

In conclusion, there are few data that directly address the question of why neocentromeres form in particular chromosomal locations and what triggers their formation. Possible factors include a weak sequence bias towards AT-rich DNA, but otherwise DNA sequence seems to play little role. In one studied case, localized late replication may play a role (Lo *et al.* 2001a), although additional studies of replication timing at other neocentromeres would help to assess the significance of this observation. Neocentromeres may form preferentially in regions that already possess 'heterochromatic qualities', such as a domain of HP1 (perhaps involved in gene regulation) or even a small amount of orphaned CENP-A in a non-centromeric location. Neocentromeres may be capable of forming in essentially any chromosomal location, and the observed 'hotspots' are due to the clinical ascertainment bias for survival, or regions prone to chromosomal rearrangements. However, regardless of the mechanism of their formation, once neocentromeres have formed and are established in stable cell lines, they provide an invaluable system to examine human centromere structure in relation to the underlying DNA sequence. As more and more neocentromeres are found on

different types of chromosomes, they will continue to provide important insight into human centromere structure and function.

### Acknowledgements

The author would like to thank Ellen Mak-Tam (North York General Hospital, Canada), John Engelen (Academic Hospital Maastricht, Netherlands), Peter Papenhausen and Jim Tepperberg (Labcorp, North Carolina, USA) and Lia Knecht (University of Amsterdam, Netherlands) for unpublished neocentromere cases, and Andy Choo (University of Melbourne) and Mariano Rocchi (University of Bari) for work prior to publication. This work was supported by NIH grants GM61150 and HG02919.

### References

- Aagaard L, Schmid M, Warburton PE, Jenuwein T (2000) Mitotic chromatin association of SUV39H, a novel component of active centromeres, correlates with cell cycle dependent phosphorylation. *J Cell Sci* **113**: 817–829.
- Alonso A, Mahmood R, Li S *et al.* (2003) Genomic microarray analysis reveals distinct locations for the CENP-A binding domains in three human chromosome 13q32 neocentromeres. *Hum Mol Genet* **12**: 2711–2721.
- Amor DJ, Choo KH (2002). Neocentromeres: role in human disease, evolution, and centromere study. *Am J Hum Genet* **71**: 695–714.
- Amor D, Bentley K, Ryan J *et al.* (2004) Human centromere repositioning "in progress". *Proc Natl Acad Sci USA* **101**: 6542–6547.
- Ando S, Yang H, Nozaki N, Okazaki T, Yoda K (2002) CENP-A, -B, and -C chromatin complex that contains the I-type alpha-satellite array constitutes the prekinetochore in HeLa cells. *Mol Cell Biol* **22**: 2229–2241.
- Barbi G, Spaich C, Adolph S, Kehrer-Sawatzki H (2003) Anaphoid *de novo* marker chromosome inv dup(3)(q28qter) with neocentromere in a dysmorphic and developmentally retarded girl. *J Med Genet* **40**: e27.
- Barwell J, Mazzaschi R, Bint S, Ogilvie C, Elmsie F (2004) A new neocentromere locus on chromosome 13 resulting in mosaic tetrasomy for distal 13q and an asymmetric phenotype. *Am J Med Genet* (in press).
- Bernard P, Allshire R (2002) Centromeres become unstuck without heterochromatin. *Trends Cell Biol* **12**: 419.
- Bernard P, Maure JF, Partridge JF *et al.* (2001) Requirement of heterochromatin for cohesion at centromeres. *Science* **294**: 2539–2542.

- Blower MD, Sullivan BA, Karpen GH (2002) Conserved organization of centromeric chromatin in flies and humans. *Dev Cell* **2**: 319–330.
- Choo KH (2001) Domain organization at the centromere and neocentromere. *Dev Cell* **1**: 165–177.
- Craig JM, Earle E, Canham P *et al.* (2003a) Analysis of mammalian proteins involved in chromatin modification reveals new metaphase centromeric proteins and distinct chromosomal distribution patterns. *Hum Mol Genet* **12**: 3109–3121.
- Craig JM, Wong LH, Lo AW, Earle E, Choo KH (2003b) Centromeric chromatin pliability and memory at a human neocentromere. *EMBO J* **22**: 2495–2504.
- Giglio S, Broman KW, Matsumoto N *et al.* (2001) Olfactory receptor-gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. *Am J Hum Genet* **68**: 874–883.
- Goshima G, Kiyomitsu T, Yoda K, Yanagida M (2003) Human centromere chromatin protein hMis12, essential for equal segregation, is independent of CENP-A loading pathway. *J Cell Biol* **160**: 25–39.
- Grimbacher B, Dutra AS, Holland SM *et al.* (1999) Anaphoid marker chromosome in a patient with hyper-IgE syndrome, autism, and mild mental retardation. *Genet Med* **1**: 214–219.
- Knecht AC, Li S, Engelen JJM, Bijlsma EK, Warburton PE (2002) Prenatal diagnosis of a karyotypically normal pregnancy in a mother with a supernumerary neocentric 13q21->13q22 chromosome and balancing reciprocal deletion. *Prenat Diagn* **23**: 215–220.
- Levy B, Papenhausen PR, Tepperberg JH *et al.* (2000) Prenatal molecular cytogenetic diagnosis of partial tetrasomy 10p due to neocentromere formation in an inversion duplication anaphoid marker chromosome. *Cyto Cell Genet* **91**: 165–170.
- Li S, Malafej P, Levy B *et al.* (2002) Chromosome 13q neocentromeres: Molecular cytogenetic characterization of three additional cases and clinical spectrum. *Am J Med Genet* **110**: 258–267.
- Lo AW, Craig JM, Saffery R *et al.* (2001a) A 330 kb CENP-A binding domain and altered replication timing at a human neocentromere. *EMBO J* **20**: 2087–2096.
- Lo AW, Magliano DJ, Sibson MC *et al.* (2001b) A novel chromatin immunoprecipitation and array (CIA) analysis identifies a 460-kb CENP-A-binding neocentromere DNA. *Genome Res* **11**: 448–457.
- Montefalcone G, Tempesta S, Rocchi M, Archidiacono N (1999) Centromere repositioning. *Genome Res* **9**: 1184–1188.
- Nagaki K, Talbert PB, Zhong CX *et al.* (2003) Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of *Arabidopsis thaliana* centromeres. *Genetics* **163**: 1221–1225.
- Ohashi H, Wakui K, Ogawa K *et al.* (1994) A stable acentric marker chromosome: Possible existence of an intercalary ancient centromere at distal 8p. *Am J Hum Genet* **55**: 1202–1208.
- Papenhausen P, Gadi I, Tepperberg J *et al.* (2003) Trisomy 3q secondary to a terminal deletion/generation of a mirror-image anaphoid marker in a neonate. *Am J Hum Genet* **73**: Suppl. A841.
- Saffery R, Irvine DV, Griffiths B *et al.* (2000) Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins. *Hum Mol Genet* **9**: 175–185.
- Saffery R, Sumer H, Hassan S *et al.* (2003) Transcription within a functional human centromere. *Mol Cell* **12**: 509–516.
- Satinover DL, Vance GH, VanDyke DL, Schwartz S (2001) Cytogenetic analysis and construction of a BAC contig across a common neocentromeric region from 9p. *Chromosoma* **110**: 275–283.
- Shen MH, Ross A, Yang J, de las Heras JI, Cooke H (2001) Neo-centromere formation on a 2.6 Mb mini-chromosome in DT40 cells. *Chromosoma* **110**: 421–429.
- Stankiewicz P, Lupski JR (2002) Genome architecture, rearrangements and genomic disorders. *Trends Genet* **18**: 74–82.
- Sugawara H, Harada N, Ida T *et al.* (2003) Complex low-copy repeats associated with a common polymorphic inversion at human chromosome 8p23. *Genomics* **82**: 238–244.
- Sullivan BA, Blower MD, Karpen GH (2001) Determining centromere identity: cyclical stories and forking paths. *Nat Rev Genet* **2**: 584–596.
- Tyler-Smith C, Gimelli G, Giglio G *et al.* (1999) Transmission of a fully functional human neocentromere through three generations. *Am J Hum Genet* **64**: 1440–1444.
- Ventura M, Archidiacono N, Rocchi M (2001) Centromere emergence in evolution. *Genome Res* **11**: 595–599.
- Ventura M, Mudge JM, Palumbo V *et al.* (2003) Neocentromeres in 15q24–26 map to duplicons which flank an ancestral centromere in 15q25. *Genome Res* **13**: 2059–2068.
- Ventura A, Weigl S, Carbone L *et al.* (2004). Recurrent sites for new centromere seeding. *Genome Res* (in press).
- Voullaire L, Saffery R, Earle E *et al.* (2001) Mosaic inv dup(8p) marker chromosome with stable neocentromere suggests neocentromerization is a post zygotic event. *Am J Med Genet* **102**: 86–94.
- Wandall A, Tranebjaerg L, Tommerup N (1998) A neocentromere on human chromosome 3 without detectable alpha satellite DNA forms morphologically normal kinetochores. *Chromosoma* **107**: 359–365.
- Warburton D (1991) *De novo* balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. *Am J Hum Genet* **49**: 995–1013.
- Warburton PE (2001) Epigenetic analysis of kinetochore assembly at human variant centromeres. *Trends Genet* **17**: 243–247.
- Warburton PE, Dolled M, Mahmood R *et al.* (2000) Molecular cytogenetic analysis of eight inversion duplications of human chromosome 13q that each contain a *Am J Hum Genet* **66**: 1794–1806.
- Warburton P, Barwell J, Splitt M *et al.* (2004) Class II neocentromeres: a putative common neocentromere site in band 4q12.2. *Am J Med Genet* **11**: 749–753.
- Williams BC, Murphy TD, Goldberg ML, Karpen GH (1998) Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nature Genet* **18**: 30–37.