



## Intrachromosomal homologous recombination between inverted amplicons on opposing Y-chromosome arms

Julian Lange<sup>a,\*</sup>, Michiel J. Noordam<sup>b,1,3</sup>, Saskia K.M. van Daalen<sup>b</sup>, Helen Skaletsky<sup>a</sup>, Brian A. Clark<sup>c</sup>, Merryn V. Macville<sup>d</sup>, David C. Page<sup>a</sup>, Sjoerd Repping<sup>b,\*\*</sup>

<sup>a</sup> Whitehead Institute, Howard Hughes Medical Institute, and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142, USA

<sup>b</sup> Center for Reproductive Medicine, Women's and Children's Hospital, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

<sup>c</sup> Center for Medical Genetics, University of Pittsburgh, Pittsburgh, PA 15213, USA

<sup>d</sup> Cytogenetics Unit, Department of Clinical Genetics, University Hospital Maastricht, Maastricht, The Netherlands

### ARTICLE INFO

#### Article history:

Received 30 January 2013

Accepted 25 April 2013

Available online 2 May 2013

#### Keywords:

Human Y chromosome

Non-allelic homologous recombination

Structural variation

Male infertility

### ABSTRACT

Amplicons – large, nearly identical repeats in direct or inverted orientation – are abundant in the male-specific region of the human Y chromosome (MSY) and provide targets for intrachromosomal non-allelic homologous recombination (NAHR). Thus far, NAHR events resulting in deletions, duplications, inversions, or isodicentric chromosomes have been reported only for amplicon pairs located exclusively on the short arm (Yp) or the long arm (Yq). Here we report our finding of four men with Y chromosomes that evidently formed by intrachromosomal NAHR between inverted repeat pairs comprising one amplicon on Yp and one amplicon on Yq. In two men with spermatogenic failure, sister-chromatid crossing-over resulted in pseudoisoyp chromosome formation and loss of distal Yq. In two men with normal spermatogenesis, intrachromatid crossing-over generated pericentric inversions. These findings highlight the recombinogenic nature of the MSY, as intrachromosomal NAHR occurs for nearly all Y-chromosome amplicon pairs, even those located on opposing chromosome arms.

© 2013 Elsevier Inc. All rights reserved.

### 1. Introduction

The male-specific region of the human Y chromosome (MSY) contains many amplicons – large, nearly identical repeats – whose sequence similarity is maintained by gene conversion [1,2]. These long segments of high sequence identity render the Y chromosome susceptible to intrachromosomal homologous recombination that can result in interstitial deletions, duplications, inversions, or isodicentric chromosomes [3–13]. Interstitial Y deletions and isodicentric Y chromosomes are associated with a wide range of sex disorders,

including male infertility, Turner syndrome, and sex reversal (reviewed in [14]).

Whereas each of the intrachromosomal homologous recombination events reported to date involve amplicons located on the same Y-chromosome arm, the reference Y chromosome also contains two sets of inverted repeats (IRs) that are composed of one amplicon on the short arm (Yp) and one amplicon on the long arm (Yq), namely IR1 and IR4 [1]. IR1 is composed of two amplicons that share >99% sequence identity over ~62 kilobases (kb), and IR4 is composed of two amplicons that share ~94% sequence identity over ~303 kb. Of note, the IR1 repeat on Yq is located within the *azoospermia factor c* (*AZFc*) region that contains genes essential for spermatogenesis and is almost entirely ampliconic [6].

We hypothesized that intrachromosomal homologous recombination between amplicons of IR1 or between amplicons of IR4 can generate two types of rearrangements: pseudoisochromosomes, in which the two chromosome ends are identical and in mirror-image orientation, and pericentric inversions (Fig. 1). For example, resolution of a double-strand break (DSB) in the Yq copy of IR1 by inter-sister-chromatid crossing-over with the Yp copy would produce a pseudoisoyp chromosome, which carries a partial duplication of Yp and a partial deletion of Yq, and a pseudoisoyq chromosome, which carries a partial duplication of Yq and partial deletion of Yp. Transmission of the former would likely result in a male offspring with impaired spermatogenesis due to the absence of multiple genes from the *AZFc* region. Alternatively, resolution

**Abbreviations:** MSY, male-specific region of the Y chromosome; NAHR, non-allelic homologous recombination; *AZFc*, *azoospermia factor c*; IR, inverted repeat; DSB, double-strand break; STS, sequence-tagged site; FISH, fluorescence in situ hybridization; PAR, pseudoautosomal region.

\* Correspondence to: J. Lange, Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. Fax: +1 212 7173627.

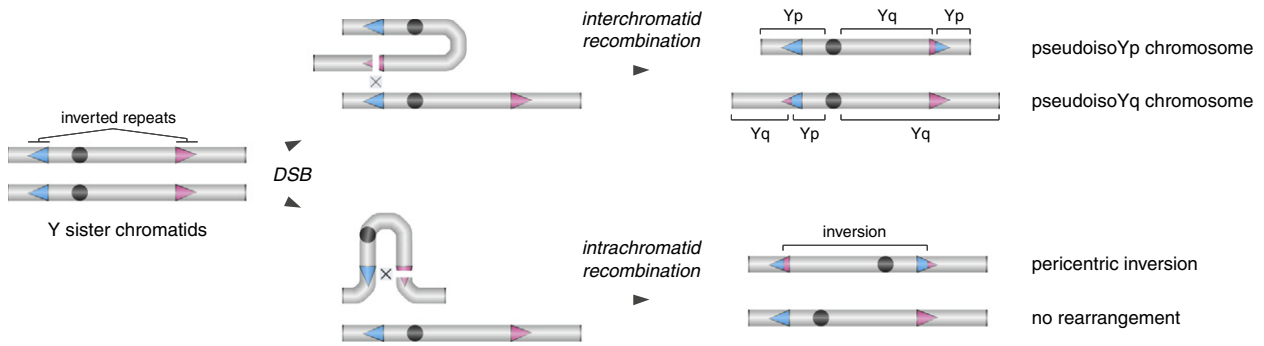
\*\* Correspondence to: S. Repping, Center for Reproductive Medicine, Women's and Children's Hospital, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. Fax: +31 0205669044.

**E-mail addresses:** [langej@mskcc.org](mailto:langej@mskcc.org) (J. Lange), [s.repping@amc.uva.nl](mailto:s.repping@amc.uva.nl) (S. Repping).

<sup>1</sup> These authors contributed equally to this work.

<sup>2</sup> Current affiliation: Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA.

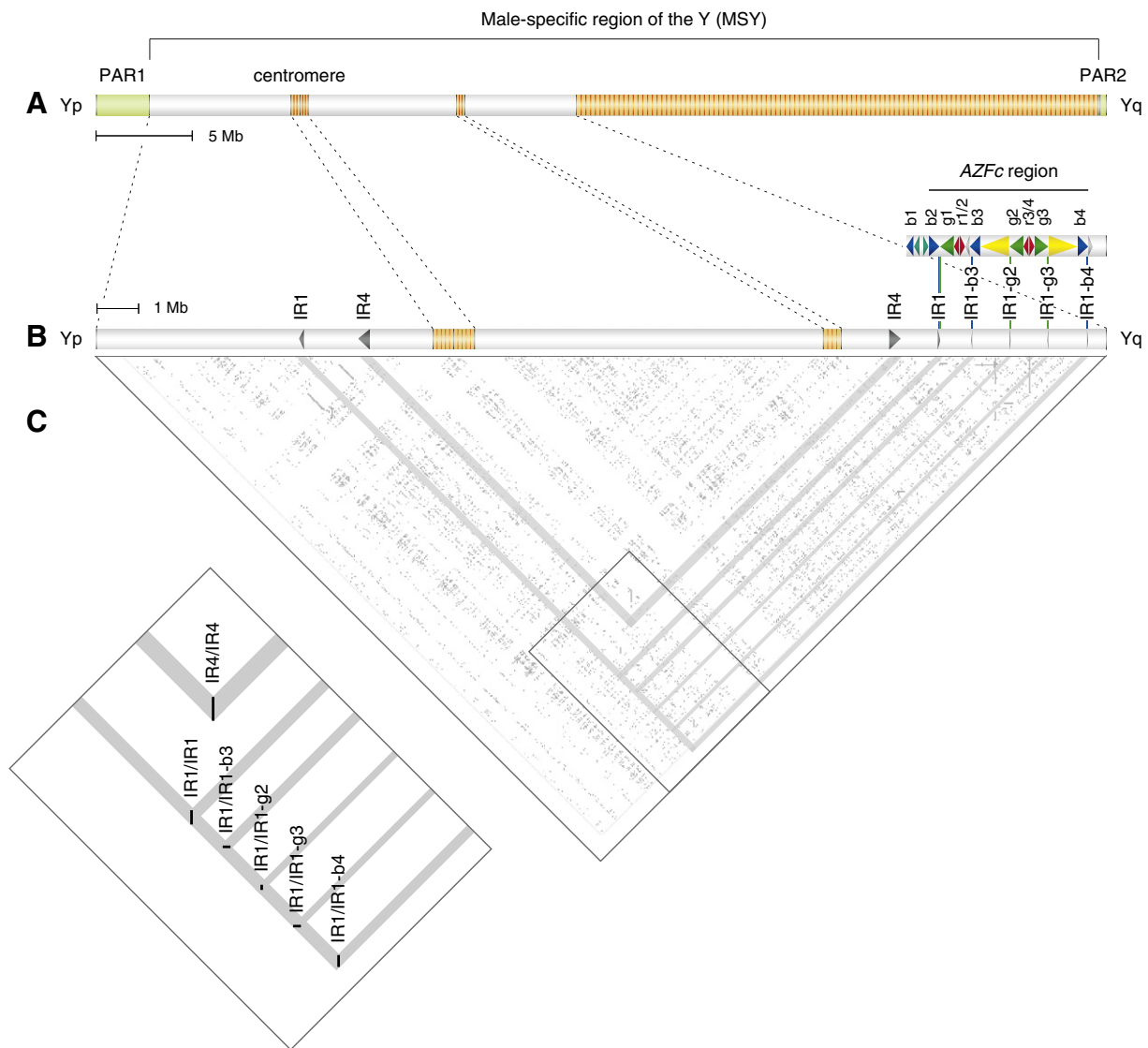
<sup>3</sup> Current affiliation: Department of Genetics, Washington University, St. Louis, MO, USA.



**Fig. 1.** Mechanisms of pseudoisochromosome formation and pericentric inversion. Ectopic homologous recombination between inverted repeats composed of one amplicon on the short arm (Yp) and one amplicon on the long arm (Yq): crossing-over between sister chromatids generates pseudoisochromosomes, intrachromatid crossing-over produces a pericentric inversion.

of a DSB in the Yq copy of IR1 by intrachromatid crossing-over with the Yp copy would lead to a pericentric inversion in that chromatid. Although pseudoisoy chromosomes and Y-chromosome pericentric

inversions have been described previously [15–26], it is unknown if these rearrangements indeed are generated via homologous recombination between inverted amplicons.



**Fig. 2.** Locations and orientations of Yp–Yq inverted repeats on the human Y chromosome. (A) Schematic representation of the human Y chromosome. The male-specific region of the Y chromosome, or MSY, is flanked by two pseudoautosomal regions, PAR1 and PAR2 (green), and contains blocks of heterochromatin (orange). (B) Inverted repeats IR1 and IR4 are composed of one amplicon on the short arm (Yp) and one amplicon on the long arm (Yq). In addition, segments of IR1 are repeated distally to IR1 on Yq in the highly ampliconic AZFc region: IR1-b3, IR1-g2, IR1-g3, and IR1-b4. (C) Triangular dot plot of the reference sequence shows length and orientation of each repeat pair. Each dot represents a 100% match over a window of 100 bp. Repeat elements have been masked. Inverted repeats appear as vertical lines, direct repeats as horizontal lines. Diagonal shading added to highlight the inverted repeats studied here. A magnification of the region bounded by the rectangle is schematized at bottom left.

Here, we report four Y chromosomes – two pseudoisoy chromosomes and two with pericentric inversions – that evidently formed by homologous recombination between amplicons of IR1 or of IR4. We then catalog all copies of IR1 and of IR4 that could be potential targets for such ectopic homologous recombination events, as well as the putative resultant pseudoisoy chromosomes and pericentric inversions. Finally, we discuss the spermatogenic phenotypes of two men with pseudoisoy chromosomes and the relationship of these findings to an inclusive model of ectopic homologous recombination in the human MSY.

## 2. Results

### 2.1. IR1 and IR4: pairs of inverted repeats on opposing Y-chromosome arms

We closely examined the MSY reference sequence to determine the precise structure of the IR1 and IR4 repeats on Yp and on Yq (Fig. 2A–B, Table 1, and Supplementary Table S1) [1]. On Yp, the IR1 and IR4 amplicons are each present in one complete copy. Similarly, on Yq, the IR4 amplicon is present in only one contiguous copy. In contrast, the IR1 amplicon on Yq is located within the 3.5-megabase (Mb) AZFc region, which comprises multiple copies of five different amplicons arrayed in direct and inverted orientations: four copies of blue (b1–b4), two copies of turquoise, three copies of green (g1–g3), four copies of red (r1–r4), and two copies of yellow [6]. IR1 on Yq partially overlaps two of the AZFc amplicons; it is composed of 30.3 kb of the 228-kb amplicon b2 and 7.6 kb of the 314.7-kb amplicon g1, as well as of 24.5 kb of intervening sequence termed u3. Thus, there are four additional Yq loci – portions of amplicons b3, g2, g3, and b4 – that share high sequence identity with segments of IR1. (The segment of IR1 that overlaps b2 and is repeated in b3 and b4 is absent in b1.) We termed these loci IR1–b3, IR1–g2, IR1–g3, and IR1–b4 (Fig. 2B and Supplementary Table S1).

To examine the sequence structure of the AZFc region in relation to Yp, we generated a triangular dot plot of the Yp region encompassing IR1 and IR4 against the Yq region encompassing AZFc (Fig. 2C). The 7.6-kb IR1–g2 and 30.3-kb IR1–b4 segments on Yq are in inverted

orientation relative to IR1 on Yp and thus present additional targets for recombination by our model (Table 1).

### 2.2. Cases

On the basis of karyotyping, we identified from our sample collections two men, WHT5557 and AMC1574, with pseudoisoy chromosomes, and two men, AMC0972 and AMC0973, with pericentric inverted Y chromosomes. WHT5557 and AMC1574 were diagnosed with azoospermia and severe oligozoospermia, respectively. In contrast, both men with pericentric inversions had normal spermatogenesis. We hypothesized that each of these rearrangements was generated through intrachromosomal homologous recombination between one of the pairs of Yp–Yq repeats. To test this hypothesis, we employed two tools to define the recombination breakpoints in these cases: 1) a series of precisely mapped sequence-tagged sites (STSs), whose presence or absence is readily assayed by PCR on genomic DNA; and 2) fluorescence in situ hybridization (FISH) of Y-specific probes, whose copy number and chromosomal location enabled us to distinguish between various MSY structures.

### 2.3. Pseudoisoy chromosomes

We had previously localized, in case WHT5557, a single breakpoint in the IR4 amplicon on Yq bounded by STSs sY1279 and sY1278: all STSs on Yp and proximal to IR4 on Yq were present, and all STSs distal to IR4 on Yq were absent [13]. We used high-resolution breakpoint mapping to further delineate the breakpoint region in this case. Making use of the ~6% sequence divergence between IR4 on Yp and IR4 on Yq, we employed plus/minus STS assays specific to the Yq repeat unit to precisely localize the breakpoint to an 800-bp interval (Fig. 3A and Supplementary Table S2) [27]. We then designed a primer pair for junction amplification by PCR, with one primer immediately proximal to the Yq breakpoint interval and a second primer immediately distal to the homologous interval on Yp. This primer pair amplified a product in WHT5557 but not in control individuals. Sequencing of the junction product confirmed the expected sequences to be those of Yq and Yp on each side of a 153-bp segment of perfect identity (Fig. 3B), indicating that the pseudoisoy chromosome in this case had formed by inter-sister-chromatid recombination between IR4 amplicons (Fig. 3C).

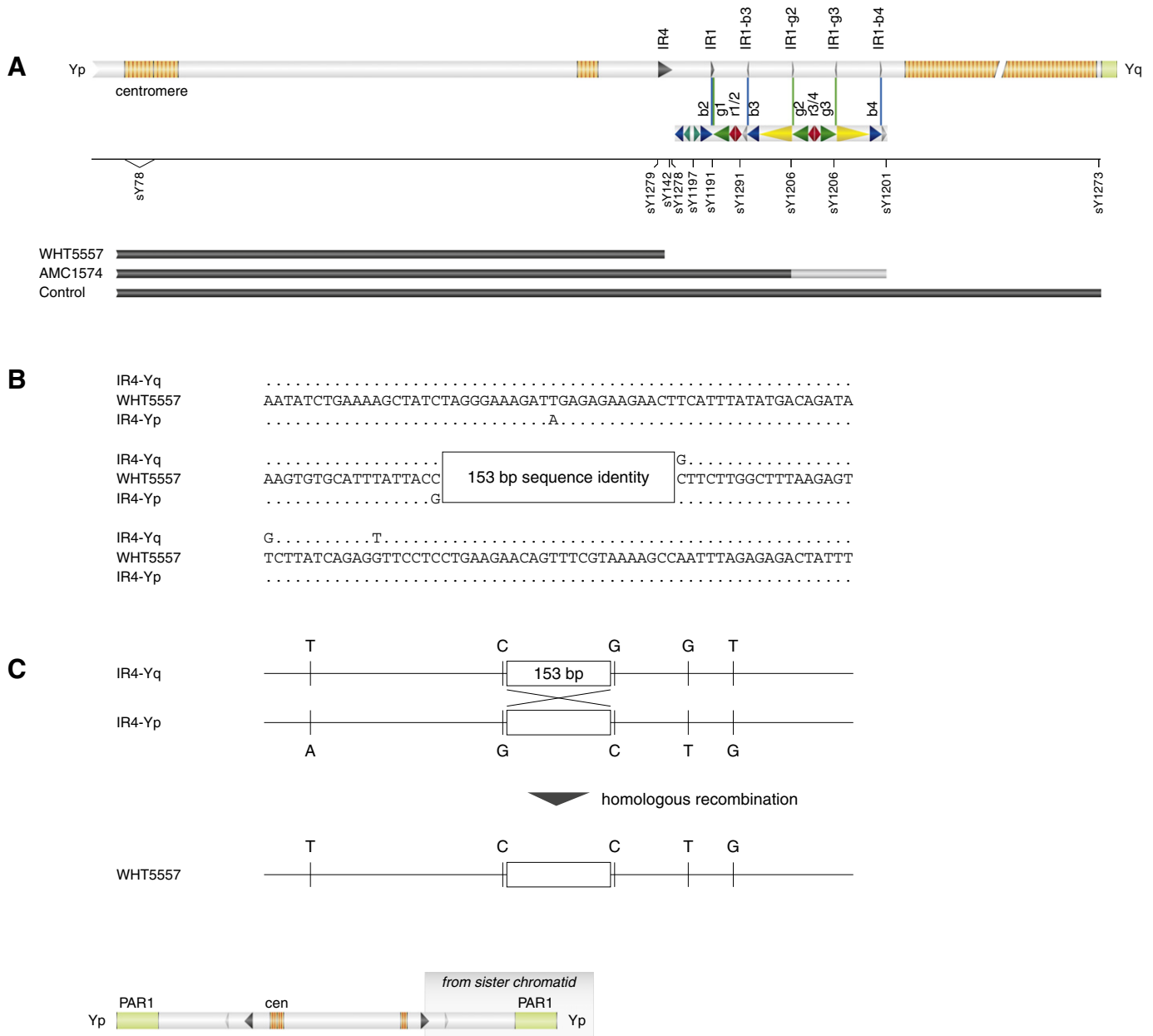
In AMC1574, STS mapping revealed a single breakpoint between the proximal copy of sY1206 and sY1201, in the distal part of AZFc (Fig. 3A and Supplementary Table S2). Since this breakpoint region contains the IR1–g2 and IR1–b4 repeat units, we reasoned that the pseudoisoy chromosome observed by karyotyping had arisen by inter-sister-chromatid recombination between either IR1–g2 or IR1–b4 on Yq and IR1 on Yp. Both such events preserve highly sequence-similar IR1 amplicons on Yq, thus high-resolution breakpoint mapping was not feasible for this case. Instead, we performed single-color FISH experiments on metaphase and interphase spreads to assay the copy number and chromosome-arm location of several loci ordinarily located only on either Yp or Yq (Fig. 4A).

Metaphase FISH in AMC1574 demonstrated that probe pDP1335, normally mapping to SRY on distal Yp, was present on both chromosome arms (Fig. 4B). Similarly, FISH probes RP11–199M2, targeting AMELY, and 17224/17225, located immediately distal to the Yp copy of IR1, were present on both chromosome arms in AMC1574, but only on Yp in a control sample. In contrast, two other FISH probes from Yp – 17228/17229, located immediately proximal to the Yp copy of IR1, and RP11–516H8, a BAC targeting the TSPY array on proximal Yp – were present exclusively on one chromosome arm in AMC1574. These results suggested that the recombination event on Yp had occurred in the region between 17224/17225 and 17228/17229, that is, in IR1. On Yq, all green, red, and yellow amplicons from the AZFc region remained exclusively on this chromosome arm in AMC1574 (Fig. 4B). Interphase FISH demonstrated reference copy

**Table 1**  
Catalog of inverted repeat targets involving IR1 or IR4 amplicons on Yp and Yq.

Structure of IR3/IR3	Structure of AZFc	Yp target	Yq target	Size (kb)	Sequence identity (%)	Supplementary Fig. S3
Reference IR3/IR3	Reference AZFc	IR1	IR1	62.4	99.66	A
		IR1	IR1–g2	7.6	99.57	B
		IR1	IR1–b4	30.3	99.76	C
		IR4	IR4	303.2	93.76	D
	b2/b3 inversion	IR1	IR1–b3	30.3	99.72	E
		IR1	IR1–g2	7.6	99.57	F
		IR1	IR1–b4	30.3	99.76	G
		IR4	IR4	303.2	93.76	H
	gr/rg inversion	IR1	IR1	62.4	99.66	I
		IR1	IR1–b3	30.3	99.72	J
		IR1	IR1–b4	30.3	99.76	K
		IR4	IR4	303.2	93.76	L
IR3/IR3 inversion	Reference AZFc	IR1	IR1–b3	30.3	99.72	M
		IR1	IR1–g3	7.6	99.54	N
	b2/b3 inversion	IR1	IR1	62.4	99.66	O
		IR1	IR1–g3	7.6	99.54	P
		IR1	IR1–g2	7.6	99.57	Q
		IR1	IR1–g3	7.6	99.54	R

Note that a 50-kb sequence gap is shared by the IR4 amplicons; sequence identity was calculated on the ~250 kb of sequence known for each amplicon. Inverted repeat pairs of IR1 or IR4 amplicons are identified in schematized dot plots in Supplementary Fig. S2. For each of the 18 inverted repeat pairs, homologous recombination can generate three derivative Y chromosomes, as shown in Supplementary Fig. S3.



**Fig. 3.** Location of the breakpoints in two men with pseudoisoy chromosomes. (A) Above, schematic representation of the long arm (Yq) of the human Y chromosome, with positions of plus/minus STSs used in low-resolution mapping of breakpoints. Below, STS results for WHT5557 and AMC1574. Solid black bars encompass STSs found to be present. Gray bars indicate intervals to which breakpoints were localized by STS mapping. See Supplementary Table S2 for high-resolution STS markers used to further delineate the breakpoint in WHT5557. (B) Junction sequence of WHT5557 aligned with Yq and Yp reference sequence. Dots indicate base pairs identical to the junction sequence. (C) Above, model of production of the Yq–Yp junction sequence in WHT5557 by homologous recombination at a 153-bp segment of perfect homology. Below, schematic of pseudoisoy chromosome in WHT5557 formed by inter-sister-chromatid crossing-over between IR4 repeats.

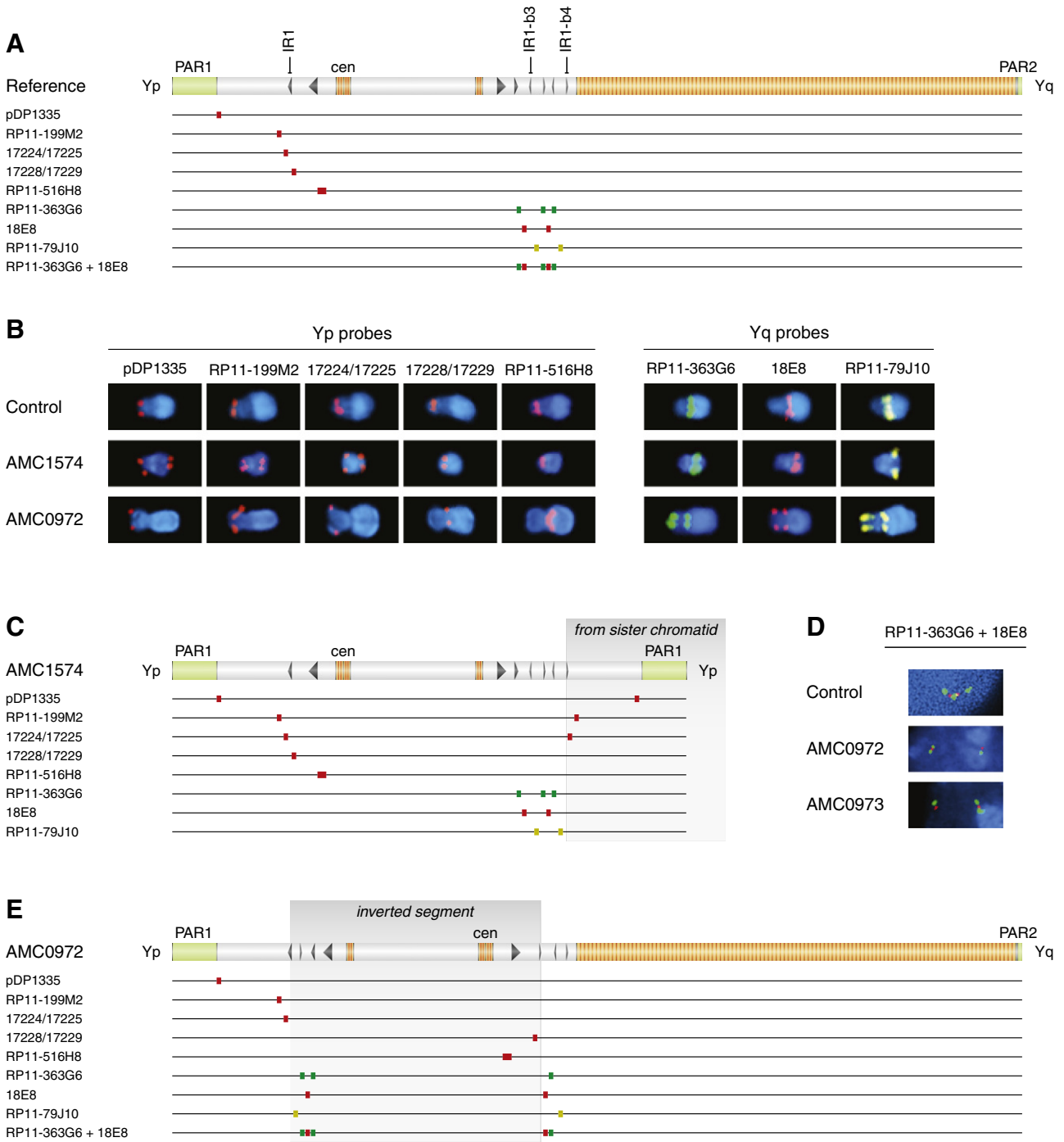
numbers of green, red, and yellow amplicons, indicating that the recombination event was distal to these amplicons, in IR1–b4 (Supplementary Table S3). In sum, these results indicate that the pseudoYp isochromosome in AMC1574 was formed by an inter-sister-chromatid recombination event between the IR1–b4 amplicon on Yq and the IR1 amplicon on Yp (Fig. 4C).

#### 2.4. Y-chromosome pericentric inversions

Because DNA material is not deleted in an inversion, STS screening was not useful for characterizing recombination events for the two patients whose karyotypes had revealed pericentric inversions in the Y chromosome. For these cases, we relied on FISH to determine

the copy number and location of several Y-chromosome loci and then to infer the recombination event that had generated each inversion. In contrast to the pseudoisoy chromosomes described above, in which the large Yq heterochromatid repeat is deleted, Y chromosomes with pericentric inversions retain this chromosomal landmark, which is clearly visible when metaphase spreads are counterstained with DAPI (Fig. 4B). This enabled us to assign metaphase FISH signals to either Yp or Yq, or to both arms.

When we assayed Yp FISH probes on metaphase spreads of AMC0972, our results indicated that the IR1 amplicon on Yp had been targeted (Fig. 4B). pDP1335, 17224/17225, and RP11-199M2, three probes that are normally distal to the Yp amplicon of IR1, remained only on Yp. In contrast, 17228/17729 and RP11-516H8,



**Fig. 4.** Characterization of Y-chromosome structures by FISH. (A) Probe hybridization sites shown below a schematic of the reference Y chromosome. (B) Metaphase FISH assays on a control sample, AMC1574, and AMC0972 demonstrate the copy number and chromosome-arm location of hybridized probes. The resolution of metaphase FISH cannot distinguish signals of closely spaced hybridization sites. (C) Schematic of pseudoisoy chromosome in AMC1574 formed by inter-sister-chromatid crossing-over between IR1 on Yp and IR1-b4 on Yq. Predicted FISH probe hybridization sites are shown. (D) Two-color interphase FISH demonstrates the rearrangement of *AZFc* amplicons in two cases with pericentric inversions. (E) Schematic of pericentric Y-chromosome inversion in AMC0972 and predicted FISH probe hybridization sites. See Supplementary Fig. S1 for two-step model: intrachromatid *gr/rg* inversion in *AZFc* followed by intrachromatid recombination between IR1 on Yp and IR1-b3 on Yq.

two probes located proximal to IR1 on Yp, hybridized only to Yq. Next, we assayed FISH probes normally in the *AZFc* region on Yq. In AMC0972, we found copies of the green, red, and yellow amplicons both on Yp and on Yq, suggesting that the inverted segment included at least one copy of each amplicon and that the Yq inversion breakpoint was located within the *AZFc* region. Indeed, two-color FISH on interphase spreads showed a green–red–green amplicon cluster clearly

separated from a red–green amplicon cluster, in contrast to a control sample where these amplicons were arranged in a single cluster with a green–red–green–red–green organization (Fig. 4D). When we performed two-color interphase FISH with probe pDP1335 to distal Yp and a probe to the green amplicon, we found that two of the three green amplicons were located on Yp in AMC0972 (data not shown). The most likely explanation for the metaphase and interphase

FISH results involves a two-step model: 1) gr/rg inversion within *AZFc* (a previously described polymorphism [9–11]), followed by 2) intrachromatid recombination between IR1 on Yp and IR1-b3 on Yq (Fig. 4E and Supplementary Fig. S1).

For AMC0973, single-color FISH on metaphase spreads showed that green, yellow and red amplicons were present on Yp and on Yq, indicating that – similar to our findings for AMC0972 – the inversion breakpoint was located within the *AZFc* region (data not shown). Two-color FISH on interphase spreads showed a red–green amplicon organization in one cluster separated from a green–red–green amplicon organization in a second cluster (Fig. 4D). In contrast to AMC0972, only a single green amplicon was located on Yp (data not shown). These results suggest that the pericentric inversion in AMC0973 was the result of intrachromatid recombination between IR1 on Yp and IR1-g2 on Yq. Insufficient cells were available for additional FISH experiments for this case.

### 2.5. Catalog of IR1 and IR4 inverted repeat pairs

Having identified four unique recombination events, including one that had evidently formed on a Y chromosome carrying an inversion polymorphism, we sought to generate a more extensive catalog of Yp–Yq inverted repeats by accounting for known inversion polymorphisms. The IR1 and IR4 amplicons on Yp and all IR1 amplicons on Yq are located in regions known to be subject to large-scale inversion (Supplementary Fig. S2). On Yp, a 3.6-Mb region containing the IR1 and IR4 amplicons and bounded by the IR3 inverted repeats has been inverted repeatedly during human history [1,11,28–31]. On Yq, the orientations of segments that contain IR1 amplicons can be polymorphic due to the frequently occurring b2/b3 and gr/rg inversions [9–11]. Thus, additional inverted repeat pairs potentially exist among extant human Y chromosomes. Taking these relatively common inversions into account, we schematized six dot plots – the two IR3 orientations plotted against each of the three *AZFc* orientations – and identified a total of 18 inverted repeat pairs involving IR1 or IR4. Each of the 18 inverted repeats is a putative substrate for the proposed model of inter-arm homologous recombination (Supplementary Fig. S2 and Table 1).

To facilitate future molecular characterization of non-reference Y chromosomes, we determined the Y-chromosome structures that could arise by our proposed model at these 18 inverted repeats (Supplementary Fig. S3). For each of these 18 repeat pairs, sister-chromatid crossing-over would generate a pseudoisoyp chromosome and a pseudoisoyq chromosome, whereas intrachromatid crossing-over would produce a pericentric inversion in one chromatid. Homologous recombination between IR1 pairs or between IR4 pairs could produce a total of 50 different non-reference Y-chromosome structures, which can be distinguished by the series of FISH probes reported here.

### 3. Discussion

We and others have shown in previous studies that nearly all amplicons located on the same Y-chromosome arm – in direct or inverted (palindromic) orientation – are substrates for ectopic homologous recombination [3–13]. Recently, we proposed an inclusive model that unites these findings as well as evidence of sequence homogenization of palindromic repeats on primate Y chromosomes [13]. In our model, a double-strand break (DSB) formed in one repeat is resolved by inter-sister-chromatid or intrachromatid repair using the other repeat as a template. Repair with crossing-over alters the structure of the Y chromosome, resulting in deletions, duplications, inversions, or isodicentric chromosomes. Alternatively, DSB repair involving noncrossover resolution leads to gene conversion, homogenizing the repeat sequences and thus maintaining opportunities for further ectopic recombination events.

The findings we report here indicate that our model extends to inverted repeats located on opposing arms of the human Y chromosome. We identified among individuals in our collections four unrelated men with non-reference Y-chromosome structures that appear to have arisen via ectopic homologous recombination between such inverted repeat pairs. Two men harbored pseudoisoyp chromosomes that apparently had formed by sister-chromatid crossing-over. In two other men, intrachromatid crossing-over apparently had generated Y-chromosome pericentric inversions. Furthermore, analysis of the reference sequence indicates that such inverted repeat pairs are also undergoing sequence homogenization. For example, the copies of IR1 on Yp and Yq share 99.66% sequence identity (0.34% sequence divergence) over 62.4 kb, an average of one difference per 295 nucleotides (Table 1). However, the amplicons include a 6142-nucleotide interval of 100% identity, which is highly unlikely without sequence homogenization. Similarly, the IR4 amplicons contain a 590-nucleotide interval of 100% sequence identity, despite overall sequence divergence of >6% over 303.2 kb (in both cases  $p < 0.0001$ , computer simulation). Thus, inverted repeat pairs located on opposing Y-chromosome arms appear to undergo NAHR-mediated sequence homogenization, which renders them susceptible to further ectopic recombination.

Whereas pericentric inversions do not impact genomic content and thus should not influence spermatogenesis, recombination events that result in pseudoisoyp chromosome formation can disrupt spermatogenesis through loss of some distal Yq DNA. Indeed, both men with pseudoisoyp chromosomes displayed severe spermatogenic failure: WHT5557 and AMC1574 were diagnosed with non-obstructive azoospermia (no sperm in semen) and severe oligozoospermia (total sperm count  $< 5 \times 10^6$ ), respectively. The difference in severity may be explained by the fact that WHT5557 lacked the entire *AZFc* region, a region essential for sperm production [6], whereas no protein-coding genes were deleted in AMC1574 (Supplementary Table S4). Reduced spermatogenesis in AMC1574 may be attributable to germline instability of the abnormal Y chromosome, as we have hypothesized previously for isodicentric Y chromosomes [13]. Alternatively, absence of the Yq pseudoautosomal region (PAR2), which can recombine during meiosis with its homologous counterpart on the X chromosome, may disrupt X–Y chromosome pairing, thereby impairing spermatogenesis [32]. In contrast and as expected, both individuals with pericentric inversions were phenotypically normal men with normal fertility.

We do not know whether the pseudoisoyp chromosomes or pericentric Y inversions reported here arose *de novo* or were inherited, because we lack requisite material (cells or genomic DNA) from patrilineal relatives. However, previous studies of individuals with pericentric Y inversions, whose frequency is estimated at 1–2 per 1000 males, have found the inversions in male relatives, indicating that such inversions are heritable and thus compatible with fertility [17–24]. Indeed, a pericentric Y inversion of a single origin was found at high frequency (~30%) among a South African population whose ancestors originated from the Indian state of Gujarat [23,24].

Of a total of 50 different non-reference Y chromosome structures predicted to arise from 18 intrachromosomal homologous recombination events between IR1 repeats or between IR4 repeats (Supplementary Fig. S3), we identified four: two pseudoisoyp chromosomes (between IR4 repeats or IR1/IR1-b4 repeats) and two pericentric Y inversions (between IR1/IR1-b3 repeats or IR1/IR1-g2 repeats). A recent report classified pericentric Y inversions found in nine unrelated men into three types, based on the Yq inversion breakpoints determined by FISH [26]. For each type, the FISH results are compatible with intrachromatid recombination between IR1 amplicons or between IR4 amplicons. It is possible that all 50 predicted structures exist in the human population, but the chance of their ascertainment is expected to vary. For example, Y-chromosome pericentric inversions do not influence spermatogenesis (nor, presumably, other phenotypes), thus

their detection depends on additional analysis of worldwide structural polymorphism [11] or on fortuitous cytological identification. In contrast, pseudoisoy chromosomes are expected to impair spermatogenesis. Thus, such chromosomes will be identified by karyotype analysis during the clinical work-up of men seeking fertility treatment. Similarly, an individual carrying a pseudoisoyq chromosome is expected to be anatomically female since *SRY*, normally located in distal Yp, is absent, and she may present with Turner syndrome stigmata.

The relative likelihood of identifying the 50 predicted structures and the isodicentric Y chromosomes we described previously [13] also depends on their relative rates of formation. In all cases, the Y chromosome must experience a DSB in an amplicon, which must then interact and recombine with a non-allelic counterpart. For pseudoisochromosomes and pericentric inversions to form, Yp and Yq – on sister chromatids or within one chromatid – must enter into proximity. On the other hand, isodicentric Y chromosomes form by crossing-over between inverted repeats located in similar positions on sister chromatids; such repeats likely remain, after replication and through to chromosome segregation, in closer physical proximity than amplicon pairs separated by the centromere. Thus, the rates at which different repeat pairs interact and recombine to form non-reference Y chromosome structures are almost certainly affected by chromosome dynamics that could sterically hinder or favor some interactions over others.

In conclusion, we have found that in addition to generating isodicentric Y chromosomes and Y chromosomes that carry deletions, duplications, or inversions, intrachromosomal homologous recombination can also generate pseudoisoy chromosomes and Y-chromosome pericentric inversions. The catalog of all possible MSY structures resulting from homologous recombination between IR1 or IR4 copies as well as the methods used to detect them will facilitate the molecular classification of all aberrant Y-chromosome structures. These methods include 1) single-locus-specific PCR-based assays that discriminate between sequence variants in sequence-similar regions and, when feasible, lead to junction amplification; and 2) FISH assays that enable the characterization of recombination events involving large amplicons with high sequence identity.

## 4. Materials and methods

### 4.1. Dot plots

Dot plots were generated using the program *self\_dot\_plot.pl* (available at <http://pagelab.wi.mit.edu/research>) [1]. To preclude showing matches caused by short interspersed repeats or by low-complexity repetitive sequences, such repeat elements were masked using RepeatMasker version open-4.0.1 (<http://www.repeatmasker.org>).

### 4.2. Cases

The men studied originated from sample collections of the Whitehead Institute, Cambridge; Onze Lieve Vrouwe Gasthuis, Amsterdam; the Maastricht University Medical Center; and the Academic Medical Center, Amsterdam. They were selected based on their karyotype data that indicated a pseudoisoyp chromosome or a pericentric inversion. Karyotyping of these men was performed as part of either fertility workup or prenatal screening. Genomic DNA was extracted from venous blood and leukocytes were isolated for FISH analysis. This study was approved by the institutional review boards of the Massachusetts Institute of Technology and the Academic Medical Center. Informed consent was obtained from all participants.

### 4.3. STS deletion screening

All four men were screened for deletions using plus/minus PCR for the following low-resolution STSs: sY142, sY1197, sY1191, sY1291, sY1206, and sY1201, as described previously [6,10]. WHT5557 was

additionally screened for deletions using MSY Breakpoint Mapper (<http://breakpointmapper.wi.mit.edu>) [27]. STSs and their GenBank accession numbers are listed in Supplementary Table S2.

### 4.4. Amplification and sequencing of WHT5557 junction sequence

The junction sequence in WHT5557 was PCR-amplified using primer 19760 (AGTGAGCCCAGTTTGCATG) specific to IR4 on Yq and primer 19761 (GGATAGCCTGAAATATAGGCAAATAT) specific to IR4 on Yp. The resulting PCR product was sequenced on an ABI3730 automated sequencer using the PCR primers and the BigDye Terminator protocol (Applied Biosystems). See GenBank accession number EU747132 for junction sequence.

### 4.5. Fluorescence in situ hybridization

Metaphase and interphase nuclei were hybridized, as previously described [13], with the following FISH probes: plasmid pDPI335 (*SRY* on distal Yp), BAC RP11-199M2 (*AMELY*), long-range PCR product 17224/17225 (immediately distal to IR1 on Yp), long-range PCR product 17228/17229 (immediately proximal to IR1 on Yp), BAC RP11-516H8 (*TSPY* array on proximal Yp), BAC RP11-363G6 (green amplicon in *AZFc*), cosmid 18E8 (red amplicon in *AZFc*), and BAC RP11-79J10 (yellow amplicon in *AZFc*). Long-range PCR product 17224/17225 was amplified with primers 17224 (GGTGTATGTGTGCA TGGATTCTGCTTG) and 17225 (AGCAGGTAGGCTTCATCAGTTGTGG TTG) using BAC RP11-305H21 as template. 17228/17229 was amplified with primers 17228 (AACTGGAATAGTGTTCCTGGGGCTGAA) and 17229 (TCTGGCCAGTGTATGGGGCTTATTAAC) using BAC RP11-109F19 as template. Long-range PCR was performed using Advantage 2 Taq polymerase (BD Biosciences) according to the manufacturer's instructions. Each primer was at 1  $\mu$ M final concentration. For a 100  $\mu$ l reaction, template DNA was 50 ng of extracted BAC DNA. Amplification conditions were 95 °C for 1 min; 30 cycles of 95 °C for 30 s, 68 °C for 10 min; 68 °C for 10 min.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.ygeno.2013.04.018>.

## Acknowledgments

We thank Lia Knegt for karyotype analysis, Laura Brown for patient data curation, and Jennifer Hughes for valuable comments on the manuscript. This work was supported by the National Institutes of Health, the Howard Hughes Medical Institute, the Netherlands Organization for Scientific Research, and the Academic Medical Center of the University of Amsterdam.

## References

- [1] H. Skaletsky, T. Kuroda-Kawaguchi, P.J. Minx, H.S. Cordum, L. Hillier, L.G. Brown, S. Repping, T. Pyntikova, J. Ali, T. Bieri, A. Chinwalla, A. Delehaunty, K. Delehaunty, H. Du, G. Fewell, L. Fulton, R. Fulton, T. Graves, S.F. Hou, P. Latrielle, S. Leonard, E. Mardis, R. Maupin, J. McPherson, T. Miner, W. Nash, C. Nguyen, P. Ozersky, K. Pepin, S. Rock, T. Rohlfling, K. Scott, B. Schultz, C. Strong, A. Tin-Wollam, S.P. Yang, R.H. Waterston, R.K. Wilson, S. Rozen, D.C. Page, The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes, *Nature* 423 (2003) 825–837.
- [2] S. Rozen, H. Skaletsky, J.D. Marszalek, P.J. Minx, H.S. Cordum, R.H. Waterston, R.K. Wilson, D.C. Page, Abundant gene conversion between arms of palindromes in human and ape Y chromosomes, *Nature* 423 (2003) 873–876.
- [3] C. Sun, H. Skaletsky, S. Rozen, J. Gromoll, E. Nieschlag, R. Oates, D.C. Page, Deletion of azoospermia factor a (*AZFa*) region of human Y chromosome caused by recombination between *HERV15* proviruses, *Hum. Mol. Genet.* 9 (2000) 2291–2296.
- [4] C. Kamp, P. Hirschmann, H. Voss, K. Huellen, P.H. Vogt, Two long homologous retroviral sequence blocks in proximal Yq11 cause *AZFa* microdeletions as a result of intrachromosomal recombination events, *Hum. Mol. Genet.* 9 (2000) 2563–2572.
- [5] P. Blanco, M. Shlumukova, C.A. Sargent, M.A. Jobling, N. Affara, M.E. Hurles, Divergent outcomes of intrachromosomal recombination on the human Y chromosome: male infertility and recurrent polymorphism, *J. Med. Genet.* 37 (2000) 752–758.

- [6] T. Kuroda-Kawaguchi, H. Skaletsky, L.G. Brown, P.J. Minx, H.S. Cordum, R.H. Waterston, R.K. Wilson, S. Silber, R. Oates, S. Rozen, D.C. Page, The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men, *Nat. Genet.* 29 (2001) 279–286.
- [7] S. Repping, H. Skaletsky, J. Lange, S. Silber, F. Van Der Veen, R.D. Oates, D.C. Page, S. Rozen, Recombination between palindromes P5 and P1 on the human Y chromosome causes massive deletions and spermatogenic failure, *Am. J. Hum. Genet.* 71 (2002) 906–922.
- [8] S. Repping, H. Skaletsky, L. Brown, S.K. van Daalen, C.M. Korver, T. Pyntikova, T. Kuroda-Kawaguchi, J.W. de Vries, R.D. Oates, S. Silber, F. van der Veen, D.C. Page, S. Rozen, Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection, *Nat. Genet.* 35 (2003) 247–251.
- [9] S. Fernandes, S. Paracchini, L.H. Meyer, G. Florida, C. Tyler-Smith, P.H. Vogt, A large AZFc deletion removes DAZ3/DAZ4 and nearby genes from men in Y haplogroup N, *Am. J. Hum. Genet.* 74 (2004) 180–187.
- [10] S. Repping, S.K. van Daalen, C.M. Korver, L.G. Brown, J.D. Marszalek, J. Gianotten, R.D. Oates, S. Silber, F. van der Veen, D.C. Page, S. Rozen, A family of human Y chromosomes has dispersed throughout northern Eurasia despite a 1.8-Mb deletion in the azoospermia factor c region, *Genomics* 83 (2004) 1046–1052.
- [11] S. Repping, S.K. van Daalen, L.G. Brown, C.M. Korver, J. Lange, J.D. Marszalek, T. Pyntikova, F. van der Veen, H. Skaletsky, D.C. Page, S. Rozen, High mutation rates have driven extensive structural polymorphism among human Y chromosomes, *Nat. Genet.* 38 (2006) 463–467.
- [12] M.A. Jobling, I.C. Lo, D.J. Turner, G.R. Bowden, A.C. Lee, Y. Xue, D. Carvalho-Silva, M.E. Hurles, S.M. Adams, Y.M. Chang, T. Kraaijenbrink, J. Henke, G. Guanti, B. McKeown, R.A. van Oorschot, R.J. Mitchell, P. de Knijff, C. Tyler-Smith, E.J. Parkin, Structural variation on the short arm of the human Y chromosome: recurrent multigene deletions encompassing amelogenin Y, *Hum. Mol. Genet.* 16 (2007) 307–316.
- [13] J. Lange, H. Skaletsky, S.K. van Daalen, S.L. Embry, C.M. Korver, L.G. Brown, R.D. Oates, S. Silber, S. Repping, D.C. Page, Isodicentric Y chromosomes and sex disorders as byproducts of homologous recombination that maintains palindromes, *Cell* 138 (2009) 855–869.
- [14] C.M. Carvalho, F. Zhang, J.R. Lupski, Structural variation of the human genome: mechanisms, assays, and role in male infertility, *Syst. Biol. Reprod. Med.* 57 (2011) 3–16.
- [15] P.A. Jacobs, M. Brunton, W.M. Brown, Cytogenetic studies in leucocytes on the general population: subjects of ages 65 years and more, *Ann. Hum. Genet.* 27 (1964) 353–365.
- [16] P.A. Jacobs, A. Ross, Structural abnormalities of the Y chromosome in man, *Nature* 210 (1966) 352–354.
- [17] R.S. Sparkes, H.M. Muller, I.C. Veomett, Inherited pericentric inversion of a human Y chromosome in trisomic Down's syndrome, *J. Med. Genet.* 7 (1970) 59–62.
- [18] U. Friedrich, J. Nielsen, Chromosome studies in 5049 consecutive newborn children, *Clin. Genet.* 4 (1973) 333–343.
- [19] E. Zeuthen, J. Nielsen, Pericentric Y inversion in the general population, *Humangenetik* 19 (1973) 265–270.
- [20] R.S. Verma, J. Rodriguez, H. Dosik, The clinical significance of pericentric inversion of the human Y chromosome: a rare “third” type of heteromorphism, *J. Hered.* 73 (1982) 236–238.
- [21] A. Toth, M. Gaal, J. Laszlo, Familial pericentric inversion of the Y chromosome, *Ann. Genet.* 27 (1984) 60–61.
- [22] L.R. Shapiro, R.O. Pettersen, P.L. Wilmot, D. Warburton, P.A. Benn, L.Y. Hsu, Pericentric inversion of the Y chromosome and prenatal diagnosis, *Prenat. Diagn.* 4 (1984) 463–465.
- [23] R. Bernstein, A. Wade, J. Rosendorff, A. Wessels, T. Jenkins, Inverted Y chromosome polymorphism in the Gujerati Muslim Indian population of South Africa, *Hum. Genet.* 74 (1986) 223–229.
- [24] A. Spurdle, T. Jenkins, The inverted Y-chromosome polymorphism in the Gujarati Muslim Indian population of South Africa has a single origin, *Hum. Hered.* 42 (1992) 330–332.
- [25] J.J. Pasantes, S. Rottger, W. Schempp, Part of the RBM gene cluster is located distally to the DAZ gene cluster in human Yq11.23, *Chromosome Res.* 5 (1997) 537–540.
- [26] S. Knebel, J.J. Pasantes, D.A. Thi, F. Schaller, W. Schempp, Heterogeneity of pericentric inversions of the human Y chromosome, *Cytogenet. Genome Res.* 132 (2011) 219–226.
- [27] J. Lange, H. Skaletsky, G.W. Bell, D.C. Page, MSY Breakpoint Mapper, a database of sequence-tagged sites useful in defining naturally occurring deletions in the human Y chromosome, *Nucleic Acids Res.* 36 (2008) D809–D814.
- [28] D.C. Page, Sex reversal: deletion mapping the male-determining function of the human Y chromosome, *Cold Spring Harbor Symp. Quant. Biol.* 51 (Pt. 1) (1986) 229–235.
- [29] N.A. Affara, M.A. Ferguson-Smith, J. Tolmie, K. Kwok, M. Mitchell, D. Jamieson, A. Cooke, L. Florentin, Variable transfer of Y-specific sequences in XX males, *Nucleic Acids Res.* 14 (1986) 5375–5387.
- [30] M.A. Jobling, G.A. Williams, G.A. Schiebel, G.A. Pandya, G.A. McElreavey, G.A. Salas, G.A. Rappold, N.A. Affara, C. Tyler-Smith, A selective difference between human Y-chromosomal DNA haplotypes, *Curr. Biol.* 8 (1998) 1391–1394.
- [31] C.A. Tilford, T. Kuroda-Kawaguchi, H. Skaletsky, S. Rozen, L.G. Brown, M. Rosenberg, J.D. McPherson, K. Wylie, M. Sekhon, T.A. Kucaba, R.H. Waterston, D.C. Page, A physical map of the human Y chromosome, *Nature* 409 (2001) 943–945.
- [32] T.K. Mohandas, R.M. Speed, M.B. Passage, P.H. Yen, A.C. Chandley, L.J. Shapiro, Role of the pseudoautosomal region in sex-chromosome pairing during male meiosis: meiotic studies in a man with a deletion of distal Xp, *Am. J. Hum. Genet.* 51 (1992) 526–533.