Dear Old Dad

Paternal age and the origin of spontaneous mutations in humans

Rivka L. Glaser and Ethylin Wang Jabs

(21 January 2004)

The origin and frequency of spontaneous mutations that occur with age in humans have been a topic of intense discussion. The mechanisms by which spontaneous mutations arise depend on the parental germ line in which a mutation occurs. In general, paternal mutations are more likely than maternal mutations to be base substitutions. This is likely due to the larger number of germ cell divisions in spermatogenesis than in oogenesis. Maternal mutations are more often chromosomal abnormalities. Advanced parental age seems to influence some mutations, although it is not a factor in the creation of others. In this review, we focus on patterns of paternal bias and age dependence of mutations in different genetic disorders, and the various mechanisms by which these mutations arise. We also discuss recent data on age and the frequency of these mutations in the human male germ line and the impact of these data on this field of research.

Introduction

The paternal age effect, a concept that took nearly a century to define, refers to the increased incidence of sporadic genetic disorders in offspring born to men of older age. The first observation of such a relationship between age and the incidence of a disorder was in 1912, when Wilhelm Weinberg noticed that children with achondroplasia, an inherited skeletal disorder, were more likely to be born later in sibships. However, at that time he could not distinguish the effects of birth order, paternal age, and maternal age (1-4). Three decades later, E.T. Mørch refined Weinberg’s initial observation and noted that parental age, not birth order, was responsible for the increased incidence of achondroplasia (2, 5, 6). The distinction between the influences of paternal and maternal ages was made nearly half a century after Weinberg’s observation: Penrose (7) observed that paternal age correlated with the incidence of de novo achondroplasia but that there was no independent correlation of maternal age with incidence.

Since Weinberg’s initial observation, approximately 20 autosomal dominant disorders have been reported to be associated with advanced paternal age (Fig. 1). Risch et al. (8) analyzed the effect of parental age on de novo cases of these disorders by comparing the distribution of parental ages of affected children for each disorder to the distribution of parental ages in the general population. The relation between parental age and the ratio of observed-to-expected (O/E) affected children for each disorder was fit to both exponential and linear models in order to determine the strength of the parental age effect.

A strong paternal age effect is evident in some disorders, such as achondroplasia and Apert, Crouzon, and Pfeiffer syndromes, in which the birth frequency of affected individuals increases rapidly with paternal age (Fig. 2A). For these disorders, the data are compatible with an exponential model of increase in the relative frequency of affected children (O/E) with increasing paternal age; the linear model is not compatible with the data. In these disorders, the average age of fathers of affected children is 5 to 7 years older than the average age of fathers in the general population. However, for other disorders, such as neurofibromatosis, a much weaker paternal age effect is seen. The rate of increase in the frequency of sporadic cases of this disorder (O/E) is much less marked, and no distinction between a linear or exponential model of increase can be made, because the data fit both models equally (Fig. 2B). In these disorders, the average age of fathers of affected children is 2 to 5 years older than the average age of fathers in the general population.

Clearly, parental age, especially paternal age, influences the development of spontaneous mutations more in some disorders than in others. What characterizes the mutations associated with these disorders? What types of spontaneous mutations are associated with paternal age and how do they arise?

Penrose’s Copy Error Hypothesis

Penrose (9) suggested that one possible cause of spontaneous mutation was DNA copy errors in chromosomal replication during mitotic cell division. These mutations, which occur during DNA replication, would most likely be base substitutions that would arise from the misincorporation of nucleotides by DNA polymerase. According to Penrose’s hypothesis, such copy errors would occur preferentially in the paternal germ line, as opposed to the maternal germ line, because of the larger number of cell divisions in spermatogenesis than in oogenesis. Furthermore, such mutations would be dependent on paternal age, as the total number of spermatogonial cell divisions increases with age.

Chromosomal Replication

In human males, gonocytes divide approximately 30 times before puberty to give rise to the spermatogonial stem cells (2) (see Walter Perspective*). During spermatogenesis, which begins at puberty, these stem cells divide every 16 days, or 23 times per year (10). The daughter cells undergo five additional divisions while developing into primary spermatocytes (11, 12). Each primary spermatocyte then enters the meiotic phase, ultimately producing four haploid spermatids, which will differentiate into mature spermatozoa. Assuming that the rate of spermatogonial stem cell division does not decrease with age, the number of cell divisions in a 30-, 40-, and 50-year-old man is 380, 610, and 840, respectively (2, 3, 13) (See “Farewell to Fatherhood†”).

The process of oogenesis, however, is quite different from that of spermatogenesis. Approximately 24 cell divisions occur,
and these are completed before birth, producing all the oocytes that will be present in a female throughout her lifetime (2). Thus, in contrast to the linear increase in total spermatogonial cell divisions with age, the total number of cell divisions in oogenesis remains constant. As the difference in the number of chromosomal replications between males and females increases with age, the proportion of paternally derived base substitutions would be expected to increase as well.

**Parental Bias in Base Substitutions**

In order to determine the parental origin of a mutation, the maternal and paternal alleles in the affected child must be distinguished from one another. The development of molecular techniques such as allele-specific polymerase chain reaction and allele-specific oligonucleotide hybridization has made it possible to use single-nucleotide polymorphisms (SNPs) near the mutations for these disorders to identify the parental alleles in the child and thereby establish the parental origin of the mutation.

Using such a strategy, several studies have demonstrated the exclusive paternal origin of point mutations in the following one-third of the disorders identified by Risch et al. as having a strong paternal age effect: achondroplasia (14) and Apert (15), Crouzon (16), Pfeiffer (16), and progeria syndromes (17) (Fig. 3). Advanced paternal age among the unaffected fathers of children with these disorders was also shown. Two other disorders not mentioned by Risch et al. (8), multiple endocrine neoplasia type 2A (MEN 2A) and medullary thyroid carcinoma (MTC), also

---

**Fig. 1.** Disorders associated with advanced paternal age [adapted from (8)]. Boldface type indicates disorders with a reported preferential paternal origin of mutation. The parental origin of mutation has not yet been determined in the remaining disorders. aMIM, Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM). bDisorders not mentioned in (8). cFrom (21). dFrom (22).
are caused by base substitutions that are exclusively paternal in origin and associated with advanced paternal age (18-21).

In contrast to these disorders for which only paternally derived mutations have been observed, mutations in MEN 2B are occasionally maternal in origin. Initially, advanced paternal age and exclusively paternal origin of one recurrent point mutation was shown in 25 sporadic cases (22). Shortly thereafter, the same mutation in two other sporadic cases was shown to be maternally derived, with no association with paternal age (23). Therefore, disorders with an exclusively paternal origin of mutation may represent the extreme end of a spectrum in which maternally derived mutations have yet to be reported.

Mutations in Disorders with a Strong Paternal Age Effect

Paternally derived base substitutions. In these eight disorders, molecular evidence shows an overwhelmingly paternal origin of base substitutions (160 out of 162, or 98.8%). What characterizes these mutations? In most of these eight disorders with a strong paternal age effect, there is a high prevalence of mutations at one or two loci in each gene, possibly due to these loci being mutational hotspots, bias of ascertainment, or selection for those mutations that cause a phenotype.

In achondroplasia, 98% of cases are caused by one mutation in the fibroblast growth factor receptor 3 gene (FGFR3): a G-to-A transition at nucleotide 1138 (1138G>A). The remaining cases result from a G-to-C transversion at the same position, or, rarely, a G-to-T transversion at nucleotide 1123 (24-27). Both mutations at nucleotide 1138 occur in a CpG dinucleotide. The germline frequencies of these mutations have been estimated to be 5.5 x 10^-6 to 28 x 10^-6 (26). Whereas only one mutation is responsible for the vast majority of cases of achondroplasia, more than 99% of all sporadic cases of Apert syndrome are caused by one of two mutations in the FGFR2 gene, 755C>G or 758C>G. The 755C>G mutation occurs in a CpG dinucleotide and has a higher mutation rate than the 758C>G mutation (5.0 x 10^-6 versus 2.7 x 10^-6), which does not occur in a CpG dinucleotide (15). Interestingly, the 755C>G mutation is a transversion, whereas most mutations in CpG dinucleotides are either C-to-T or G-to-A transitions. More than 90% of cases of progeria also are caused by a recurrent mutation within a CpG dinucleotide: a C-to-T transition at nucleotide 2036 of the lamin A gene (LMNA) [see “Lamin-tation” (http://sageke.sciencemag.org/cgi/content/full/2003/16/nw59)] (17, 28). Recurrent missense mutations in exons 11 and 16 of the RET protooncogene cause 85, 30, and 95% of all cases of MEN 2A, MTC, and MEN 2B, respectively (18, 29-33). However, unlike mutations in FGFR2, FGFR3, and LMNA, none of the recurrent mutations in RET occur in a CpG dinucleotide.

Recurrent mutations are found in Crouzon and Pfeiffer syndromes as well, but they occur at a much lower frequency than those in achondroplasia and Apert syndrome. These mutations, like those found in MEN 2A and 2B, do not occur within a CpG dinucleotide, suggesting that more than one mechanism gives rise to paternally derived mutations with a paternal age effect. The majority of the point mutations in Crouzon and Pfeiffer syndromes are found in FGFR2, the same gene mutated in Apert syndrome (27). But unlike in Apert syndrome and achondroplasia, where only one or two mutations cause the majority of cases, many different mutations have been observed in these two syndromes. Over 25 and 30 different mutations have been reported in Pfeiffer and Crouzon syndromes, respectively (27).

Mutations with unknown parental origin. The parental origin of mutations in the remaining disorders in the group with a strong paternal age effect has yet to be established, because either the gene is not known, as in acrodysostosis and fibrodysplasia ossificans progressiva, or there are many different mutations, as in basal cell nevus, cleidocranial dysostosis, Marfan, and

<table>
<thead>
<tr>
<th>Disorder</th>
<th>MIM #</th>
<th>No. of paternal cases (%)</th>
<th>No. of maternal cases (%)</th>
<th>a b</th>
<th>Paternal age effect</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autosomal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achondroplasia</td>
<td>100800</td>
<td>40 (100)</td>
<td>0</td>
<td>∞</td>
<td>Yes</td>
<td>14</td>
</tr>
<tr>
<td>Apert</td>
<td>101200</td>
<td>57 (100)</td>
<td>0</td>
<td>∞</td>
<td>Yes</td>
<td>15</td>
</tr>
<tr>
<td>CMT 1A c</td>
<td>118220</td>
<td>1</td>
<td>0</td>
<td>∞</td>
<td>Unknown</td>
<td>70</td>
</tr>
<tr>
<td>Hirschsprung</td>
<td>142623</td>
<td>0</td>
<td>3 (100)</td>
<td>0</td>
<td>Unknown</td>
<td>71</td>
</tr>
<tr>
<td>MEN 2A, MTC</td>
<td>171400</td>
<td>11 (100)</td>
<td>0</td>
<td>∞</td>
<td>Yes</td>
<td>18-21</td>
</tr>
<tr>
<td>MEN 2B</td>
<td>162300</td>
<td>25 (93)</td>
<td>2 (7)</td>
<td>12.5</td>
<td>Yes</td>
<td>22, 23</td>
</tr>
<tr>
<td>Neurofibromatosis 1</td>
<td>162200</td>
<td>33 (89)</td>
<td>4 (11)</td>
<td>8.25</td>
<td>No</td>
<td>72-75</td>
</tr>
<tr>
<td>Neurofibromatosis 2</td>
<td>101000</td>
<td>13 (57)</td>
<td>10 (42)</td>
<td>1.3</td>
<td>Unknown</td>
<td>36</td>
</tr>
<tr>
<td>Progeria</td>
<td>176670</td>
<td>5 (100)</td>
<td>0</td>
<td>∞</td>
<td>Yes</td>
<td>17</td>
</tr>
<tr>
<td>Tuberous sclerosis</td>
<td>191092</td>
<td>2</td>
<td>3</td>
<td>.67</td>
<td>No</td>
<td>76</td>
</tr>
<tr>
<td>von Hippel-Lindau</td>
<td>193300</td>
<td>4 (57)</td>
<td>3 (43)</td>
<td>1.3</td>
<td>No</td>
<td>77</td>
</tr>
<tr>
<td>X-linked</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelizaeus-Merzbucher</td>
<td>312080</td>
<td>4 (80)</td>
<td>1 (20)</td>
<td>4</td>
<td>No</td>
<td>78</td>
</tr>
<tr>
<td>Rett syndrome</td>
<td>312750</td>
<td>29 (91)</td>
<td>3 (9)</td>
<td>9.7</td>
<td>No</td>
<td>42-44</td>
</tr>
</tbody>
</table>

α for autosomal disorders 8.5
α for X-linked disorders 8.25
α for all disorders 8.5
α for disorders with a paternal age effect 69

Fig. 3. Parent-of-origin and parental age effects in base substitutions in dominant disorders [adapted from (38)]. aMIM, Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM). bRatio of the number of paternally derived mutations to maternally derived mutations for each disorder. cCMT 1A, Charcot-Marie-Tooth syndrome type 1A.
oculodentodigital dysplasia (ODDD), and Waardenburg syndromes. Among the latter group of disorders, a range of mutation types—base substitutions, deletions, insertions, and duplications—are found with fairly equal frequency (34). Although a strong paternal age effect has been observed for all of these disorders, it remains to be seen whether these different types of mutations are of paternal origin.

**Mutations Weakly Associated with Advanced Paternal Age**

Bilateral retinoblastoma, multiple exostoses, neurofibromatosis (NF1), Sotos syndrome, and Treacher Collins syndrome were shown by Risch et al. to have a weak paternal age effect (Figs. 1 and 2B). The mutations associated with these disorders are less homogeneous with regard to type and parental origin than are mutations associated with disorders that have a strong paternal age effect and an almost exclusively paternal origin of mutation.

Base substitutions, as well as deletions, insertions, rearrangements, and translocations, are found with approximately equal frequency in these disorders (34). Large deletions in bilateral retinoblastoma and Sotos syndrome are more often paternally derived than maternally derived (Fig. 4). In contrast, small deletions in Treacher Collins syndrome do not show a preference for parental origin (Fig. 5) (35). In NF1, 89% of base substitutions are paternal in origin, whereas only 20% of large deletions are maternal in origin (see citations in Figs. 3, 4, and 5). Conversely, in NF2, base substitutions do not show a bias of parental origin, whereas deletions do; 90% of deletions are paternal in origin (Figs. 3 and 5) (36). The facts that ratios of paternally derived mutations to maternally derived mutations differ depending on mutation type and that some mutations are not strongly associated with parental age argue for the existence of more than one mechanism underlying the origin of these mutations.

**Mutation Mechanisms**

**Higher frequency of base substitutions in males.** In his study of X-linked hemophilia, Haldane suggested that a higher mutation frequency exists in the germ cells of human males compared to females (37). He based this hypothesis on his observation that affected males were more likely to be born to carrier mothers (heterozygous) than to unaffected mothers (homozygous), meaning that the mutation occurred in an earlier generation. Although other studies have found a higher mutation frequency in human males and in males of other species, there is disagreement about how much higher the mutation frequency is in males as compared with in females [reviewed in (38)]. This is an important issue because the answer will provide information about the mutational mechanism(s) of base substitutions: Are DNA copy errors the primary source of mutations or are other mechanisms involved?

The magnitude of the ratio of mutation frequency in males to mutation frequency in females (ε) is calculated by determining the ratio of the number of paternally derived mutations to maternally derived mutations for each disorder. 

<table>
<thead>
<tr>
<th>Disorder</th>
<th>MIM#</th>
<th>No. of paternal cases (%)</th>
<th>No. of maternal cases (%)</th>
<th>α</th>
<th>Paternal age effect</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autosomal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilateral retinoblastoma</td>
<td>180200</td>
<td>17 (89)</td>
<td>2 (11)</td>
<td>8</td>
<td>No</td>
<td>79-82</td>
</tr>
<tr>
<td>Cri-du-chat</td>
<td>123450</td>
<td>85 (87)</td>
<td>13 (13)</td>
<td>6.5</td>
<td>Unknown</td>
<td>83-85</td>
</tr>
<tr>
<td>Del 1p36</td>
<td>18 (35)</td>
<td>34 (65)</td>
<td>.53</td>
<td>Unknown</td>
<td>86, 87</td>
<td></td>
</tr>
<tr>
<td>Del 9p</td>
<td>9 (70)</td>
<td>4 (30)</td>
<td>2.25</td>
<td>Unknown</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Del 14</td>
<td>5 (71)</td>
<td>2 (29)</td>
<td>2.5</td>
<td>Unknown</td>
<td>89-93</td>
<td></td>
</tr>
<tr>
<td>Del 22q11</td>
<td>192430</td>
<td>65 (50)</td>
<td>153 (70)</td>
<td>.42</td>
<td>Unknown</td>
<td>94-101</td>
</tr>
<tr>
<td>217095</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>188400</td>
</tr>
<tr>
<td>Hirschsprung</td>
<td>142623</td>
<td>1</td>
<td>0</td>
<td>∞</td>
<td>Unknown</td>
<td>71</td>
</tr>
<tr>
<td>HNPPa</td>
<td>162500</td>
<td>3 (43)</td>
<td>4 (57)</td>
<td>.75</td>
<td>Unknown</td>
<td>102-106</td>
</tr>
<tr>
<td>Miller-Dieker syndrome</td>
<td>247200</td>
<td>12 (71)</td>
<td>5 (29)</td>
<td>2.4</td>
<td>Unknown</td>
<td>107-109</td>
</tr>
<tr>
<td>Neurofibromatosis 1</td>
<td>162200</td>
<td>18 (20)</td>
<td>74 (80)</td>
<td>.24</td>
<td>No</td>
<td>75, 110-113</td>
</tr>
<tr>
<td>Sotos syndrome</td>
<td>117550</td>
<td>18 (90)</td>
<td>2 (10)</td>
<td>9</td>
<td>Unknown</td>
<td>114</td>
</tr>
<tr>
<td>Spinal muscular atrophy</td>
<td>7 (88)</td>
<td>1 (12)</td>
<td>7</td>
<td>No</td>
<td>115-117</td>
<td></td>
</tr>
<tr>
<td>TRPS IIb</td>
<td>150230</td>
<td>6 (67)</td>
<td>3 (33)</td>
<td>2</td>
<td>Unknown</td>
<td>118, 119</td>
</tr>
<tr>
<td>Tuberculous sclerosis</td>
<td>191092</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>No</td>
<td>76</td>
</tr>
<tr>
<td>von Hippel Lindau</td>
<td>193300</td>
<td>3 (50)</td>
<td>3 (50)</td>
<td>1</td>
<td>No</td>
<td>77</td>
</tr>
<tr>
<td>Williams</td>
<td>194050</td>
<td>93 (44)</td>
<td>118 (56)</td>
<td>.79</td>
<td>No</td>
<td>120</td>
</tr>
<tr>
<td>Wolf-Hirschhorn/ Pitt-Roger-Danks</td>
<td>194190</td>
<td>27 (87)</td>
<td>4 (13)</td>
<td>6.75</td>
<td>No</td>
<td>121</td>
</tr>
<tr>
<td>X-linked</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duchenne and Becker muscular dystrophies</td>
<td>310200</td>
<td>5 (14)</td>
<td>31 (86)</td>
<td>.25</td>
<td>Unknown</td>
<td>122</td>
</tr>
<tr>
<td>300376</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rett</td>
<td>312750</td>
<td>1</td>
<td>0</td>
<td>∞</td>
<td>No</td>
<td>43</td>
</tr>
</tbody>
</table>

α for autosomal disorders: .93
α for X-linked disorders: .26
α for all disorders: .89

Fig. 4. Parent-of-origin and paternal age effects in large deletions in dominant disorders. aMIM, Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM). bRatio of the number of paternally derived mutations to maternally derived mutations for each disorder. cHNPP, hereditary neuropathy with liability to pressure palsies. dTRPS, trichorhinophalangeal syndrome type II. Also known as Langer-Giedion syndrome.
of these disorders. As discussed earlier, point mutations in several disorders were found to be overwhelmingly paternal in origin and associated with older age. As shown in Fig. 3, these disorders have the highest $\alpha$ (>12.5), because there were virtually no maternally derived cases observed. Therefore, there appears to be a good correlation between disorders identified by Risch et al. as having a strong paternal age effect and disorders with the highest $\alpha$. Is this the case for base substitutions in other disorders? Surveying the literature for all reported cases of known parental origin of base substitutions revealed that 89.5% (222 out of 248 mutations) were paternally derived, and for all disorders $\alpha$ was calculated to be 8.5 (Fig. 3). Although point mutations in general may be more prone to occurring in the male germ line, it should be noted that $\alpha$ is close to 1 (there is no parent-of-origin effect) in three disorders: NF2, tuberous sclerosis, and von Hippel-Lindau disease. However, in tuberous sclerosis and von Hippel-Lindau disease, the estimated $\alpha$ may not be correct because of the small number of cases reported for each of these disorders. There were no disorders in which there was a significantly greater number of maternally derived base substitutions. Three maternally and no paternally derived mutations were found for Hirschsprung disease, but nothing definitive can be said with such a small sample size.

This higher prevalence of mutations in males may explain the observed lack of affected males and reduced fitness of females with any of several X-linked dominant disorders once thought to be lethal in males (39). For these disorders, the higher frequency of mutation in males would result in paternally derived de novo mutations on the X chromosome, which would manifest only in heterozygous (affected) females, because of the lack of X chromosome transmission from father to son. In the absence of maternally derived mutations, affected males would be born only to affected mothers, but because of the reduced fitness of such females, familial cases are very rare.

The role of methylation. Differences in the methylation status of the DNA in male and female germ cells may play a role in the observed higher frequency of mutation in males than in females. CpG dinucleotides are considered hypermutable because in most of these dinucleotides the cytosine is methylated, rendering it prone to undergoing spontaneous deamination to thymine. In humans and mice, sperm DNA has been shown to be more methylated than oocyte DNA, and this has been suggested to account for the greater number of paternally derived than maternally derived point mutations occurring within a CpG dinucleotide (40, 41). Indeed, the ratio of paternally derived to maternally derived base substitutions in a CpG dinucleotide is 1.5 to 2 times as high as the ratio for base substitutions at all locations (Fig. 6). This mechanism could explain paternally derived base substitutions in CpG sites with no age effect, as seen in Rett syndrome (42-44). However, this mechanism would not explain the association of some CpG mutations with older paternal age, as seen in achondroplasia, Apert syndrome, and progeria. Few studies have examined the relation between age and the extent of DNA methylation in sperm as an explanation for the paternal age effect seen in these disorders (45). One such study found that methylation was stable with age at nucleotide 1138 of FGFR3, the locus implicated in most cases of achondroplasia. The existence of CpG mutations that are mostly paternal in origin and not associated with older paternal age, such as those in Rett syndrome, suggests that the mechanisms that generate base substitutions with both parent-of-origin and paternal age effects differ from the mechanisms that generate base substitutions with only a parent-of-origin effect.

Mutations arising during meiosis. In contrast to base substitutions, large deletions (more than 20 base pairs) do not demonstrate a strong parent-of-origin effect (Fig. 4). Analysis of all reported cases of large deletions of known parental origin shows that 392 out of 847, or 46%, are paternal in origin, with an $\alpha$ of 0.86. This type of mutation can be generated by unequal recombination between repeated sequences or repetitive sequence elements (for example, Alu sequences) or, in some cases, nonhomologous sequences in the genome. Recombination during meiosis has been postulated to account for many of the deletions listed in Fig. 4. Because a limited number of meiotic divisions occur in both spermatogenesis and oogenesis, mutations arising in this stage of germ cell development would not be expected to show either a parent-of-origin or parental age effect (4, 46). This same mechanism may generate other chromosomal abnormalities such as large insertions and duplications (46), for which we would not expect to see parent-of-origin or parental age effects.

Overall, large deletions appear not to have a parent-of-origin effect, but a clear parent-of-origin effect is observed in some disorders associated with large deletions. For example, in bilateral retinoblastoma and cri-du-chat, Sotos, and Wolf-Hirschhorn syn-

<table>
<thead>
<tr>
<th>Disorder</th>
<th>MIM # a</th>
<th>No. of paternal cases (%)</th>
<th>No. of maternal cases (%)</th>
<th>$\alpha$ b</th>
<th>Paternal age effect</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hirschspring</td>
<td>142623</td>
<td>2</td>
<td>0</td>
<td>$\infty$</td>
<td>Unknown</td>
<td>71</td>
</tr>
<tr>
<td>Neurofibromatosis 2</td>
<td>101000</td>
<td>8 (89)</td>
<td>1 (11)</td>
<td>8</td>
<td>Unknown</td>
<td>36</td>
</tr>
<tr>
<td>Rett syndrome</td>
<td>312750</td>
<td>2</td>
<td>0</td>
<td>$\infty$</td>
<td>No</td>
<td>43, 44</td>
</tr>
<tr>
<td>Treacher Collins</td>
<td>154500</td>
<td>6 (67)</td>
<td>3 (33)</td>
<td>2</td>
<td>No</td>
<td>35</td>
</tr>
<tr>
<td>Tuberous sclerosis</td>
<td>191092</td>
<td>3 (60)</td>
<td>2 (40)</td>
<td>1.5</td>
<td>No</td>
<td>76</td>
</tr>
</tbody>
</table>

$\alpha$ for all disorders 3.33

Fig. 5. Parent-of-origin and parental age effects in small deletions (fewer than 20 base pairs) in dominant disorders. aMIM, Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM). bRatio of the number of paternally derived mutations to maternally derived mutations for each disorder.
dromes, 87 to 90% of the deletions are paternal in origin (Fig. 4). On the other hand, 65 to 85% of deletions in del 1p36 and del 22q11 syndromes, NF1, and Duchenne and Becker muscular dystrophies are maternal in origin (Fig. 4). Although a parent-of-origin effect is observed for deletions associated with some disorders, a strong paternal age effect has yet to be reported for deletions. These observations contrast with what is seen for point mutations; namely, a strong paternal age effect and a lack of disorders having mostly maternally derived base substitutions (Fig. 3). The nature of the repeated sequences that mediate deletions may influence the parental origin of the deletions, because some sequences may be predisposed to recombination and deletion during oogenesis, whereas other deletions may be detrimental to not only mature sperm, but to other stages of spermatogonial development (47, 48).

When a variety of mutations, such as point mutations and deletions, are observed in one disorder, as is the case in NF1, a paternal age effect for one type of mutation (point mutations) may be diluted by the presence of other types of mutations (deletions) that are not associated with age (4). It would be interesting to see whether a strong paternal age effect is observed for only the point mutations in these disorders.

In contrast to gross deletions, small deletions (fewer than 20 base pairs) are thought to arise during replication by mispairing and misalignment of direct repeats or short runs of identical bases (49). This type of mutation tends to be paternal in origin (α = 3.33), similar to what is observed for point mutations (Fig. 5). This type of mutation may occur during replication, which may explain the higher mutation frequency seen in males as compared with females. However, small deletions are not associated with older paternal age, whereas point mutations, which also are thought to arise during replication, are.

**Discrepancy between the frequency of mutations predicted by the copy error hypothesis and the frequency of children born with disorders with a strong paternal age effect.** A number of different mechanisms have been mentioned that could give rise to mutations associated with a parent-of-origin effect alone or in conjunction with a paternal age effect. One such mechanism, copy error, predicts a paternal origin of base substitutions associated with older paternal age. The linear increase in the number of chromosomal replications in sperm with age alone is expected to produce a linear increase in the number of mutations with age. However, the increase in the frequency of affected children born to older fathers actually is exponential (2, 4, 8) (Fig. 7A). This suggests that mechanisms in addition to replication errors contribute to the paternal age effect. Such mechanisms may be the decreased efficiency and/or fidelity of DNA replication.

<table>
<thead>
<tr>
<th>Disorder</th>
<th>MIM #</th>
<th>No. of paternal cases (%)</th>
<th>No. of maternal cases (%)</th>
<th>α</th>
<th>Paternal age effect</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achondroplasia</td>
<td>100800</td>
<td>40 (100)</td>
<td>0</td>
<td>∞</td>
<td>Yes</td>
<td>14</td>
</tr>
<tr>
<td>Apert syndrome</td>
<td>101200</td>
<td>388 (100)</td>
<td>0</td>
<td>∞</td>
<td>Yes</td>
<td>15</td>
</tr>
<tr>
<td>Neurofibromatosis 2</td>
<td>101000</td>
<td>4 (50)</td>
<td>4 (50)</td>
<td>1</td>
<td>No</td>
<td>36</td>
</tr>
<tr>
<td>Rett syndrome</td>
<td>312750</td>
<td>26 (90)</td>
<td>3 (10)</td>
<td>8.7</td>
<td>No</td>
<td>42-44</td>
</tr>
<tr>
<td>von Hippel-Lindau syndrome</td>
<td>193300</td>
<td>2</td>
<td>0</td>
<td>∞</td>
<td>No</td>
<td>77</td>
</tr>
</tbody>
</table>

α for all disorders except Apert syndrome = 12.7

**Fig. 6.** Parent-of-origin and parental age effects in CpG mutations in dominant disorders. aMIM, Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM). bRatio of the number of paternally derived mutations to maternally derived mutations for each disorder. cThe original paper did not state how many of the 57 paternally derived mutations occurred in a CpG dinucleotide. However, we assumed that this mutation was present in 67% of these 57 cases (38 cases) because two-thirds of live born children with Apert syndrome have the 755C>G mutation. dAlthough occurring in a CpG dinucleotide, the 755C>G mutation is a transversion and therefore probably not the result of spontaneous deamination.

**Fig. 7.** Direct measurement of mutations in sperm to assess a potential mechanism for the paternal age effect. (A) Comparison of the fold increase in the number of chromosomal replications (solid squares) to the relative frequency of children born with Apert syndrome (open squares) with paternal age. Data are from (8). Comparison of the sperm mutation frequency (diamonds) and birth frequency in achondroplasia (open circles) (B) and Apert syndrome (open squares) (C). Graph in (B) adapted from (45). Data in (C) are from (52).
polymerases and proteins involved in DNA repair pathways or decreased functional apoptotic mechanisms with age (4, 50). These mechanisms, in conjunction with the mechanisms generating the base substitutions, may result in the exponential accumulation of mutations with age.

**Direct measurement of mutations in sperm to assess a potential mechanism for the paternal age effect.** Base substitutions are thought to occur mainly in spermatogonial stem cells, which divide throughout life. Regardless of the mechanism by which these mutations are generated, the increase in incidence of disorders with a paternal age effect should be a reflection of the increasing number of mutations present in sperm as men age. Screening the entire genome for base substitutions, each present at a very low frequency, would be ideal for determining overall mutation frequency. But no existing technology can feasibly do so. However, samples can be screened for the presence of one or two recurrent mutations associated with older paternal age and an exclusively paternal origin. These criteria are met by the mutations most commonly found in achondroplasia and Apert syndrome.

Achondroplasia was the model disorder in the first of three studies determining the age-related frequency of base substitutions in sperm (45). Surprisingly, the slight age-related increase in the frequency of the most common mutation in achondroplasia did not correspond to the age-related increase in the incidence of the disorder (Fig. 7B). Furthermore, large variances in the frequency of mutations were observed in sperm from donors over 60 years old.

Two subsequent studies examined the age-related frequency of mutations in Apert syndrome. The first study, done in our lab, analyzed the frequency of the two most common transversion mutations (755C>G and 758C>G) in Apert syndrome in sperm (51). Using a different experimental technique than that used in the achondroplasia study, we found that these mutations also increased in frequency in sperm as donors’ age increased. As in the achondroplasia study, this increase in frequency occurred much later in life than expected and did not correspond in timing with the age-related increase in incidence of offspring with the disorder (Fig. 7C). Also as in the achondroplasia study, large variances were seen in mutation frequency in older men. Blood samples from a subset of men were analyzed for both mutations, but no increase in mutation frequency was found with age.

The third study used yet another technique and assayed for the three possible mutations at the most commonly mutated locus in Apert syndrome, nucleotide 755 (C-to-T, C-to-A, and C-to-G) (52). The C-to-G mutation at nucleotide 758 was not analyzed. An age-related increase in the frequency of the most common mutation, a C-to-G transversion at nucleotide 755, as well as the much rarer C-to-T transition at the same locus, was observed in sperm. No age-related increase in the frequency of the three mutations was observed in white blood cells. Two results in this study contrasted with observations from the other two studies. First, the variance among individuals in the oldest age groups was smaller than the variance among younger individuals. Second, mutation frequency in sperm corresponded well with the frequency of live-born offspring with Apert syndrome. Differences in study design, sample recruitment, sample number, and sensitivity of the experimental techniques could have contributed to these differences.

In all three studies, sperm DNA from fathers of affected children were also analyzed in the same manner as sperm from the other donors to address the hypothesis that fathers of affected children may have different mutation frequencies. Of the four fathers of children with achondroplasia who were studied, two had mutation frequencies similar to those of age-matched controls, one had a mutation frequency that was twice as high as in age-matched controls, and the oldest father had an overall mutation frequency that was lower than that of his age-matched controls. As this was such a small sample size, no conclusions could be drawn. Of the six fathers of children with Apert syndrome that Goriely et al. analyzed, one (approximately 45 years old) had a mutation frequency similar to that observed in older men who had not fathered an affected child (52). This finding is similar to our observation that 3 out of 15 fathers, all younger than 45 years old, had elevated mutation frequencies, similar to those observed in 60- to 80-year-old men who had not fathered affected children (51). We therefore thought that this group of men may be a distinct subgroup with either a higher degree of mosaicism in their sperm cells, possibly resulting from mutations occurring early in gametogenesis, or a predisposition to mutations, as compared to fathers of unaffected children. Unfortunately, blood samples were not available from these fathers; therefore, we could not determine whether they were somatic or germinal mosaics. However, germline mosaicism in Apert syndrome and achondroplasia is rare (33, 54). Alternatively, these fathers may have been more prone to spontaneous mutations. Additional studies on larger groups of fathers of affected children are necessary to confirm these observations. Thorough genetic and epidemiologic studies of fathers of children with disorders for which paternal age effects have been observed may help to identify environmental factors that may influence mutation rate.

It has been suggested that FGFR genes, specifically FGFR2 and FGFR3, may be predisposed to mutational events and that the high frequencies of several of these nucleotide changes may be due to a selective advantage of the mutated germ cell (15, 55). Both Goriely et al. and our lab found indirect evidence for selection (51, 52). Of the two common mutations in Apert syndrome, the 755C>G mutation is found in 67% of all cases of Apert syndrome, whereas the 758C>G mutation is found in approximately 32% of all cases. Therefore, we were surprised to find these two mutations present at similar frequencies in sperm, suggesting selection for the 755C>G mutation or selection against the 758C>G mutation. Goriely et al. found that the 755C>G mutation was more prevalent in sperm than the C-to-T or a C-to-A mutations at the same locus (52). Because the cytosine of this nucleotide is methylated and occurs in a CpG dinucleotide, we would expect a mutation at this locus to be a C-to-T transition caused by spontaneous deamination. However, the less common C-to-G transversion was found to be more prevalent in sperm. The phase of these mutations with respect to a G/A SNP approximately 100 base pairs upstream of the mutations was determined in the sperm of men who were heterozygous for this SNP. In these men, the C-to-G and C-to-T mutations did not arise preferentially on one allele. However, marked variation was seen between men. In some men, the C-to-G mutation appeared more often on the G allele, whereas in other men the mutation appeared more often on the A allele. These authors contend that these findings can both be explained by positive selection of the C-to-G mutation in sperm.

How might these mutations be advantageous to sperm? FGFRs have been localized to the developing testes and seem to be important in maintaining spermatogenesis (56-58). Dimer-
ization of the receptors upon ligand binding results in autophosphorylation of the intracellular tyrosine kinase domains (59). This phosphorylation triggers downstream signaling pathways, resulting in changes in gene expression and other biological consequences. Phosphorylation of tyrosine kinases is important in sperm motility and capacitation (60, 61). Gain-of-function mutations in these genes may confer some selective advantage on the sperm in terms of motility and capacitation, through constitutive or altered activation of the FGF receptors. In fact, all the disorders with a strong paternal age effect in which an almost exclusively paternal origin of mutation has been demonstrated are caused by gain-of-function mutations in genes encoding for receptor tyrosine kinases (FGFR2, FGFR3, and RET) (62-66).

Selection may also occur after fertilization, during implantation. Penrose (7) originally rejected the idea that maternal age had any effect on the incidence of achondroplasia, but it was later found that a maternal age contribution to the incidence of disorders with a strong paternal age effect could not be discounted (8). Subsequent molecular studies showing an exclusively paternal origin of mutations and advanced paternal age in these disorders once again shifted the focus back to the effect of paternal age. Perhaps embryos with these mutations are more compatible with the hormonal environment in older women.

It is intriguing that two studies found that mutation frequency in sperm increased with a man's age, but mutation frequency in white blood cells did not, suggesting cell type-specific selection. Both cell types arise from populations of rapidly regenerating stem cells, so why would mutations accumulate in only one cell type? The age-related accumulation of mutations in several genes in lymphocytes is well documented (67-69). However, those studies analyzed the age-related increases in deletions as well as point mutations, whereas only point mutations were analyzed by Gorile et al. and us. Perhaps only some mutation types accumulate with age. An alternative explanation is that the mutations in Apert syndrome may be selected against in sperm, resulting in a decrease of these mutations in sperm to decrease with age, with a concomitant increase in DNA damage (50).

Closing Remarks

The paternal age effect is much more complex than originally thought. Base substitutions in CpG dinucleotides, as well as those not in CpG dinucleotides, are associated with older paternal age. However, not all mutations occurring in CpG dinucleotides are associated with older paternal age, suggesting a number of different mechanisms at work. There is no simple relation between the number of chromosomal replications and the number of mutations in the male germ line. The three recent studies on the frequency of mutations in the male germ line all agreed that the frequency of mutations in sperm increases with age. However, there is disagreement over how these mutations increase with age—does the increase correspond to the increase in frequency of affected children born to older fathers, or is there a discrepancy in timing between the two? These differences highlight the need for future studies in this field. Two of these studies provide evidence for a relative selective advantage of pathogenic mutations to the sperm. We think this would be interesting for future studies to concentrate on how these mutations affect sperm. In addition, determining whether fathers of affected children have different mutation rates has important consequences for genetic counseling.

References

25. R. Shiang, L. M. Thompson, Y. Z. Zhu, D. M. Church, T. J. Fielder, M. Bo-


9. H. J. Ludecke, R. Burdick, G. Senger, U. Clausen, E. Passarge, B. Hors-


123. The authors thank J. Crow, G. Thomas, and B. Migeon for helpful discussions and R. Schulman and J. Cai for help in preparing the manuscript. This article is based on R.L.G.’s doctoral dissertation, “The role of paternal origin and frequency of spontaneous mutations in the paternal age effect.”