Alu in Lynch Syndrome: A Danger SINE?

Megan P. Hitchins¹ and John Burn²

Abstract

Lynch syndrome is a hereditary cancer predisposition syndrome caused by germline loss of a DNA mismatch repair gene. In a significant proportion of cases, loss of function of the MSH2 mismatch repair gene is caused by large heterogeneous deletions involving MSH2 and/or the adjacent EPCAM gene. These deletions usually result from homologous recombination events between Alu elements, a family of short interspersed nuclear elements (SINE). Recent recognition that the extent of these deletions influences phenotypic outcome provided new impetus for fine-mapping the breakpoints. In doing so, Pérez-Cabornero and colleagues uncovered new evidence for Alu-mediated ancestral founder deletions within MSH2 in the Spanish Lynch syndrome population (as reported beginning on pages 1546 and 1556 in this issue of the journal). This is the first such finding to date and prompted a revisitation of the role of Alu elements in the causation of Lynch syndrome. Whether Alu density is a danger sign for genomic regions prone to rearrangement and what additional factors may be required to actuate these events remain to be discovered. Cancer Prev Res; 4(10); 1527–30. ©2011 AACR.

Lynch syndrome, the most common hereditary cancer syndrome, is caused by heterozygous loss of a DNA mismatch repair gene. These loss-of-function mutations are most frequently reported in the mutant L homolog 1, colon cancer, and nonpolyposis type 2 Escherichia coli (MLH1) gene (accounting for 50% of all mismatch repair mutations in Lynch syndrome) and in the mutant S homolog 2, colon cancer, and nonpolyposis type 1 E. coli gene (MSH2; 40%), followed by the mutant S homolog 6 and E. coli gene (MSH6; 10%) and the postmeiotic segregation increased 2 Saccharomyces cerevisiae gene (PMS2; rare). These germline mutations predispose carriers to the early development of mismatch repair–deficient (typically microsatellite unstable) cancers, particularly colorectal and endometrial (1). Identification of the pathogenic mutation in afflicted families is crucial for their clinical management, determining which relatives are carriers and thus in need of regular clinical surveillance, while alleviating noncarriers from as invasive and expensive medical procedures. However, the plethora of different mismatch repair mutations that underlie Lynch syndrome cases presents a major diagnostic challenge (1).

Where available, immunohistochemistry of the mismatch repair proteins in the tumors of affected individuals can help in directing the search for the causative germline mutation, but the relevant pathogenic mutations are disseminated throughout the 4 mismatch repair genes associated with Lynch syndrome. Furthermore, large alterations affecting the structure of MLH1 and MSH2 occur in 10% to 20% of Lynch syndrome cases. These structural rearrangements largely comprise intragenic deletions encompassing at least one exon, accounting for as many as one third to half of the mutations affecting these 2 genes in certain populations (2–8). MLH1 and MSH2 duplications, insertions, and inversions have also been reported (9–11). In the case of MSH2, deletions frequently encompass the first exon and, in some cases, extend upstream into the adjacent epithelial cell adhesion molecule (EPCAM) gene as well (3, 5). The heterogeneity of mismatch repair mutations thus necessitates comprehensive genetic analysis of the relevant mismatch repair gene(s) on a case-by-case basis. Screening for point mutations by exonic sequencing, as well as dosage analysis by multiplex ligation probe amplification (MLPA) to detect structural changes within MLH1 and MSH2 (12), is now routine practice. Founder mutations have been identified in a few ethnic groups, which allow for more cost effective, directed first-pass mutation scanning in relevant populations (2, 13, 14).

In recent years, the finding that epigenetic dysregulation of the MLH1 and MSH2 genes in normal tissues can also give rise to a Lynch syndrome phenotype, in the absence of a sequence mutation within either gene, has added further complexity to the detection of the underlying molecular defect in specific cases (15, 16). Germline deletions of the 3’ end of the EPCAM gene have been identified in families with MSH2-deficient tumors, but

References

1. Pérez-Cabornero et al., p. 1546, and p. 1556

©2011 American Association for Cancer Research.

www.aacrjournals.org

Downloaded from cancerpreventionresearch.aacrjournals.org on June 20, 2017. © 2011 American Association for Cancer Research.
with a genetically intact MSH2 gene located immediately downstream. These interstitial deletions juxtapose MSH2 such that it is placed under the transcriptional regulation of the EPCAM gene, which is most active in epithelial tissues such as the colonic mucosa. Therein, loss of the EPCAM polyadenylation (transcription stop) signal results in transcriptional elongation from EPCAM into MSH2, producing fusion transcripts and ultimately resulting in MSH2 promoter methylation (17). Interestingly, because the functional impact that terminal deletions of EPCAM exert on MSH2 is confined primarily to epithelial tissues, they seem to confer a more restricted phenotype.

Carriers of deletions that are limited to the terminal portion of EPCAM are associated with a significantly reduced life-time risk of developing endometrial cancer compared with carriers of deletions that extend from EPCAM into MSH2, or with carriers of MSH2 mutations, but they have a similar life-time risk of colorectal cancer (18, 19). The recent realization of the phenotypic variance accorded by the extent of the deletions involving EPCAM has provided a renewed rationale for precisely defining their boundaries. Therefore, not only is the detection of EPCAM deletions important for a genetic diagnosis in cases with a clinical suspicion of Lynch syndrome, but genetic counseling and clinical surveillance strategies may also be influenced by the extent of the deletions.

In 2 papers published back-to-back in this issue of the journal, Pérez-Cabornero and colleagues report their comprehensive survey of a large Spanish Lynch syndrome cohort for causative mutations within MLH1, MSH2, and MSH6, as well as deletions involving EPCAM (20, 21). Collectively, they identified the pathogenic mutation in 25 of 160 index cases and describe the associated phenotypes (20, 21). Using the method of heteroduplex analysis by capillary array electrophoresis followed by direct sequencing of exons displaying altered bands, they identified point mutations, including several novel mutations, within all 3 major Lynch syndrome–associated mismatch repair genes (MLH1, MSH2, and MSH6) across 17 families (20). Eight of the index cases had MLH1 mutations, of which 3 carried the c.306+5G>A founder mutation that these investigators had previously identified in other Spanish Lynch syndrome families (13). A novel frameshift mutation within MSH6 was identified in one proband who had developed a "microsatellite-stable" endometrial cancer showing MSH6 loss, providing further evidence for the low incidence and variant molecular features of the tumors associated with mutations of this particular gene in Lynch syndrome (22).

Despite the presence of an MLH1 founder mutation, the incidence of MSH2 mutations in this cohort exceeded that of MLH1. This finding was attributable to the detection (by MLPA) of two recurring intragenic deletions within MSH2 in 7 families; one involved a deletion of exon 7 (3 families) and the other encompassed exons 4 to 8 (4 families; ref. 20). This finding was interesting because, although large deletions within MSH2 are frequently identified in Lynch syndrome, recurring deletions within this gene have not previously been described, and the families harboring the same deletion in this report were seemingly unrelated to one another.

In their second paper, Pérez-Cabornero and colleagues describe an index case with a deletion spanning the final 2 exons of EPCAM and the first 3 exons of MSH2, also initially detected by MLPA, but no deletions confined to EPCAM alone (21). Each of the 3 deletions was more precisely and efficiently mapped by a custom-designed comparative genomic hybridization (CGH) array comprising a linear set of probes located at 500-bp intervals across EPCAM and MSH2. Probes flanking each of the regions that had shown contiguous reduced dosage (indicative of the region deleted) were then employed for amplification of the junction fragments and sequence analysis across the deletion breakpoints. This series of genetic analyses revealed that each of the 3 deletions was most likely to have arisen via homologous Alu-Alu unequal recombination events between intronic Alu elements located kilobases apart (20, 21).

Alu elements are a family of short interspersed nuclear elements (SINE) found only in primates and comprise approximately 10.5% of the human genome (23). These are mobile retrotransposable elements that also contain a recombinogenic motif, priming them for recombinational activity. However, their mobility and propensity for recombination is presumably tempered by host-defensive methylation of these CpG-rich elements (23). Deletions formed through unequal Alu-mediated homologous recombination involve a cross-over at regions of shared sequence identity between 2 parental Alu elements located in cis in the same orientation, with loss of the loop of intervening genomic sequence during the exchange. These deletions are identifiable by signatory tracts of perfect Alu-derived sequence identity overlapping or adjacent to the deletion breakpoints, such that the deletion junction cannot be mapped to a precise nucleotide. The deletions reported by Pérez-Cabornero and colleagues bore the hallmarks of prior Alu-mediated recombination events. They mapped each of the deletion breakpoints to within short stretches of fused Alu sequences that shared close homology to their respective parental Alu elements.

The finding of inter-Alu–mediated deletions affecting the mismatch repair genes is not new in itself. It has been well established over the past 15 years that Alu elements are frequently responsible for genomic rearrangements in a number of genetic-based diseases. Indeed, the first observed germline deletion caused by an inter-Alu malrecombination event in any autosomal dominant familial cancer syndrome was reported in 1995 and occurred within the MLH1 gene (2). Interestingly, this particular deletion represented an ancestral founder mutation that accounted for a significant proportion of Lynch syndrome kindreds in the Finnish population (2). Subsequently, it was shown that deletions within MSH2 are frequent in Lynch syndrome (7), and characterization of their breakpoints showed that the vast majority of these are attributable to Alu-mediated malrecombination events (3–5, 7).
The mechanism by which structural alterations (primarily deletions in Lynch syndrome) arise is important because it will affect their relative frequency among mutation carriers. Deletions affecting the mismatch repair genes tend to originate via distinct mechanisms (24). Although a proportion is Alu-mediated, MLH1 deletions are mainly caused by chromosomal breakage events followed by reunion of the 2 ends, with loss of some intervening sequence (3, 24, 25). Such deletions typically have well-defined breakpoints, whereby sequence analysis will usually reveal the point of junction to precise nucleotides. Deletions of MSH2 and EPCAM, whether overlapping or confined to either singular gene, are almost invariably Alu-mediated (3-5, 7, 26). The heterogeneity found among deletions of MSH2, and now EPCAM as well, suggests that they arise from multiple independent events. These 2 neighboring genes and their intergenic sequence are littered with Alu elements, and the propensity for Alu-mediated structural changes has been correlated with their density (3, 5). Another example of a deletion-prone gene is the Alu-rich VHL gene, which underlies von Hippel–Lindau syndrome (27). However, although the high density of Alu repeats within a given region increases the likelihood of structural alterations, other factors may also play a contributing role to actuate these events. Retroelements are densely methylated, contributing to a substantial portion of the cytosine methylation in the human genome (28). The methylation status of an Alu element may determine the likelihood of the site serving as a trigger for increased recombinational activity. Supporting this notion, the evolutionarily youngest and least-methylated subfamily of Alu elements has been shown to drive chromosomal rearrangements during evolution in primates (29). Dnmt3L-deficient germ cells, in which de novo methylation of endogenous genes and retrotransposons has failed to occur, show major chromosomal alterations during meiosis (30). Furthermore, genome-wide hypomethylation is associated with widespread genomic instability in neoplasia (28).

Although it was originally hypothesized that Alu-mediated deletions affecting MSH2 might display founder effects among Lynch syndrome cases from the same geographic region or ethnic group (7), this has not transpired until now. To our knowledge, the first paper by Pérez-Cabornero and colleagues is the first report of Alu-mediated founder deletions within MSH2 (20). Of key interest in the Spanish cohort is the finding that both the intragenic exon 7 and exons 4 to 8 MSH2 deletions were recurrent in 2 sets of seemingly unrelated families, although genealogic information was limited. Deletions may recur independently because of intrinsic instability, for example, via the mechanisms discussed earlier, rendering a particular genomic segment susceptible to alteration. Alternatively, recurring deletions may occur through a founder effect; the homologue bearing the deletion is transmitted through successive generations during the expansion of isolated populations. The founder effect is demonstrable through the sharing of a common haplotype among deletion carriers due to linkage disequilibrium in the vicinity of the deletion. Pérez-Cabornero and colleagues showed that the sequences encompassing the breakpoint junction in each of the 2 deletions were identical in carriers from both sets of families. Furthermore, haplotyping at various polymorphic markers flanking MSH2 showed that each set of families shared a particular haplotype in association with their deletion. These congruent findings provided irrefutable evidence that both the MSH2 exon 7 and exons 4 to 8 deletions represent ancestral founder mutations in the Spanish Lynch syndrome population.

Pérez-Cabornero and colleagues raise the issue of anticipation in kindreds carrying the MSH2 exons 4 to 8 deletion, because disease severity in terms of cancer number, spectrum, and age of onset seemed to worsen with each successive generation. The phenomenon of anticipation, whereby the symptoms of a disease arise at a younger age and show increased severity in later generations, is exemplified by disorders associated with triplet repeat expansion such as myotonic dystrophy (31). Although initially dismissed as a consequence of ascertainment bias, this phenomenon now has a well-established molecular basis as the continued expansion in copy number of these unstable repeats during gametogenesis, with the phenotypic severity in offspring correlating with repeat length (31). However, the issue of anticipation in the context of Lynch syndrome has been controversial because there is no obvious molecular basis for it and phenotypic heterogeneity has been well documented. The issue is further confounded by generational shifts in the environment, most notably the introduction of refrigeration and the increase in consumption of red meat. It will be interesting to see whether surveillance of Lynch syndrome families in the next generation supports the anticipation theory, or whether this turns out to be a mixture of ascertainment biases and the effect of the changing environment.

It is now apparent that mapping the deletion breakpoints where the EPCAM gene is involved will be important for personalized clinical management because the extent of these deletions influence phenotypic expression (18, 19). The finding of founder mutations will facilitate future genetic screening for the pathogenic cause in pertinent populations. Furthermore, study of pedigrees carrying mutations in common will allow for the study of environmental and other potential modifiers of cancer risk in the Lynch syndrome population. With evidence of a better prognosis in mismatch repair–deficient cancers (32) and different responses to chemotherapy (33), including mutation-specific responses (34), and the prospect of targeted chemoprevention (35), this work is a valuable contribution to the molecular genetics knowledge base. The further demonstration of a direct relationship between Alu repeat density and propensity to intragenic deletion draws attention to a mechanism of general importance to those who are developing strategies for the expansion of molecular diagnostic services.
Disclosure of Potential Conflicts of Interest

The authors have no conflicts of interest to declare.

References
